1. Random forest results. Charts A, B and C compare the NMSE (Normalized Mean Square Error) distributions of rea[l] and randomized data. Chart D plots the distribution of paired differences between NMSEs obtained from the environmental (ENV) and original abundance (Operational Taxonomic Units, OTU) data (blue) versus ENV and randomized OTU data (gray). Chart E shows the AUC (Area under the curve) values of the binary classifiers obtained by thresholding the output of the regression model (class 0 associated to the 60% of the samples with the lowest OTU abundance and class 1associated to the remaining 40%).