**Data Analysis Description for Supplementary Materials**

Given the large number of predictive variables compared to the number of observations in the data, for all analyses we perform a variable selection procedure which takes into account and removes collinear variables, thereby preserving only those that are more predictive for each response considered. Therefore we use the Least Absolute Shrinkage and Selection Operator, hereinafter Lasso regression, before proceeding with the appropriate regression models for each specific problem. Indeed the Lasso is particularly effective in high-dimensional settings where multicollinearity may be present (Tibshirani, 1996). When Lasso is not directly applicable to a response type in our analysis, a pseudo-response is used to facilitate variable selection. For example, Multinomial Lasso regression is implemented for compositional response variables, such as categorical pathogen species proportions. This method extends Lasso to multiclass classification problems (Friedman et al., 2010) and, since compositional data inherently sum to one, multinomial models ensure that the structure of the response variable is appropriately maintained. This being done, once the most relevant predictors are identified for all responses in the analysis, they are utilized in the final regression models tailored to the data type and research objective.

To apply the Lasso, the data is preprocessed by encoding categorical variables, such as temporal and commercial scale indicators, to be integrated into the following regression models. Numeric predictors are standardized, particularly to ensure balanced contributions in Lasso regression, which is sensitive to variable scale. The optimal regularization parameter is determined using cross-validation, ensuring an appropriate trade-off between model complexity and predictive performance. Once the important variables for each analysis have been selected, we proceed to the most appropriate form of regression to test the importance and impact of these variables.

* For response variables that fall within the (0,1) range, such as proportions or relative abundance measurements, beta regression is employed (Ferrari & Cribari-Neto, 2004). This model is preferred due to its ability to accommodate non-normality while ensuring that predicted values remain within the constrained range. Maximum likelihood estimation (MLE) is used relying on the logit link function.
* Ordinal logistic regression, commonly referred to as the proportional odds model, is used when the response variable is ordinal, such as for the disease severity scores (McCullagh, 1980). This method appropriately accounts for the ordered nature of the response variable while assuming that the relationship between predictor variables and the response follows a proportional odds structure. Maximum likelihood estimation is applied and the model coefficients provide insights into the log-odds of transitioning between severity levels.
* For relative abundance data constrained to sum to one, Dirichlet regression provides a robust analytical framework (Hijazi & Jernigan, 2009). This model accounts for the multivariate nature of compositional data, ensuring that predictions remain valid within the unit simplex. Data transformations are employed where necessary to avoid numerical instability due to zero values. The Dirichlet distribution effectively handles overdispersion, making it an appropriate choice for modeling count-based compositions.

**References**

* Ferrari, S. L., & Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *Journal of Applied Statistics, 31*(7), 799-815.
* Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*(1), 1-22.
* Hijazi, R. H., & Jernigan, R. W. (2009). Modelling compositional data using Dirichlet regression models. *Journal of Applied Probability and Statistics, 4*(1), 77-91.
* McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological), 42*(2), 109-142.
* Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological), 58*(1), 267-288.