

FIG S1 Mixed perturbations effect on prediction error in the *SIM* dataset. 100 variations of the *SIM* dataset were generated by combining a random extent of each perturbation (except P#5). Each circle corresponds to a dataset. For each dataset, 10 training sets of 400 samples were subsampled and the rest of the samples served as a test set. Classifiers were trained on unperturbed data and tested on perturbed data. The values on the x-axis correspond to mean cosine distance across all samples in a dataset, calculated between the perturbed version of a sample and its unperturbed origin. The size of each circle reflects the number of peaks that were deleted on average on this dataset. Outlier datasets, with relatively high error rates for samples that were on average still quite similar to their origin (reflected by low mean cosine dissimilarity) are encircled.

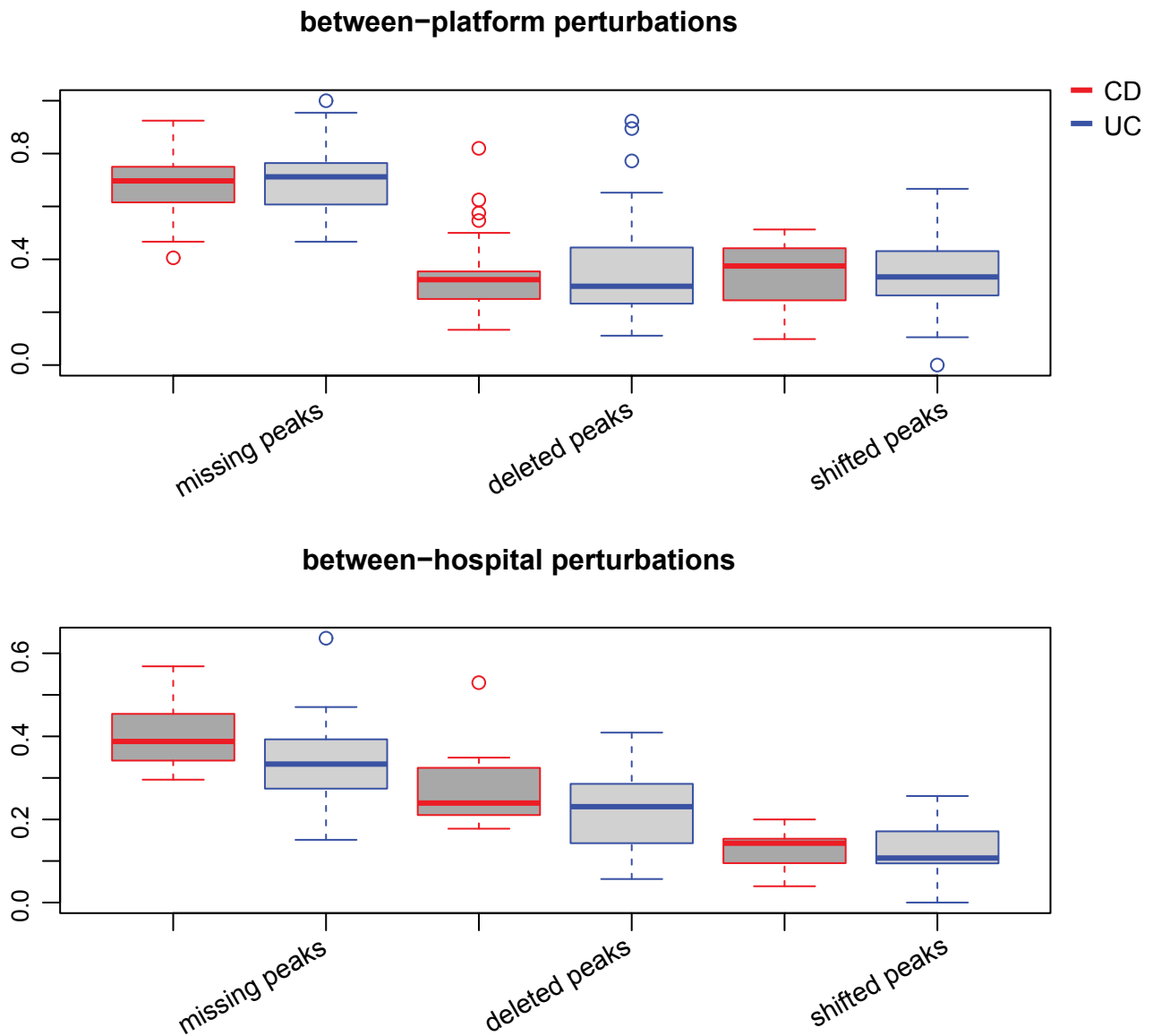


FIG S2 UC and CD variation. Comparison of variation observed between CD and UC samples. No between-platform (top) or between-hospital (bottom) differences were observed between the two IBD subtypes by means of missing, shifted or deleted peaks based on replicates pairwise comparison. A peak is defined missing if it is detected in the reference profile but is missing from its replicate. A missing peak can be either shifted or deleted. It is shifted if it is detected in +1 or -1 position in the replicate profile. Otherwise, it is defined “deleted”.