Appendix S3

3. Why does site-based statistical testing in a GLM fail?
The reason of the failure of the test is that the importance of e in the species-by-sites interaction is not a sufficient condition for trait-environment association, even when the test-statistic tries to capture the t-e interaction only. Consider the GLMM model
\[ \log(\mu_{ij}) = R_i + C_j + \beta_j e_i \] with \( \beta_j \) a species-specific slope with respect to e, modelled as a linear model of trait t, with intercept \( b_0 \) and slope \( b_{te} \), and with \( \epsilon_{\beta_j} \) a normally distributed error term with mean 0 and variance \( \sigma_{\beta}^2 \). By inserting the model for \( \beta_j \) in the log-linear model, we see that \( b_{te} \) is indeed the coefficient of the interaction \( t_j e_i \). The term \( b_0 e_i \) can be absorbed in the row main effect \( R_i \), and \( \epsilon_{\beta_j} e_i \) represents additional species-specific random variation that interacts with the observed environment. For judging the significance of \( b_{te} \) one can plot estimates of \( \beta_j \) against \( t_j \) and add the fitted line to the plot, as Jamil et al. (2013) did in their Fig. 3. Clearly, the scatter of the species-specific slopes around the fitted line gives an insight about the importance of the relationship. This scatter, in Equation (2) represented by \( \epsilon_{\beta_j} \), is taken into account in the GLMM model, but not in the GLM model with site-based resampling as in anova.traitglm. Therefore, resampling of sites, as in anova.traitglm, is expected to generate inflated type I error rates, unless the parameter \( \sigma_{\beta} = 0 \).