



Figure 3. External validation of the disease specificity of PDAC fecal microbiome models.

False positive rate (FPR) of metagenomic unconstrained model-1 and enrichment-constrained model-2 in 25 external test sets is shown as barplot (see Supplementary Table S4 for list of all studies included). Validation datasets were profiled and normalized in the same way as the initial dataset (see Methods). Validation datasets were profiled and normalized in the same way as the initial dataset (see Methods). Each study was stratified according to health status and models were tested to predict in the given group with 90% specificity cut-off. Low FPR on metagenomes from patients with other disorders and healthy individuals indicates that the model is specific to the PDAC which was trained to recognise. The number of subjects in each group is displayed as color coded circles below. BRCA: breast cancer, CRC: colorectal cancer, CD: Crohn's disease, CP: chronic pancreatitis, CTR: controls, LD: liver disease, NAFLD: non-alcoholic fatty liver disease, PC: Pancreatic cancer, T1D: type-1 diabetes, T2D: type-2 diabetes, UC: ulcerative colitis, ES: Spanish, DE: German.

Figure 3