Complexity vs linearity: relationships between functional traits in a protist

Nils Alain Svendsen (nils.svendsen@uclouvain.be)
Université Catholique de Louvain
Viktoriia Radchuk
Leibniz Institute for Zoo and Wildlife Research
Thibaut Morel-Journel
Center for Interdisciplinary Research in Biology
Virginie Thuillier
Université Catholique de Louvain
Nicolas Schtickzelle
Université Catholique de Louvain

Research Article

Keywords: functional traits, linearity assumption, soft/hard traits framework, Tetrahymena thermophila, trait correlations

Posted Date: June 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1741945/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Ecology and Evolution on January 11th, 2023. See the published version at https://doi.org/10.1186/s12862-022-02102-w.

Abstract

Background

The mechanisms underlying the relationship between biodiversity and ecosystem functioning are still poorly understood. Although species richness is commonly used as a biodiversity measure, recent studies showed that functional diversity, i.e. the diversity of functional traits, might be a better proxy. Functional traits are defined as phenotypic traits that affect an organism's performance and shape ecosystem-level processes. The main challenge when using those traits to quantify biodiversity is to choose which ones to measure, since effort and money are limited. As one way of dealing with this, Hodgson et al. (1999) introduced the idea of two types of traits, with soft traits that are easy and quick to quantify, and hard traits that are directly linked to ecosystem functioning but difficult to measure. If a link exists between the two types of traits, then one could use soft traits as a proxy for hard traits for a quick but meaningful assessment of biodiversity. However, this framework is based on two assumptions: (1) hard and soft traits must be tightly connected to allow reliable prediction of one using the other; (2) the relationship between traits must be monotone and linear to be detected by the most common statistical techniques (e.g. GLM, PCA).

Results

Here we addressed those two assumptions by focusing on six functional traits of the protist species *Tetrahymena thermophila*, which vary both in their measurement difficulty and functional meaningfulness. They were classified as: easy traits (morphological traits), intermediate traits (movement traits) and hard traits (oxygen consumption and population growth rate). We were able to detect a high number (> 60%) of non-linear relationships between the traits, which can explain the low number of significant relationships found using PCA and GLM analysis. In the end, these analyses did not detect any relationship strong enough to predict one trait using another, but that does not imply there are none.

Conclusions

Our results highlighted the need for more complex statistical analyzes than the ones commonly used by the scientific community, to account for all the factors that might blur the relationships between traits (e.g. plasticity, non-linearity), and state about the soft/hard framework.

Background

Biodiversity is declining at an alarming rate (1–5), requiring more than ever to be carefully measured in different ecosystems. Traditionally the focus when measuring biodiversity was on taxonomical diversity, e.g. species richness or evenness. However, such an approach has been criticized for its inability to bring

a mechanistic understanding of the effects that species composing the community have on ecosystem functioning (1, 6, 7). As an alternative, measuring functional diversity was suggested, whereby one measures the functional traits defined as characteristics of an organism's phenotype that affect its performance (8), on the one hand, and that shape ecosystem-level processes, on the other hand (9–12). Although focusing on functional diversity is appealing, quantifying it remains difficult.

The main challenge when measuring functional diversity relates to the choice of the functional traits to measure. There are too many traits to measure them all, and efforts are limited, thus usually only a subset of possible traits are measured (13, 14), often those that are rather easy to get (15, 16). However, the measured traits must properly capture the effects of an organism on ecosystem functioning and its own fitness. As one way of dealing with this, Hodgson et al. (1999) introduced the idea of soft and hard traits, where the former ones are relatively easy and quick to quantify, while the latter ones are meaningful but hard to measure. Ideally, we would measure hard traits (e.g. metabolic or physiological traits) when quantifying functional diversity, but since they are by definition difficult to measure, one could instead measure soft traits (e.g. leaf area) that are assumed to be linked to these hard traits. Such use of the soft traits as proxies for hard ones is promising but it is based on a strong assumption: hard and soft traits must be tightly connected. Another implicit assumption lurking behind the most common correlation measurement methods (e.g. Pearson correlation, PCA) is that the relationship between soft and hard traits is monotone and linear (**Fig. 1**). However, these assumptions are rarely checked.

Here we address the assumptions underlying the soft/hard framework by focusing on *Tetrahymena thermophila*, a ciliate unicellular that has widely been studied as a model system in cellular and molecular biology for more than 80 years (17–19) and in ecology and evolutionary biology for over a decade. Over these years, numerous studies provided a lot of information about *T. thermophila* metabolism (18, 20–23), reproduction (24–26), movement (27, 28) and morphology (19), allowing us to carefully assert the functional traits for this species. Based on the literature and our own existing data, we have chosen the following six functional traits of *T. thermophila* cells: two morphological traits (cell size and shape), two movement traits (movement speed and trajectory tortuosity), oxygen consumption and population growth rate. In our experimental microcosm system, these traits vary in their measurement difficulty and functional meaningfulness: easy (morphological traits), intermediate (movement traits) and hard (oxygen consumption and population growth rate).

The morphological traits are considered functional because they relate to the resource use of *T. thermophilla*. Indeed, an increase in cell size is often a consequence of resource accumulation. These resources could then be mobilized if environmental conditions become harsh. For example, when oxygen is present, the cells will produce and accumulate glycogen using a part of this oxygen (22, 23, 29); glycogen is here used as a storage for energy and can be used to produce ATP (i.e. energy needed for the cell survival) through fermentation when oxygen is lacking (20), allowing cells to survive for few hours without breathing. The shape of the cell is an indicator of wellness for *T. thermophila* (30). When the environment is stressful, *T. thermophila* cells tend to adopt a rounder shape, possibly because they exhaust all their metabolites (e.g. glycogen) in reserves to survive until the environmental conditions

become suitable again. We classify these two morphological traits as “easy” since they remained indirect proxies of glycogen accumulation or wellness, and their quantification in our system only requires a snapshot picture of cells.

The two movement traits are expected to play a major role in resource foraging, hence survival, reproduction, and dispersal strategy (31, 32). Swimming fast gives the advantage of quickly exploring spaces, allowing the cells to potentially find a better environment, at the cost of the energy needed to move and the risk of exhausting themselves to death. The same reasoning applies to the trajectory linearity, since a tortuous trajectory could enhance local foraging by maximizing resources exploited in the neighborhood, while a straight trajectory will allow access to distant patches with possible better resources and escaping harsh local conditions. We considered these two movement traits as having an “intermediate” level both regarding the measurement difficulty, since measuring them requires recording a video with trajectories of moving cells, but also functionally, since they directly impact the foraging abilities of *T. thermophila* cells.

Regarding the last group of traits, oxygen consumption is a direct proxy of the cell metabolic rate, and one of the major factors driving protist community structure (33). Population growth rate is directly proportional to the individual clonal cell reproduction rate and is the main driver of biomass production, which is often used as a proxy for ecosystem functioning or species wellness (34–39). These two traits are the most difficult to measure in our microcosm system because they cannot be measured from a snapshot data recording (picture or video), which is possible to acquire even in the field, but instead involve a time series of measurements using specialized equipment in the lab. However, they are also more directly connected to the ecological parameters of the population (i.e. metabolism and biomass production), making their estimation very desirable in functional diversity studies. Thus, according to the soft/hard framework (40), if we detect a significant relationship between these hard traits and the intermediate/easy ones, it would allow for indirect estimation of these hard traits based on snapshot picture or video measurements, which are even possible in the field.

Here, we measured the above-described six functional traits on 40 genetically distinct strains (i.e. clonally reproducing genotypes) of the protozoa *T. thermophila*, which differ in geographic origin and time since extraction from the field (Pennekamp et al., 2014). These strains were previously shown to exhibit clear differences in several life-history characteristics such as growth rate, maximum density, and survival under starvation conditions (Fjerdingstad et al., 2007; Pennekamp, 2014; Pennekamp et al., 2014); which have been demonstrated to be reliable phenotypic traits at the strain level because of the high repeatability of their measures through time (Chaine et al., 2009; Fjerdingstad et al., 2007; Pennekamp, 2014; Schtickzelle et al., 2009). The use of several strains gave us some intraspecific variation, without which it is impossible to establish trend between traits valid over the whole species.

We expect some of our chosen easy or intermediate traits to correlate with the hard traits. For example, cell shape and cell size could correlate with population growth rate as the faster a strain reproduces, the less time its cells have to accumulate resources, to become longer and larger. The oxygen consumption

rate is also expected to correlate with both movement traits and population growth rate since these processes require energy, creating a complex relationship between three traits. Further, since bigger cells may have a higher metabolism, we also expect cell size and shape to be related to oxygen consumption. However, these examples are not exhaustive, and to account for any possible pattern, we looked first of all for general trend between all the traits through a PCA. Secondly, we used GAM to look at if the predictions were improved by considering possible non-linear or monotonous relationships between the traits. Specifically, we assessed the shapes (i.e. form and standard deviation, see **Fig. 1**) of the best fit, for all possible pairwise relationships between the six traits, regardless of the difficulty of taking measurement.

Results

Among the 15 pairwise relationships between the six functional traits, 8 exhibited a higher deviance explained when permitted to be non-linear through GAMs (**Fig. 2**). Thus, considering the non-linearity significantly improved the model fit in half of the cases. However, for the other 7 relationships, the simple linear model remains the best fit. Now, if we look at the deviance explained of the best model (linear or not) for the 15 relationships, only 3 showed one above 25%: both cell size (40.5%) and cell shape (36.3%) when predicting NGDR with a non-linear model, and NGDR itself predicting population growth rate (31.2%) with a linear model (**Fig. 2**). As the D14 strain seemed to have an important leverage on some relationships (i.e. outlier), especially the ones involving cell speed, we also performed the analysis with that strain removed and still found 8 significant non-linear relationships, and similar deviance explained across all 15 best-fitted models (**See Supplementary Fig. 1, Additional File 1**). Overall, all models explained a limited proportion of the deviance, making the predictions based on a single trait quite unreliable.

From the PCA analysis on the six traits, we chose to keep the first 3 dimensions (or Principal Components), because all their eigenvalues were superior to 1, and summed up to a total of 76.6% of the original dataset inertia. According to the square cosine, the two hard traits were well represented by the first two dimensions, capturing together 89% of the variation among strains in oxygen consumption and 77% in population growth rate (**See Supplementary Table 2, Additional File 3**). For both traits, dimension 3 did not provide additional representation. On these dimensions, the other traits were represented at: 78% for NGDR, 41% for speed, 62% for cell size and only 5% for cell shape. Due to the poor representation of cell speed and shape, we could not assess their relationship with the two hard traits. These dimensions showed a weak correlation between cell size and oxygen consumption, and a strong correlation between NGDR and population growth rate (**Fig. 3.A & B**). Otherwise, the traits seemed rather independent from each other when plotted on the two first axes. To assess the relationship of cell shape and speed, we plotted dimensions 1 and 3 together to maximize their representation (**See Supplementary Table 2, Additional File 3**). On this plot, only three variables reached a square cosine higher than 50% and thus can be confidently assessed: NGDR (64%), speed (63%) and cell shape (85%). Among those, none displayed strong relationships with each other (**Fig. 3.C & D**). In conclusion, without regard to the plotted dimensions, the PCA did not show any significant relationships between three or more traits. Furthermore,

even if the data were summarized with only 3 dimensions, among that space each trait was showing little redundancy with the others, and nothing indicated that including more than one trait in the models would improve the prediction.

Discussion

The difficulty of estimating functional diversity has raised several concerns over the years (14, 57–59), mainly due to the different challenges related to it, such as measuring dozens of traits. In this study, we aimed to assess the relationship between several functional traits of a protozoa, *T. thermophila*, using a solution proposed by Hodgson et al. (1999) to deal with this challenge, the soft/hard framework. Here, one uses soft traits that are easy to measure but not always meaningful to predict hard traits that are functionally very desirable but difficult to obtain. The soft/hard framework assumes that the functional space can be reduced to a small number of traits, and so the presence of strong trade-offs or relationships among traits. This is definitely an appealing and intuitive framework that theoretically has the power to simplify the assessment of functional diversity, which might, in turn, be a key feature to understand the functioning of ecosystems. However, most of the statistical methods commonly used with this framework assume the relationships between the traits to be monotonic and linear (60), and this might be the reason for the mitigated results of the framework as of today in functional ecology studies (61, 62).

In our case study on *T. thermophila*, we detected several significant pairwise relationships between the soft and hard traits, with over 60% of them being non-linear. Thus, the non-linearity was definitely an issue, and forbid us to conclude about the presence or absence of relationships between the traits using such methods as GLM. This problem worsens with multivariate analyses like PCA, which do not only assume simple linearity between two traits, but among all of them. Those results support the idea that, despite those methods being standard when it comes to detecting relationships, one should not use them without checking the linearity first (56, 60, 63). Still, even with a method that does not assume linearity (i.e. GAM), we could not find pairwise relationships strong enough alone to allow reliable predictions of one trait using another. That absence of “good enough” relationships (i.e. which meets the points from the introduction: (i) a form that allows prediction and (ii) low standard deviation) could also be caused by a lack of statistical power.It is unlikely in our case since the analyzes were performed with a large sample size (n = 40), each data point furthermore being the average of many real replicate measures, a dataset that can be defines as of high-quality for all the traits. However, we can see another two possible causes

The first one is an absence of tight relationships between these functional traits in *T. thermophila*. In this paragraph, we will take the example of trade-offs as an illustration of phenomena that can bind traits to each other. According to Garland (2014), a trade-off exists when one trait cannot increase without a decrease in another. Such a situation can be caused by several physical and biological mechanisms, like limiting resources, allowing the improvements in one trait only at the expense of another (Alexander, 1985; Bennett & Lenski, 2007; Garland, 2014). Thus, trade-offs limit the possibilities of trait association (Bennett & Bever, 2009) and lead to a phenotype made of a whole suite of coadapted traits, that one can

define as the organism strategy (68). This idea of co-dependence between traits, which can be the results of several non-exclusive phenomena, lies in the core of the soft/hard framework, with several traits defining the strategy and function of an organism within the ecosystem.

At first sight, the case of *T. thermophila* fits this idea of strategy, with three dimensions representing more than 75% of the variability of six functional traits. However, a closer look at the PCA analysis does not show any clusters of more than two traits (**Fig. 3.A & C**), with the link between those traits remaining weak, and the strains being scattered over the multivariate space (**Fig. 3.B & D**). The GAM and GLM analyses deliver the same message, with the non-linearity sometimes improving the assessment of the link between traits, but those remain mostly weak, and again an absence of any cluster of more than two traits. For example, the models using cell size, shape and speed as predictors explained between 20 and 40% of NGDR value; however, they were not good predictors of each other, despite all being linked to the same trait (i.e. NGDR). The pattern described here is consistent with a situation where there are several different strategies (i.e. viable combination of traits) among the strains, each one not required to invest in several functional traits, but only in a few. Thus, the bond between traits, despite existing, would remain too loose to allow reliable prediction using the soft/hard framework, because of the number of possible combinations.

The second possible cause is the presence of factors that complexify the correct assessment of the relationships between traits. In this paper, we focused on tackling the problem of the non-linearity between traits, since correlation and linear model assume a simple and direct relationship (68–70). We demonstrated that taking into consideration this non-linearity can significantly improve the prediction of the model (**Fig. 2**). However, other factors can also blur the relationships between traits in the analysis, with for example the plasticity of the traits (68), their phenology (68, 71), the absence of a trait in the analysis (68), and the scale considered (71). Among those different examples, plasticity is likely to play a role in the relationship between traits for *T. thermophila*, since several traits of this species have been proven to be plastic (Morel-Journel et al., 2020; Pennekamp et al., 2014). In that case, the functional trait values will vary based on both the environment and the genetics of the strains. Thus instead of having their traits value set to only maximize the efficiency for their strategy (i.e. set of functional traits), the strains will optimize their traits value to also deal with the environment and genetic constraints. This variation of the trait values due to plasticity might then loosen the link between traits, or at least introduce a delay for the change of value in one trait to affect the other traits it is tight to. If this reasoning is correct, it introduces a notion of trait variance in the study on functional traits, which needs to be taken into account since performing analyzes only on the mean values can be misleading about the real relationship between traits.

This introduction of plasticity in the field of functional traits also advocates for considering the non-linearity, since the relationships between traits and environmental variables (i.e. reaction norms) are often non-linear (e.g. Morel-Journel et al., 2020; Rocha & Klaczko, 2012), and often follow a hump shape with an optimum, or a sigmoidal shape, reflecting the existence of environmental thresholds beyond which the trait performance changes drastically (73). These non-linear relationships might be rather common in

nature (60), as between environmental variables and demographic rates (e.g. emperor penguin adult survival and sea ice concentration (74), Euraisan Oystercatcher fecundity and resource availability (75), red kangaroo survival and rainfall (76)). It is likely that the non-linear relationships among traits are also very frequent, and assessing their prevalence empirically across species is an avenue for future research. Importantly, should the commonness of non-linear relationships among traits and environmnental variables be confirmed across a broader range of taxa and locations, this would imply that further methodological developments are needed for reliable quantification of functional diversity.

Conclusions

In the end, our study did not allow us to detect any “good enough” relationships between traits, but this does not imply there is none. It simply highlighted the need for more complex statistical analyzes than simple GLM and PCA to account for possible factors that might blur traits relationships. For example, using an analysis that accounts for several traits at the same time, the non-linearity of the relationships and the trait variation due to plasticity might greatly increase the deviance explained by models in our case. However, even in the case of a successful model working on our data, it would only be the first step for using this framework as originally intended in functional ecology. To expand those results to the whole species, one would need to consider the genetic diversity within it (61, 68), which in the case of *T. thermophila* means verifying the effectiveness of the model on as many strains as possible. In addition, one should verify if relationships between traits change in different environmental conditions, both biotic and abiotic (61, 77). Indeed, consideration of the relationships between traits under different environmental conditions has the potential of being an important bias (30), because if traits use the same resources, which would be scarce in harsh conditions, then investment choices have to be made by the organisms (Bennett & Bever, 2009; Dorken & Van Drunen, 2018). On the contrary, if there are enough resources available to invest in several traits, it might blur the relationship between those traits, making the use of the soft/hard framework more difficult. One should not assume that the relationship between the traits will stay the same across the whole range of viable environmental conditions.

Taking all these factors into account might look overwhelming at first, but it is in our opinion a primordial condition for a deeper understanding of the relationships between the traits, and their constancy in different conditions. Such an understanding is a must before we could use a trait as a predictor of some other(s) (78–80). Our study is just a preliminary step along that line, but we believe it illustrates convincingly that one should not look for simplicity at all cost (72), but more for a reasoned simplicity, which results from experiments, to identify which factors could be removed from the equation, and which should definitely be a part of it.

Methods

Culture conditions & experimental design

All strains were maintained, before and during the experiment, under standardized culture conditions that allow only clonal reproduction (46, 47): axenic liquid culture in a nutrient medium (consisting of 2% Proteose peptone and 0.2% yeast extract, diluted in ultrapure water), kept at constant 27°C temperature under a 14:10 h light/dark cycle. Stability of culture conditions is an important requirement for both the experiment and the maintenance of the cultures since *T. thermophila* shows a high degree of plasticity depending on the environmental factors (18, 28, 30). Culture stocks were renewed every seven days by inoculating a 2 mL sample of fresh medium with 20 µL of the old culture and maintained in 24-multiwell plates (ref: 662102, Greiner BioOne). All manipulations of axenic cultures were conducted under sterile conditions in a laminar flow hood (Ultrasafe 218 S, Faster).

The experimental design involved measuring the six traits of interest for the 40 strains with at least ten replicates each. The number of replicates was indeed adjusted to both the risk of losing some replicates due to bacterial contamination and the intrinsic measurement error for each trait: 18 for growth measurements (measured over five days); 10 for measuring both the morphology, movement (snapshot measures) and oxygen consumption (measured over two hours). For each strain, replicates each originated from a different mother culture, created from the stock culture at different times, ensuring they fully reflect the natural variability within each strain. To detect bacterial and fungal contamination, which can put cells under stress and alter their traits, we ran contamination tests after each measurement. These contamination tests were done by inoculating a Petri Dish containing a nutrient medium (Bacto Tryptone 2.5 g, Yeast extract 1.25 g, Glucose 0.5 g, Bacto Agar 9 g) with a few drops of the experimental cultures; contamination was detected by the presence of some fungal or bacterial development after three days at 27°C. Contaminated samples were discarded from the analyses. We first created mother cultures for each strain and replicate by putting 500 µL of the corresponding culture stock and 6 mL of fresh axenic medium in a 30 mL cell culture container (ref: 201170, Greiner BioOne). Then, each mother culture was placed horizontally to favor oxygen ingress into the medium, and thus population growth. After a 2-day growth period, the measure of the six traits was performed on each replicate according to protocols detailed below.

Trait measurement

Morphology and movement traits were measured from pictures and videos recorded under a dark field microscope by placing two 10 µL samples from each mother culture into individual chambers on counting slides (ref: 630–2048, VWR). We recorded one picture for each chamber and one 20 seconds video of one of the two chambers, randomly chosen. For all identified cells from the two pictures, cell size (its area in µm²) and shape (i.e. aspect ratio: major/minor ratio of the two axes of a fitted ellipse; minimum is 1 for round cells, > 1 for more elongated cells) were quantified (Pennekamp & Schtickzelle, 2013). From the video, we estimated cell speed (in µm/s) and trajectory tortuosity (as net to gross displacement ratio, NGDR) using the BEMOVI R-package (Pennekamp et al., 2015). NGDR is computed as a ratio of the net displacement (i.e. straight line between the starting and the ending position) over the gross displacement (i.e. the total traveled distance during the measurement), meaning NGRD is 1 for a perfectly straight trajectory and < 1 for more tortuous ones.

Oxygen consumption was measured by quantifying the decay in dissolved oxygen concentration in a sealed high-density culture. Cell abundance was first quantified in each mother culture (i.e. one replicate of a given strain) using optical density at 550 nm (Genesys 20 spectrophotometer, Thermo Fisher Scientific), to standardize the experimental cultures at 1.5 mL and 200 000 cells/mL through dilution. Then, the experimental cultures were loaded in a 24-multiwell plate containing an individual oxygen sensor in each well (Oxodish OD24, PreSens). This plate was placed on a reader device (SDR SensorDish Reader, PreSens), and its 24 wells were sealed by covering the plate (without its lid) with a silicone mat and pressing the whole system (reader, plate and mat) between two plastic plates screwed together. As soon as the plate was sealed, the oxygen stopped flowing into the well, and we started recording the concentration of dissolved oxygen (in mg/L) within the experimental cultures every two minutes for two hours. The measure of oxygen concentration using this technique is very sensitive to temperature, so the temperature was kept at precisely 27°C during all steps, from the mother cultures to the whole measurement phase, and the temperature was recorded together with each measurement of oxygen concentration as a further check. After a short lag period, the recorded oxygen concentration started a linear decrease as the cells consumed oxygen, until reaching an asymptote below 20% oxygen saturation when almost all of the oxygen was consumed. As the trait quantifying oxygen consumption, we used the slope (in mg/L*min) of the linear decrease estimated using linear regression in the R software v3.6.1 (50).

To measure population growth rate, we created 2 mL experimental cultures by diluting each mother culture (i.e. one replicate of a given strain) in its stationary phase by a factor of five, to allow exponential growth; experimental cultures were placed in 24-multiwell plates (ref: 662102, Greiner BioOne). Cell density was estimated every two hours over five days using optical density at 550 nm (Synergy H1 microplate reader with robotic plate feeder, Biotek). The growth rate μ of each experimental culture was estimated as the slope of the optical density increase over time in its linear phase using the *gcfit* function (*grofit* R package (51)).

Statistical analyses

To explore relationships between the six traits, we averaged all replicates at the strain level because we were interested in functional traits exhibited by the strains without integrating the within strain variation between the replicates. This also compensated for the unequal number of replicates for the six traits, due to experimental design and discard of contaminated replicates. We were left with 51 discarded out of 721 replicates for growth and 18 discarded out of 400 replicates for morphology/movement/oxygen consumption, with a minimum replicate per strain of 12 for the growth and 8 for the rest of the traits. All statistical analyses were performed on this set of 40 means at the strain level for each of the six traits. Morphology and movement traits, measured at the cell level, were in addition first averaged for every replicate to give the same weight to each replicate estimate without regards for the number of cells identified on pictures and videos. Then, we performed two complementary analyses on our results.

In a first analysis, we explored the existence and linearity of simple pairwise relationships (i.e. a line) between functional traits, comparing GAM (*gam* function in *mgcv* R package (52)), against a GLM (*glm* (50)). GAM are an extension of GLM adding nonparametric smooth functions to a model, which allow the

model to be non-linear (53). This allows GAM to capture common nonlinear patterns that a classic linear model would miss. Adding an extra bending point (or knot) in the GAM allows for a better fit (as adding an extra parameter in a GLM) but comes with a penalty, meaning that if the bending does not significantly improve the fit to the data (increasing the deviance explained), it will not be retained. The presence of a significant pairwise relationship in the model can be checked with its deviance explained, and the linearity with the value of the effective degree of freedom (e.d.f). Any e.d.f value above 1 indicates a significant non-linear relationship, and the higher the value, the more non-linear it is.

In a second analysis, to examine the global dimensionality and test if the predictions of the models could be improved by combining several traits together, we conducted a principal component analysis (PCA) on standardized data (*FactoMineR* (54) and *factoextra* (55) R packages). To assess the individual representativity of our traits on the dimensions, we used the squared cosine (56). A high squared cosine indicates a good representation of the variable on the dimension under consideration. If a variable is perfectly represented by just two dimensions, the sum of squared cosine on these two axes is equal to 1, thus if we plot these two dimensions, the variable will be positioned directly on the correlation circle. Since our main goal was to study the relationships between the hard and the soft traits, we used the square cosine to identify which dimensions to plot to best reflect the relationships between our traits. However, unlike the GAM, the main weakness of PCA here is its assumption of simple linear relationships between traits, hence the complementarity of the two approaches.

Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and materials: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests

Funding: The equipment used in that study was supported by an equipment grant from UCLouvain (U.N035.16). N.A.S was supported by a PhD grant from UCLouvain (FSR) and by a FRIA PhD grant. T.M.-J. was supported by a Move-In-Louvain Marie Curie Action postdoctoral fellowship and by an ERC grant (EvoComBac: 949208). V.T. was supported by a PhD grant from UCLouvain (ARC 10-15/31) and by a FRIA PhD grant. N.S. is Senior Research Associate of the F.R.S.-FNRS and acknowledges financial support from UCLouvain (ARC 18-23/095: DIVERCE). This paper is contribution BRCXXX of the Biodiversity Research Centre at UCLouvain.

Authors' contributions: N.A.S., V.R. and N.S. conceived and designed research questions and the methodology for the experiment; N.A.S. collected the data, using protocols developed by N.A.S., V.T. and N.S.; N.A.S. and V.R. analyzed the data with contributions from T.M.-J. and N.S.; N.A.S. led the writing on

the manuscript, with substantial contributions from V.R., T.M.-J. and N.S. All the authors contributed critically to the drafts and gave final approval for publication.

Acknowledgements: Not applicable

References

1. Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, et al. Biodiversity loss and its impact on humanity. Nature [Internet]. 2012 Jun 6;486(7401):59–67. Available from: http://www.nature.com/doifinder/10.1038/nature11373
2. Hooper DU, Chapin FS, Ewel JJ, Hector A, Inchausti P, Lavorel S, et al. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecol Monogr [Internet]. 2005 Feb;75(1):3–35. Available from: http://www.esajournals.org/doi/pdf/10.1890/04-0922
3. Balvanera P, Pfisterer AB, Buchmann N, He JS, Nakashizuka T, Raffaelli D, et al. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. Ecol Lett [Internet]. 2006 Oct;9(10):1146–56. Available from: http://doi.wiley.com/10.1111/j.1461-0248.2006.00963.x
4. Balvanera P, Siddique I, Dee L, Paquette A, Isbell F, Gonzalez A, et al. Linking Biodiversity and Ecosystem Services: Current Uncertainties and the Necessary Next Steps. Bioscience [Internet]. 2014 Jan 1;64(1):49–57. Available from: http://bioscience.oxfordjournals.org/cgi/doi/10.1093/biosci/bit003
5. Cardinale BJ, Srivastava DS, Emmett Duffy J, Wright JP, Downing AL, Sankaran M, et al. Effects of biodiversity on the functioning of trophic groups and ecosystems. Nature [Internet]. 2006 Oct 26;443(7114):989–92. Available from: http://www.nature.com/doifinder/10.1038/nature05202
6. Reiss J, Bridle JR, Montoya JM, Woodward G. Emerging horizons in biodiversity and ecosystem functioning research. Trends Ecol Evol [Internet]. 2009 Sep;24(9):505–14. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0169534709001803
7. Cadotte MW, Carscadden K, Mirotchnick N. Beyond species: functional diversity and the maintenance of ecological processes and services. J Appl Ecol [Internet]. 2011 Oct;48(5):1079–87. Available from: http://doi.wiley.com/10.1111/j.1365-2664.2011.02048.x
8. Violle C, Navas ML, Vile D, Kazakou E, Fortunel C, Hummel I, et al. Let the concept of trait be functional! Oikos. 2007;116(5):882–92.
9. Díaz S, Purvis A, Cornelissen JHC, Mace GM, Donoghue MJ, Ewers RM, et al. Functional traits, the phylogeny of function, and ecosystem service vulnerability. Ecol Evol [Internet]. 2013 Sep;3(9):2958–75. Available from: http://doi.wiley.com/10.1002/ece3.601
10. Luck GW, Lavorel S, Mcintyre S, Lumb K. Improving the application of vertebrate trait-based frameworks to the study of ecosystem services. J Anim Ecol. 2012;81(5):1065–76.
11. Lavorel S, Storkey J, Bardgett RD, De Bello F, Berg MP, Le Roux X, et al. A novel framework for linking functional diversity of plants with other trophic levels for the quantification of ecosystem services. J Veg Sci. 2013;24(5):942–8.

12. Lavorel S, Grigulis K. How fundamental plant functional trait relationships scale-up to trade-offs and synergies in ecosystem services. J Ecol. 2012;100(1):128–40.
13. McGill BJ, Enquist BJ, Weiher E, Westoby M. Rebuilding community ecology from functional traits. Trends Ecol Evol. 2006;21(4):178–85.
14. Nock CA, Vogt RJ, Beisner BE. Functional Traits. In: eLS [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2016. p. 1–8. Available from: http://doi.wiley.com/10.1002/9780470015902.a0026282
15. Weiher E, Werf A, Thompson K, Roderick M, Garnier E, Eriksson O. Challenging Theophrastus: A common core list of plant traits for functional ecology. J Veg Sci [Internet]. 1999 Oct 24;10(5):609–20. Available from: https://onlinelibrary.wiley.com/doi/abs/10.2307/3237076
16. Bartomeus I, Gravel D, Tylianakis JM, Aizen MA, Dickie IA, Bernard-Verdier M. A common framework for identifying linkage rules across different types of interactions. Funct Ecol. 2016;30(12):1894–903.
17. Collins K, Gorovsky MA. Tetrahymena thermophila. Curr Biol [Internet]. 2005 May;15(9):R317–8. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0960982205004392
18. Ormsbee RA. The normal growth and respiration of Tetrahymena geleii. Biol Bull [Internet]. 1942 Jun;82(3):423–37. Available from: http://www.journals.uchicago.edu/doi/10.2307/1537988
19. Collins K. Tetrahymena thermophila. [Internet]. Methods in. Collins K, editor. New York: Academic Press; 2012. 1–452 p. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780123859679000074
20. Manners DJ, Ryley JF. Studies on the metabolism of the Protozoa. 2. The glycogen of the ciliate Tetrahymena pyriformis (Glaucoma piriformis). Biochem J. 1952;52(3):480–2.
21. Blum JJ. Metabolic pathways in Tetrahymena. Biol Chem. 1972;247(16):5199–209.
22. Blum JJ. Effect of AMP and related compounds on glycogen content ofTetrahymena. J Cell Physiol [Internet]. 1972 Dec;80(3):443–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/4630641
23. Levy MR, Scherbaum OH. Glyconeogenesis in Growing and Non-growing Cultures of Tetrahymena pyriformis. J Gen Microbiol [Internet]. 1965 Feb 1;38(2):221–30. Available from: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-38-2-221
24. Elliott AM, Hayes RE. Mating Types in Tetrahymena. Biol Bull [Internet]. 1953 Oct;105(2):269–84. Available from: http://www.jstor.org/stable/1538642?origin=crossref
25. Doerder FP. Abandoning sex: multiple origins of asexuality in the ciliate Tetrahymena. BMC Evol Biol [Internet]. 2014;14(1):112. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4045964&tool=pmcentrez&rendertype=abstract
26. Cervantes MD, Hamilton EP, Xiong J, Lawson MJ, Yuan D, Hadjithomas M, et al. Selecting One of Several Mating Types through Gene Segment Joining and Deletion in Tetrahymena thermophila. Umen JG, editor. PLoS Biol [Internet]. 2013 Mar 26;11(3):e1001518. Available from: http://dx.plos.org/10.1371/journal.pbio.1001518

27. Ferracci J, Ueno H, Numayama-Tsuruta K, Imai Y, Yamaguchi T, Ishikawa T. Entrapment of Ciliates at the Water-Air Interface. Humphries S, editor. PLoS One [Internet]. 2013 Oct 10;8(10):e75238. Available from: https://dx.plos.org/10.1371/journal.pone.0075238
28. Ishikawa T. Swimming of ciliates under geometric constraints. J Appl Phys. 2019;125(20).
29. Raugi GJ, Liang T, Blum JJ. Effect of oxygen on the regulation of intermediate metabolism in Tetrahymena. J Biol Chem [Internet]. 1975 Jan 25;250(2):445–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/4356248
30. Morel-Journel T, Thuillier V, Pennekamp F, Laurent E, Legrand D, Chaine AS, et al. A multidimensional approach to the expression of phenotypic plasticity. Funct Ecol. 2020;34(11):2338–49.
31. Fronhofer EA, Altermatt F. Eco-evolutionary feedbacks during experimental range expansions. Nat Commun [Internet]. 2015 Nov 22;6(1):6844. Available from: http://www.nature.com/articles/ncomms7844
32. Fronhofer EA, Gut S, Altermatt F. Evolution of density-dependent movement during experimental range expansions. J Evol Biol [Internet]. 2017 Dec;30(12):2165–76. Available from: http://doi.wiley.com/10.1111/jeb.13182
33. Fenchel T. Protozoa and oxygen. Acta Protozool. 2014;53(1):3–12.
34. Steudel B, Hector A, Friedl T, Löfke C, Lorenz M, Wesche M, et al. Biodiversity effects on ecosystem functioning change along environmental stress gradients. Gessner M, editor. Ecol Lett [Internet]. 2012 Dec [cited 2016 Apr 4];15(12):1397–405. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22943183
35. Fischer FM, Wright AJ, Eisenhauer N, Ebeling A, Roscher C, Wagg C, et al. Plant species richness and functional traits affect community stability after a flood event. Philos Trans R Soc B Biol Sci [Internet]. 2016 May 19;371(1694):20150276. Available from: http://rstb.royalsocietypublishing.org/lookup/doi/10.1098/rstb.2015.0276
36. Shipley B, Vile D, Garnier E. From Plant Traits to Plant Communities: A Statistical Mechanistic Approach to Biodiversity. Science (80-) [Internet]. 2006 Nov 3;314(5800):812–4. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1132595
37. Viaene KPJ, De Laender F, Van den Brink PJ, Janssen CR. Using additive modelling to quantify the effect of chemicals on phytoplankton diversity and biomass. Sci Total Environ [Internet]. 2013 Apr;449(2013):71–80. Available from: http://dx.doi.org/10.1016/j.scitotenv.2013.01.046
38. Kersting K, van den Brink PJ. Effects of the insecticide Dursban®4e (active ingredient chlorpyrifos) in outdoor experimental ditches: Responses of ecosystem metabolism. Environ Toxicol Chem [Internet]. 1997 Feb;16(2):251–9. Available from: http://doi.wiley.com/10.1002/etc.5620160222
39. Mcmahon TA, Halstead NT, Johnson S, Raffel TR, Romansic JM, Crumrine PW, et al. Fungicide-induced declines of freshwater biodiversity modify ecosystem functions and services. Ecol Lett. 2012;15(7):714–22.
40. Hodgson JG, Wilson PJ, Hunt R, Grime JP, Thompson K. Allocating C-S-R Plant Functional Types: A Soft Approach to a Hard Problem. Oikos [Internet]. 1999 May;85(2):282. Available from:

https://www.jstor.org/stable/3546494?origin=crossref

41. Pennekamp F, Mitchell KA, Chaine A, Schtickzelle N. Dispersal propensity in tetrahymena thermophila ciliates-A reaction norm perspective. Evolution (N Y) [Internet]. 2014 May;68(8):n/a-n/a. Available from: https://onlinelibrary.wiley.com/doi/10.1111/evo.12428
42. Pennekamp F. Swimming with ciliates - Dispersal and movement ecology of Tetrahymena thermophila. Vol. PhD. 2014.
43. Fjerdingstad EJ, Schtickzelle N, Manhes P, Gutierrez A, Clobert J. Evolution of dispersal and life history strategies – Tetrahymena ciliates. BMC Evol Biol [Internet]. 2007;7(1):133. Available from: http://www.biomedcentral.com/1471-2148/7/133
44. Chaine AS, Schtickzelle N, Polard T, Huet M, Clobert J. Kin-based recognition and social aggregation in a ciliate. Evolution (N Y) [Internet]. 2009 Dec;64–5:1290–300. Available from: http://doi.wiley.com/10.1111/j.1558-5646.2009.00902.x
45. Schtickzelle N, Fjerdingstad EJ, Chaine A, Clobert J. Cooperative social clusters are not destroyed by dispersal in a ciliate. BMC Evol Biol [Internet]. 2009;9(1):251. Available from: http://www.biomedcentral.com/1471-2148/9/251
46. Bruns PJ, Brussard T. Pair Formation in Tetrahymena pyriformis, an Inducible Developmental System. J Exp Zool. 1974;188(3):337–44.
47. Wellnitz WR, Bruns PJ. The pre-pairing events in Tetrahymena thermophila. Exp Cell Res [Internet]. 1979 Mar;119(1):175–80. Available from: http://linkinghub.elsevier.com/retrieve/pii/001448277990346X
48. Pennekamp F, Schtickzelle N. Implementing image analysis in laboratory-based experimental systems for ecology and evolution: a hands-on guide. Hodgson D, editor. Methods Ecol Evol [Internet]. 2013 May;4(5):483–92. Available from: http://doi.wiley.com/10.1111/2041-210X.12036
49. Pennekamp F, Schtickzelle N, Petchey OL. BEMOVI, software for extracting behavior and morphology from videos, illustrated with analyses of microbes. Ecol Evol [Internet]. 2015 Jul;5(13):2584–95. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84937023585&partnerID=tZOtx3y1
50. R Core Team. R: A Language and Environment for Statistical Computing. 2021; Available from: https://www.r-project.org/
51. Kahm M, Hasenbrink G, Lichtenberg-Fraté H, Ludwig J, Kschischo M. Grofit: Fitting biological growth curves with R. J Stat Softw. 2010;33(7):1–21.
52. Wood S. Generalized Additive Models: An Introduction with R. 2nd ed. Chapman and Hall/CRC; 2017.
53. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Things are not Always Linear; Additive Modelling. In 2009. p. 35–69. Available from: http://link.springer.com/10.1007/978-0-387-87458-6_3
54. Lê S, Josse J, Husson F. FactoMineR: An R package for multivariate analysis. J Stat Softw. 2008;25(1):1–18.

55. Kassambara A, Mundt F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. 2020; Available from: https://cran.r-project.org/package=factoextra
56. Abdi H, Williams LJ. Principal component analysis. Wiley Interdiscip Rev Comput Stat. 2010;2(4):433–59.
57. Mlambo MC. Not all traits are 'functional': insights from taxonomy and biodiversity-ecosystem functioning research. Biodivers Conserv [Internet]. 2014 Mar 17;23(3):781–90. Available from: http://link.springer.com/10.1007/s10531-014-0618-5
58. Mouillot D, Mason WHN, Dumay O, Wilson JB. Functional regularity: a neglected aspect of functional diversity. Oecologia [Internet]. 2005 Jan 20;142(3):353–9. Available from: http://link.springer.com/10.1007/s00442-004-1744-7
59. Petchey OL, Gaston KJ. Functional diversity: back to basics and looking forward. Ecol Lett [Internet]. 2006 Jun;9(6):741–58. Available from: http://doi.wiley.com/10.1111/j.1461-0248.2006.00924.x
60. Arnold PA, Kruuk LEB, Nicotra AB. How to analyse plant phenotypic plasticity in response to a changing climate. New Phytol [Internet]. 2019;222(3):1235–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30632169
61. Yang J, Cao M, Swenson NG. Why Functional Traits Do Not Predict Tree Demographic Rates. Trends Ecol Evol [Internet]. 2018;33(5):326–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29605086
62. Laughlin DC, Gremer JR, Adler PB, Mitchell RM, Moore MM. The Net Effect of Functional Traits on Fitness. Trends Ecol Evol [Internet]. 2020;35(11):1037–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32807503
63. Onwuegbuzie AJ, Daniel LG. Uses and Misuses of the Correlation Coefficient. Mid-South Educ Educ Res Assoc [Internet]. 1999;(7):58. Available from: http://eric.ed.gov/?id=ED437399
64. Garland T. Trade-offs. Curr Biol [Internet]. 2014;24(2):R60–1. Available from: http://dx.doi.org/10.1016/j.cub.2013.11.036
65. Alexander RM. The ideal and the feasible: physical constraints on evolution. Biol J Linn Soc [Internet]. 1985 Dec;26(4):345–58. Available from: https://academic.oup.com/biolinnean/article-lookup/doi/10.1111/j.1095-8312.1985.tb02046.x
66. Bennett AF, Lenski RE. An experimental test of evolutionary trade-offs during temperature adaptation. Proc Natl Acad Sci [Internet]. 2007 May 15;104(Supplement 1):8649–54. Available from: http://www.nap.edu/catalog/11790
67. Bennett AE, Bever JD. Trade-offs between arbuscular mycorrhizal fungal competitive ability and host growth promotion in Plantago lanceolata. Oecologia. 2009;160(4):807–16.
68. Nylin S, Wiklund C, Wiklund P-O, Garcia-Barros E. Abscence of trade-off between sexual size dimorphism and early emergence in a buttefly. Ecology. 1993;74(5):1414–27.
69. Brown GE, Smith RJF. Foraging Trade-offs in Fathead Minnows (Pimephales promelas, Osteichthyes, Cyprinidae): Acquired Predator Recognition in the Absence of an Alarm Response. Ethology [Internet].

1996 Apr 26;102(5):776–85. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1439-0310.1996.tb01166.x

70. Dorken ME, Van Drunen WE. Life-history trade-offs promote the evolution of dioecy. J Evol Biol. 2018;31(9):1405–12.
71. Van Drunen WE, Dorken ME. Trade-offs between clonal and sexual reproduction in sagittaria latifolia (alismataceae) scale up to affect the fitness of entire clones. New Phytol. 2012;196(2):606–16.
72. Rocha FB, Klaczko LB. Connecting The Dots Of Nonlinear Reaction Norms Unravels The Threads Of Genotype-Environment Interaction In Drosophila. Evolution (N Y). 2012;66(11):3404–16.
73. Saatkamp A, Römermann C, Dutoit T. Plant Functional Traits Show Non-Linear Response to Grazing. Folia Geobot [Internet]. 2010 Sep 29;45(3):239–52. Available from: http://link.springer.com/10.1007/s12224-010-9069-2
74. Jenouvrier S, Holland M, Stroeve J, Barbraud C, Weimerskirch H, Serreze M, et al. Effects of climate change on an emperor penguin population: Analysis of coupled demographic and climate models. Glob Chang Biol. 2012;18(9):2756–70.
75. Van De Pol M, Vindenes Y, Sæther BE, Engen S, Ens BJ, Oosterbeek K, et al. Effects of climate change and variability on population dynamics in a long-lived shorebird. Ecology. 2010;91(4):1192–204.
76. Jonzén N, Pople T, Knape J, Sköld M. Stochastic demography and population dynamics in the red kangaroo Macropus rufus. J Anim Ecol. 2010;79(1):109–16.
77. Pérez-Ramos IM, Matías L, Gómez-Aparicio L, Godoy Ó. Functional traits and phenotypic plasticity modulate species coexistence across contrasting climatic conditions. Nat Commun [Internet]. 2019;10(1):2555. Available from: http://dx.doi.org/10.1038/s41467-019-10453-0
78. Boyer AG, Jetz W. Extinctions and the loss of ecological function in island bird communities. Glob Ecol Biogeogr [Internet]. 2014 Jun;23(6):679–88. Available from: http://doi.wiley.com/10.1111/geb.12147
79. Laughlin DC, Messier J. Fitness of multidimensional phenotypes in dynamic adaptive landscapes. Trends Ecol Evol [Internet]. 2015;30(8):487–96. Available from: http://dx.doi.org/10.1016/j.tree.2015.06.003
80. Poff NL, Olden JD, Vieira NKM, Finn DS, Simmons MP, Kondratieff BC. Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships. J North Am Benthol Soc [Internet]. 2006 Dec;25(4):730–55. Available from: http://www.journals.uchicago.edu/doi/10.1899/0887-3593(2006)025[0730:FTNONA]2.0.CO;2

Figures

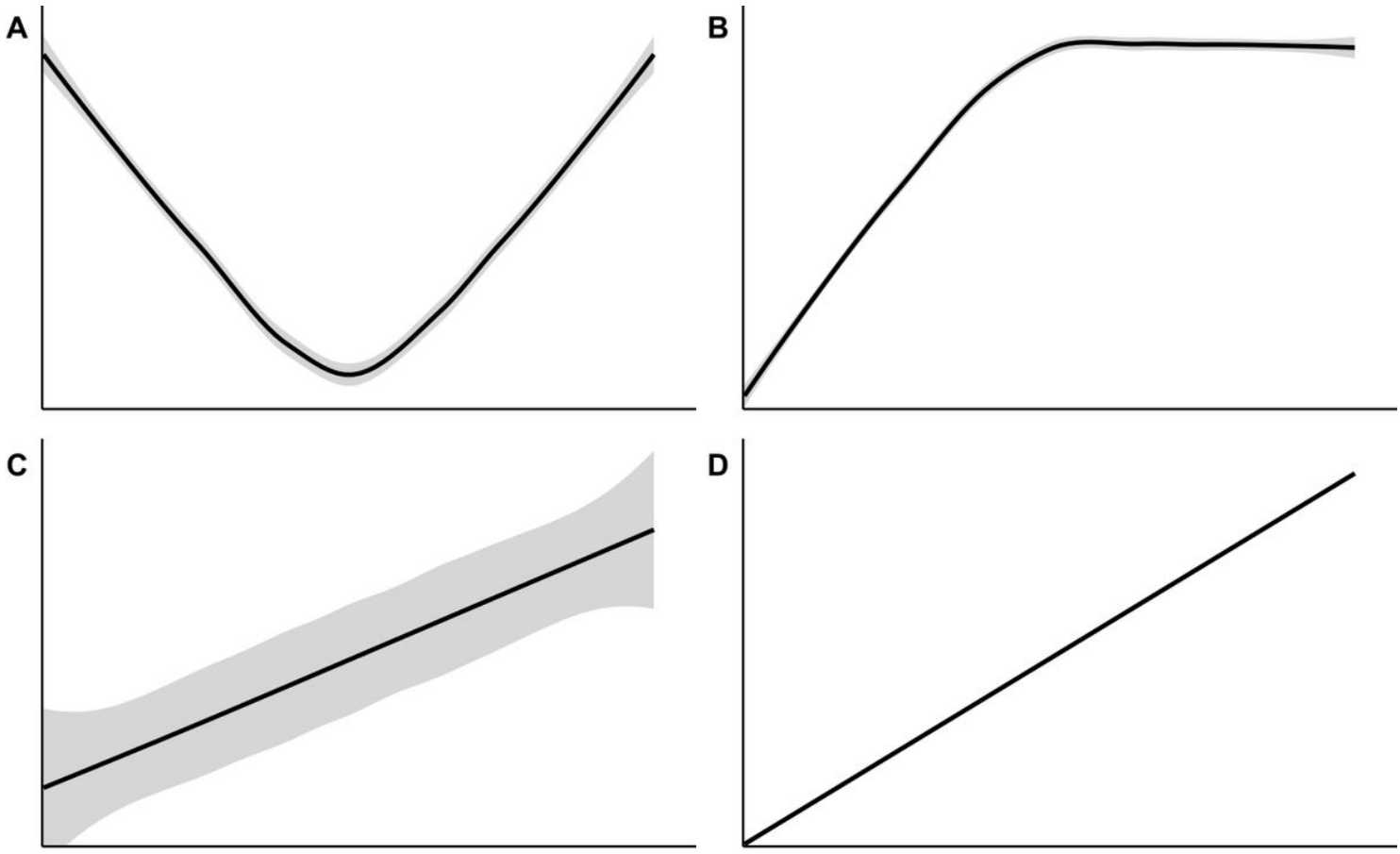

Figure 1

Four scenarios illustrating a spectrum of possible relationships between hard and soft traits: A) This non-monotonic relationship allows prediction of the hard trait using the soft one, but not through a simple linear method. (B) An example of a monotonic relation, where the two traits are linearly related on a portion of their variation domain, only allowing accurate predictions on this part. (C) The traits are here linearly related, but reliable prediction cannot be achieved because of the high standard error. (D) The ideal linear, strong and monotonic relationship, that is assumed to exist in most cases. Thus, one can use a trait as proxy for another one only if there is a well-known relationship that is correctly estimated, implying (1) the knowledge of the form of the relationship between the two traits, (2) a relationship that is not completely flat, since any flat part is unusable, and (3) a standard error on the model parameter small enough to give a reliable prediction.

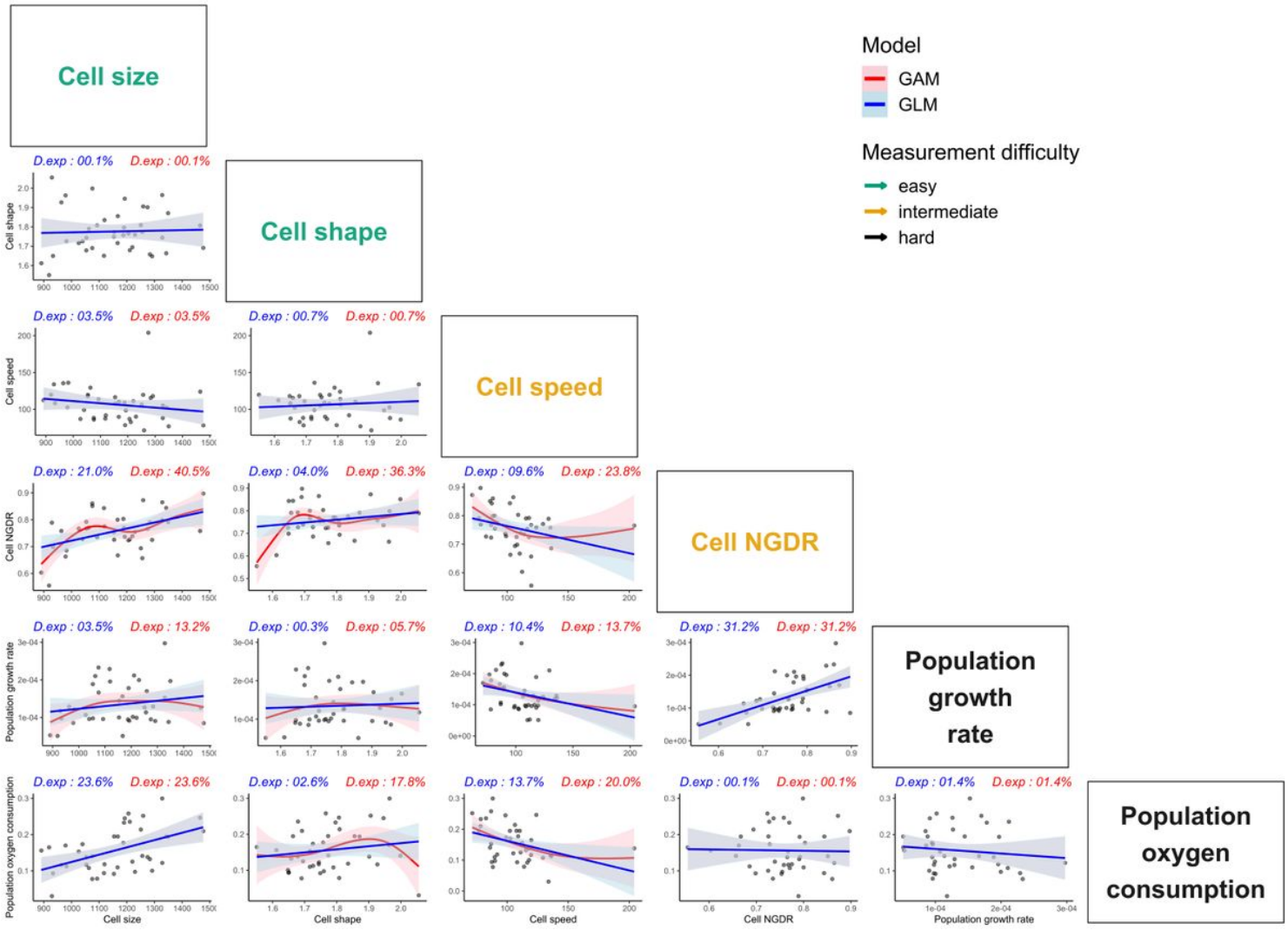

Figure 2

Pairwise relationships among the six functional traits measured for the 40 *T. thermophila* strains. Each dot represents the average value of all replicates at the strain level, the blue line a linear model (GLM), and the red curve a GAM. Both are represented with their respective 95% confident interval. Above every graph is displayed the deviance explained (D.exp) by each model. When outlined, the non-linear GAM is significantly better than the GLM (e.d.f. > 1), otherwise the non-linearity does not improve the model's fit and the GAM is behaving exactly as the GLM, hence the same D.exp.

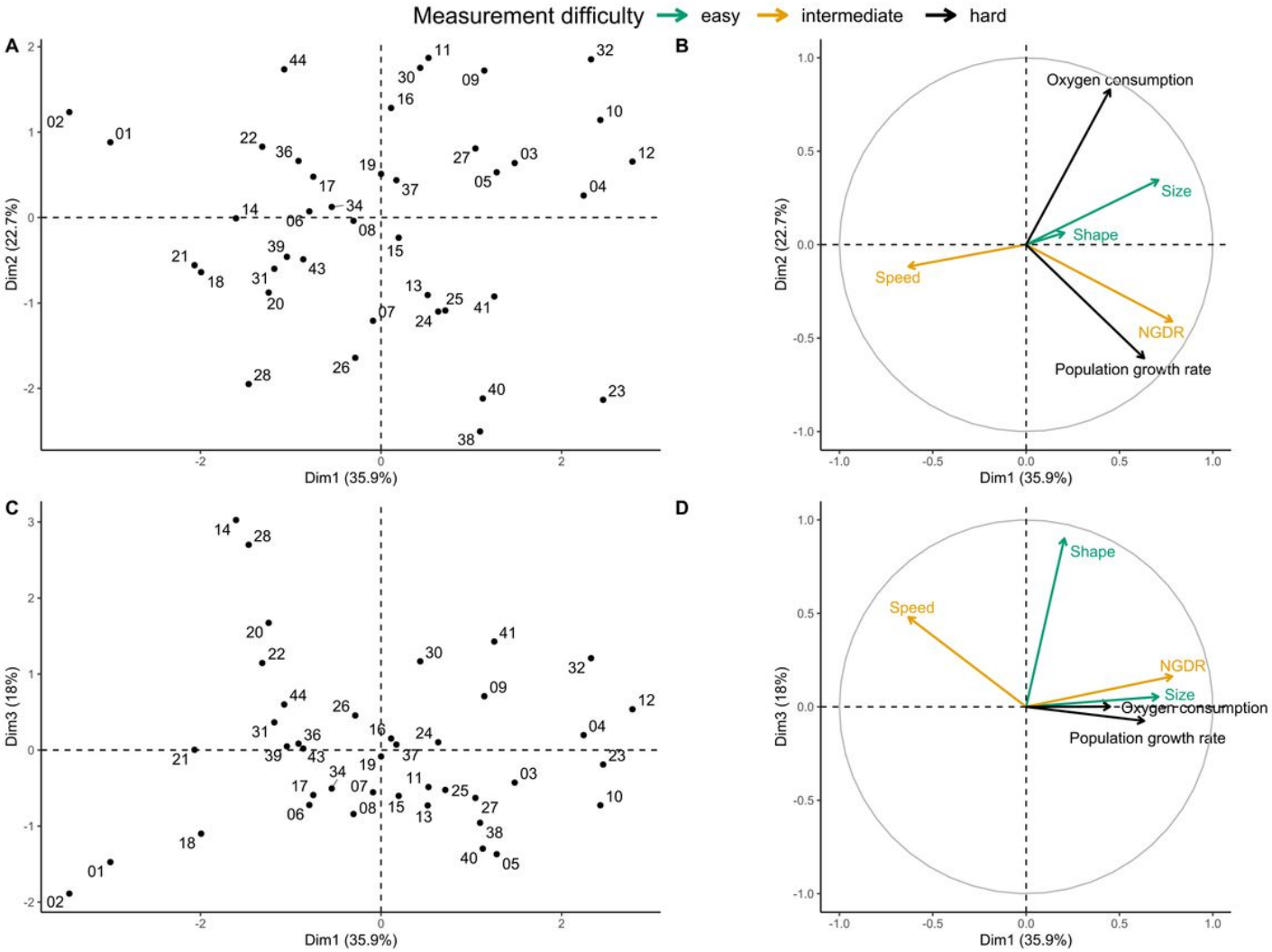

Figure 3

PCA analysis performed on the six functional traits (averaged at the strain level) for the 40 *T. thermophila* strains. The left panel Fig.s represent the distribution of the strains along dimensions 1 and 2 (A) and along dimensions 1 and 3 (C). The numbers stand for the labels of the strains. The right panel Fig.s represent the associated correlation circles along the first and second dimensions (B) and the first and third dimensions (D). The functional traits are colored based on their measurement difficulty.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.pdf
- Additionalfile2.pdf
- Additionalfile3.pdf