

S3 Table. Classification of feature sets

	No upscaling or amplification in the train and test sets	Only upscaling in the train and test sets	With upscaling and amplification in the train and test sets
Selected microbiome data dominantly expressed in pCR response	01	07	13
Selected microbiome data dominantly expressed in RD response	02	08	14
Microbiome with significant differences	03	09	15
Microbiome data dominantly expressed in pCR response and immune cell fraction data	04	10	16
Microbiome data dominantly expressed in RD response and immune cell fraction data	05	11	17
Microbiome with significant differences and immune cell fraction data	06	12	18

pCR, pathological complete response; RD, residual disease.