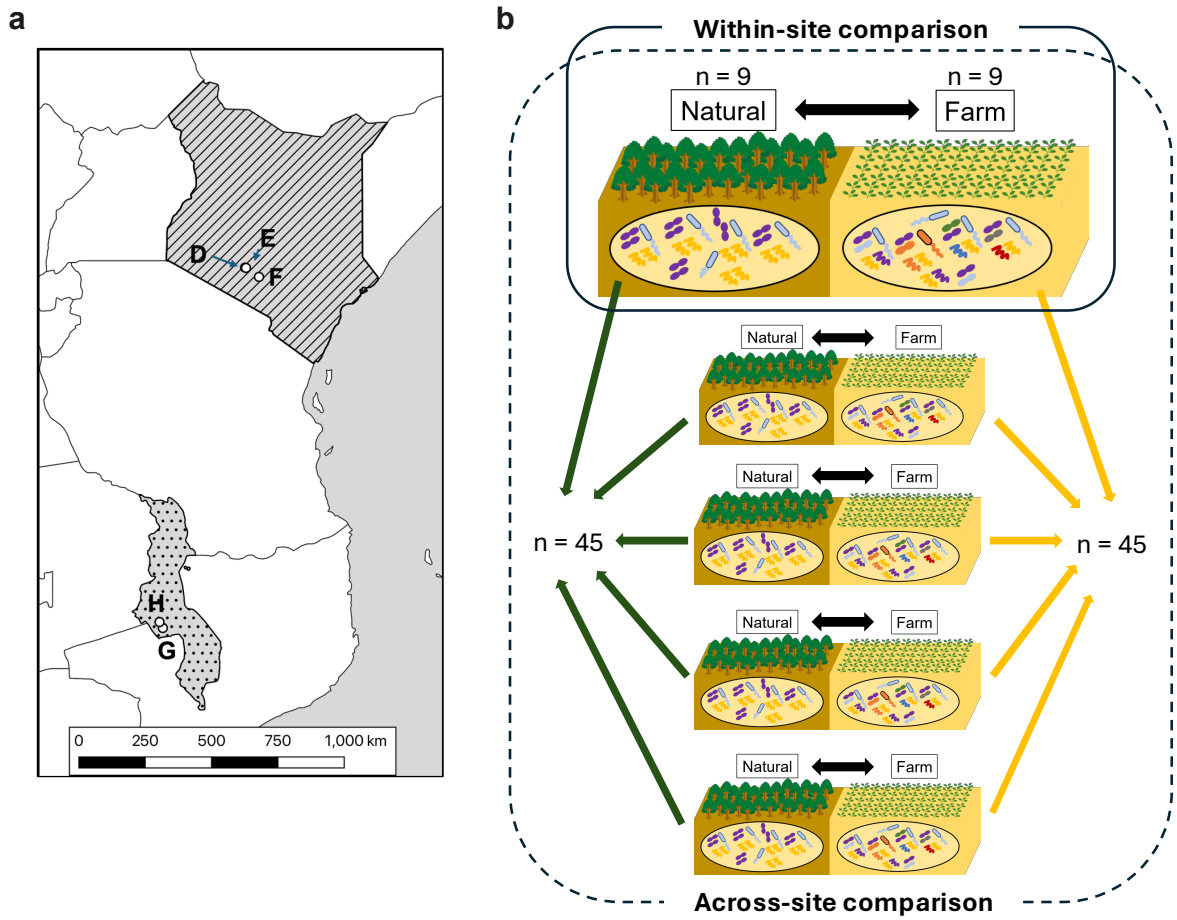

Figure 1. Location of sampling sites and concept of the analysis. (**a**) The sampling locations, Sites D—F in Kenya and Sites G and H in Malawi, are annotated in the map. (**b**) All sites had a pair of neighboring natural (unmanaged) land and farmland. Nine soil samples were collected at each land use at one site, resulting in 45 samples for each land use across five sites.

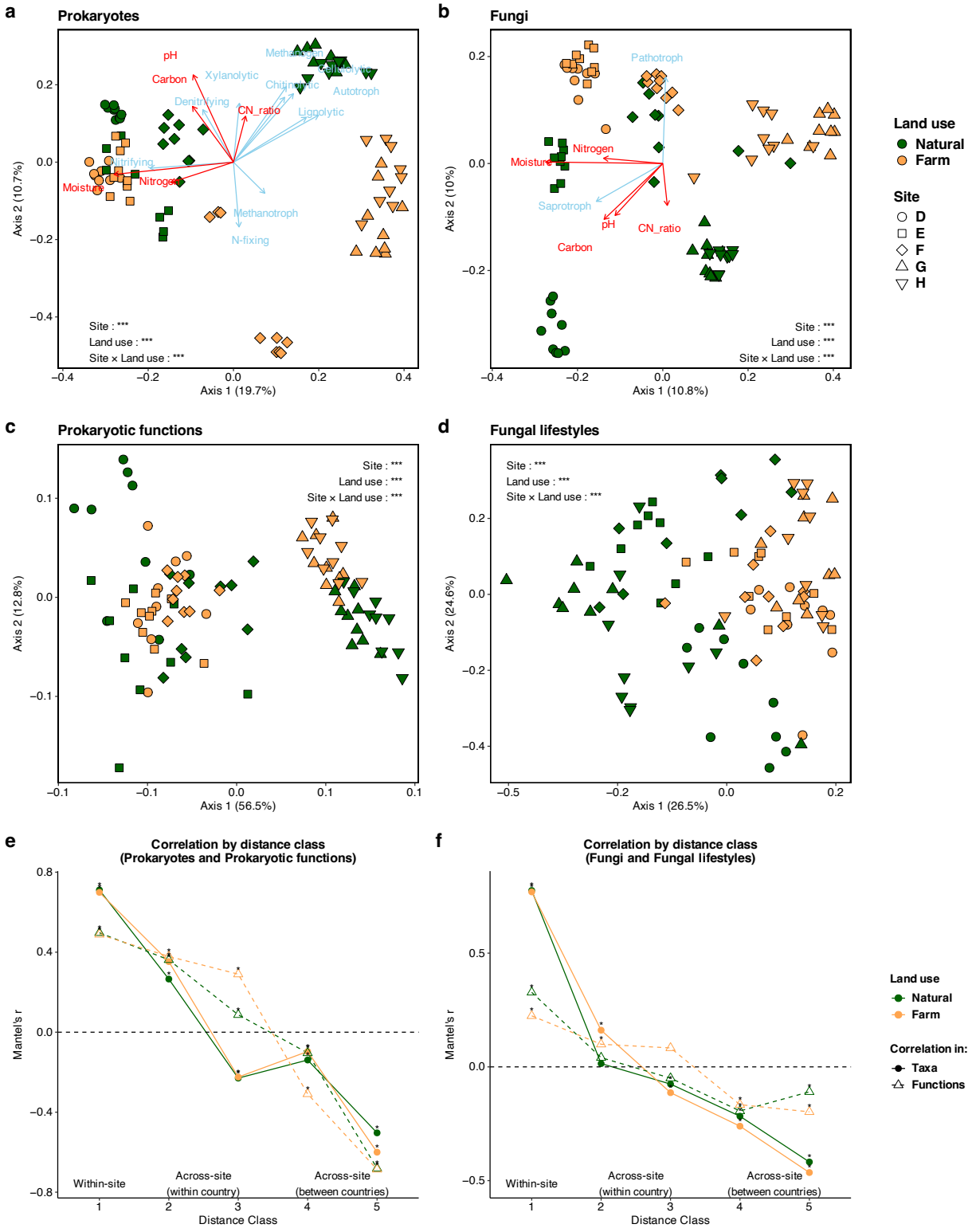

Figure 2. Dissimilarity of soil microbial taxonomic and functional compositions in natural lands and farmlands. PCoA plots for (**a**) prokaryotic and (**b**) fungal taxa. Prokaryotic or fungal functions and the environmental variables that significantly correlated with the communities are illustrated with arrows, colored by red and skyblue, respectively. PCoA plots for (**c**) prokaryotic and (**d**) fungal functions are shown. The p-values in the PERMANOVA on sites, land uses and their interaction are indicated with "*", "**", or "***", representing $p < 0.05$, $p < 0.01$, or $p < 0.001$, respectively. Mantel correlograms of microbial taxonomic and functional compositions by distance class are plotted for (**e**) prokaryotes and prokaryotic functions and (**f**) fungi and fungal lifestyles. The significant autocorrelations (corrected $p < 0.05$) in each distance class tested by permutation is indicated with *.

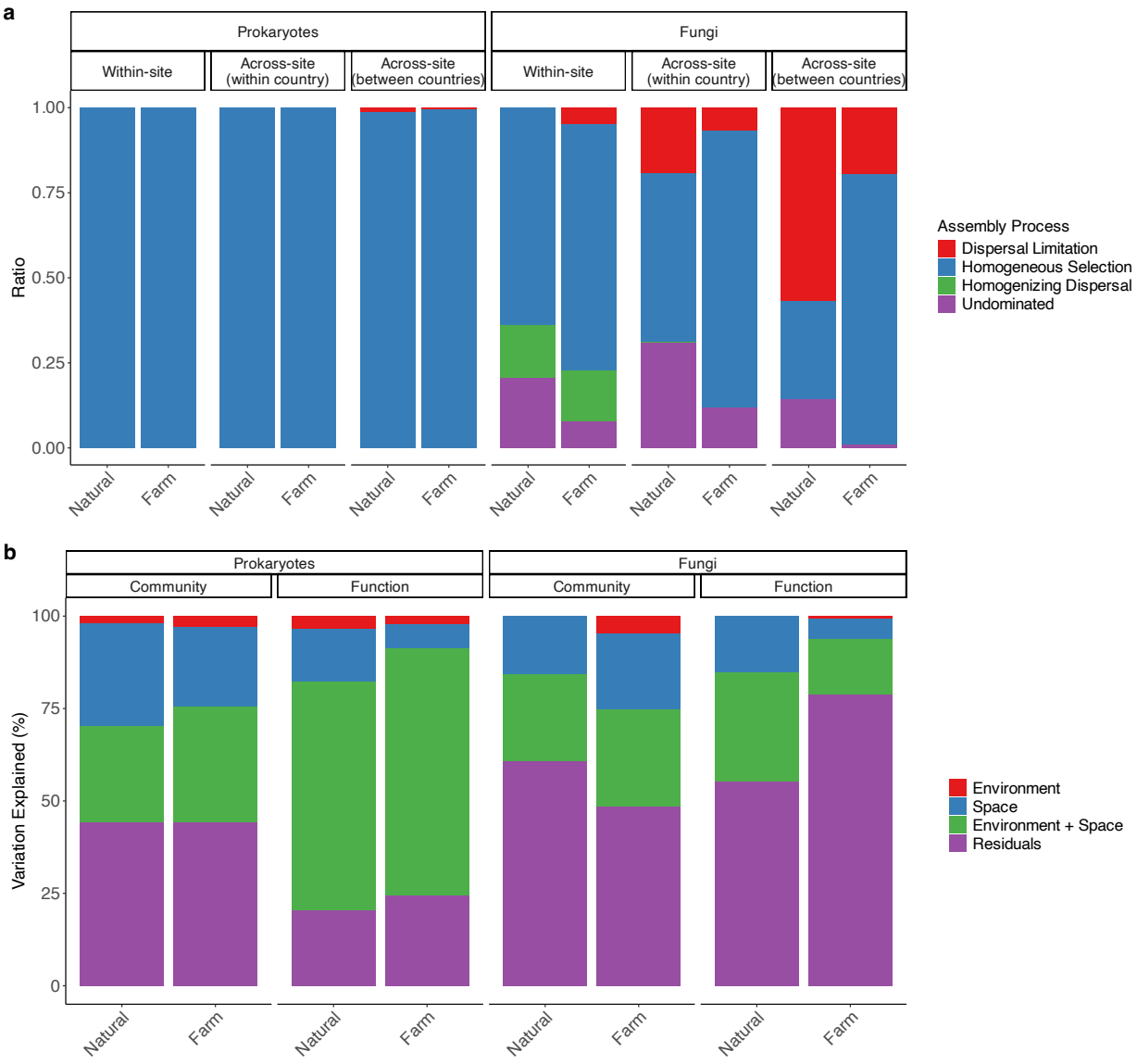

Figure 3. Community assembly process and drivers of taxonomic and functional variation. (**a**) The community assembly process estimated by βNTI and RC_{bray} was grouped by the within-site, across-site (within the country), and across-site (between the countries) scales. (**b**) The variations in taxonomic and functional compositions that were explained by environmental factors, space, their overlap, and residuals are shown.

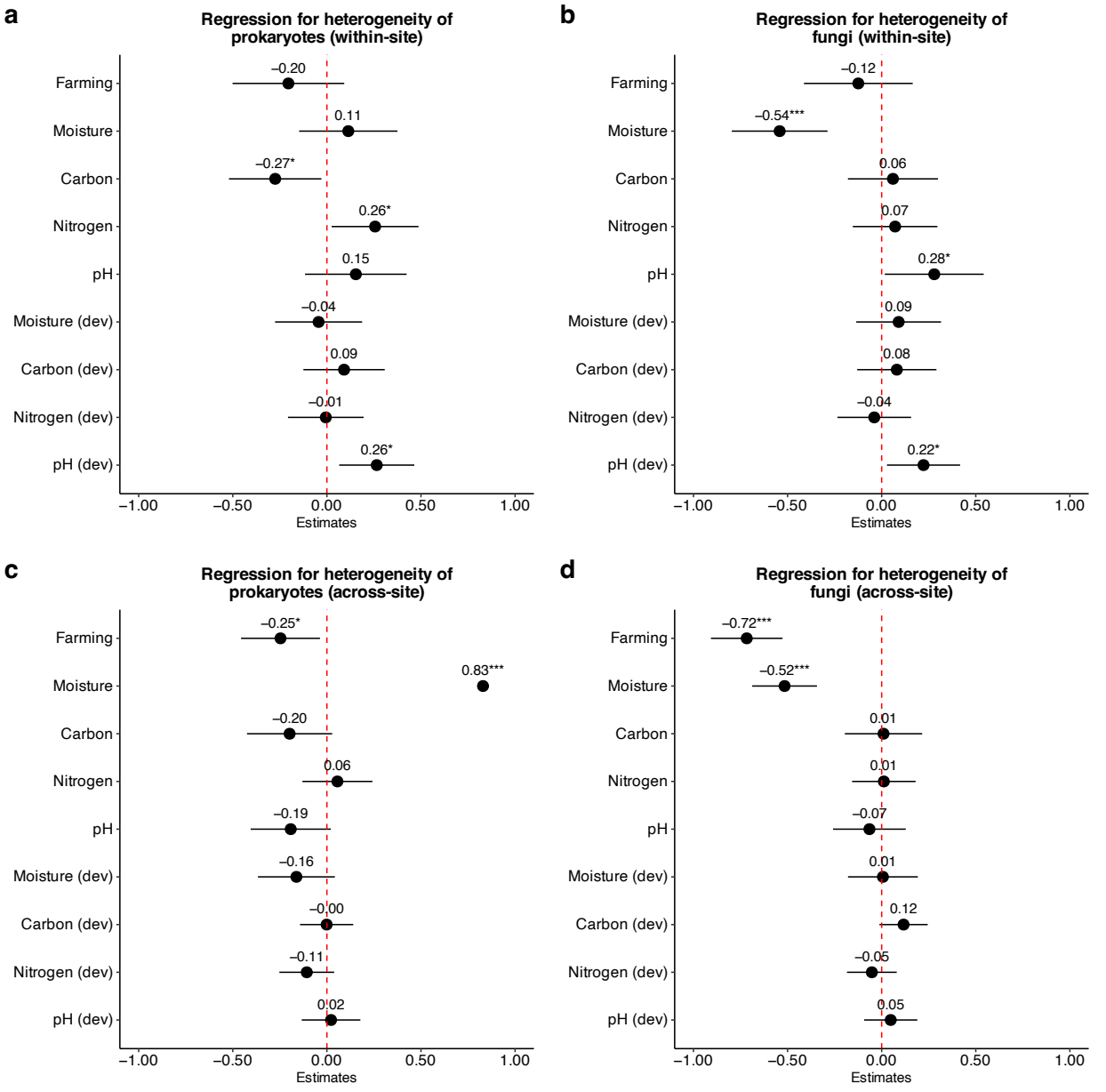

Figure 4. Contributions of farming activity and environmental variables to the heterogeneity of microbial community. The 95% confidence intervals of the regression coefficients, calculated in the multiple regression from farming activity, environmental variables, and absolute deviations of environmental variables to the distance to centroids of (**a, c**) prokaryotes and (**b, d**) fungi in the (**a, b**) within-site and (**c, d**) across-site scales, are shown. The values of coefficients and significance as indicated with "*", "**", or "***", representing $p < 0.05$, $p < 0.01$, or $p < 0.001$, respectively, are noted above each point.

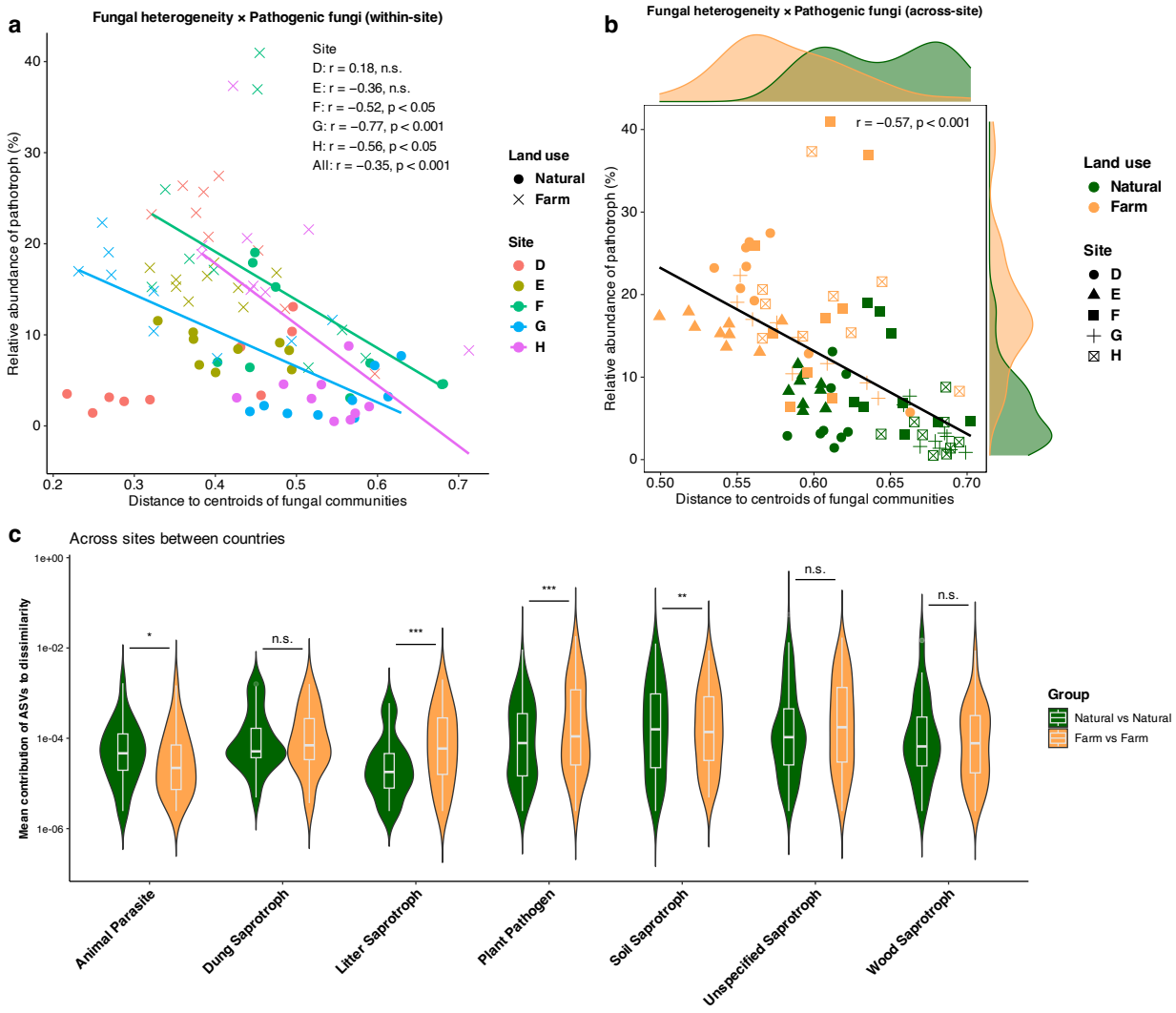

Figure 5. Relation between the heterogeneity of fungal communities and abundance of pathogenic fungi. Correlations between distance to centroids of the fungal communities and relative abundance of pathogenic fungi are plotted for the (**a**) within-site and (**b**) across-site scales. The correlation coefficients and p-values in the Pearson's correlation tests are indicated with "*", "**", or "***", representing $p < 0.05$, $p < 0.01$, or $p < 0.001$, respectively. Regression lines are shown for within sites (in **a**) or for across sites (in **b**) where significant correlations were observed ($p < 0.05$). (**c**) The contributions of each fungal ASV to the Bray–Curtis dissimilarity among samples within each land use across site between countries were averaged and grouped by fungal lifestyle. Asterisks ("*", "**", or "***") and n.s. indicate the significance levels of adjusted p-values: < 0.05, $p < 0.01$, $p < 0.001$, or no significant difference respectively, tested by 1000 permutations to assess differences in the mean values between natural lands and farmlands.