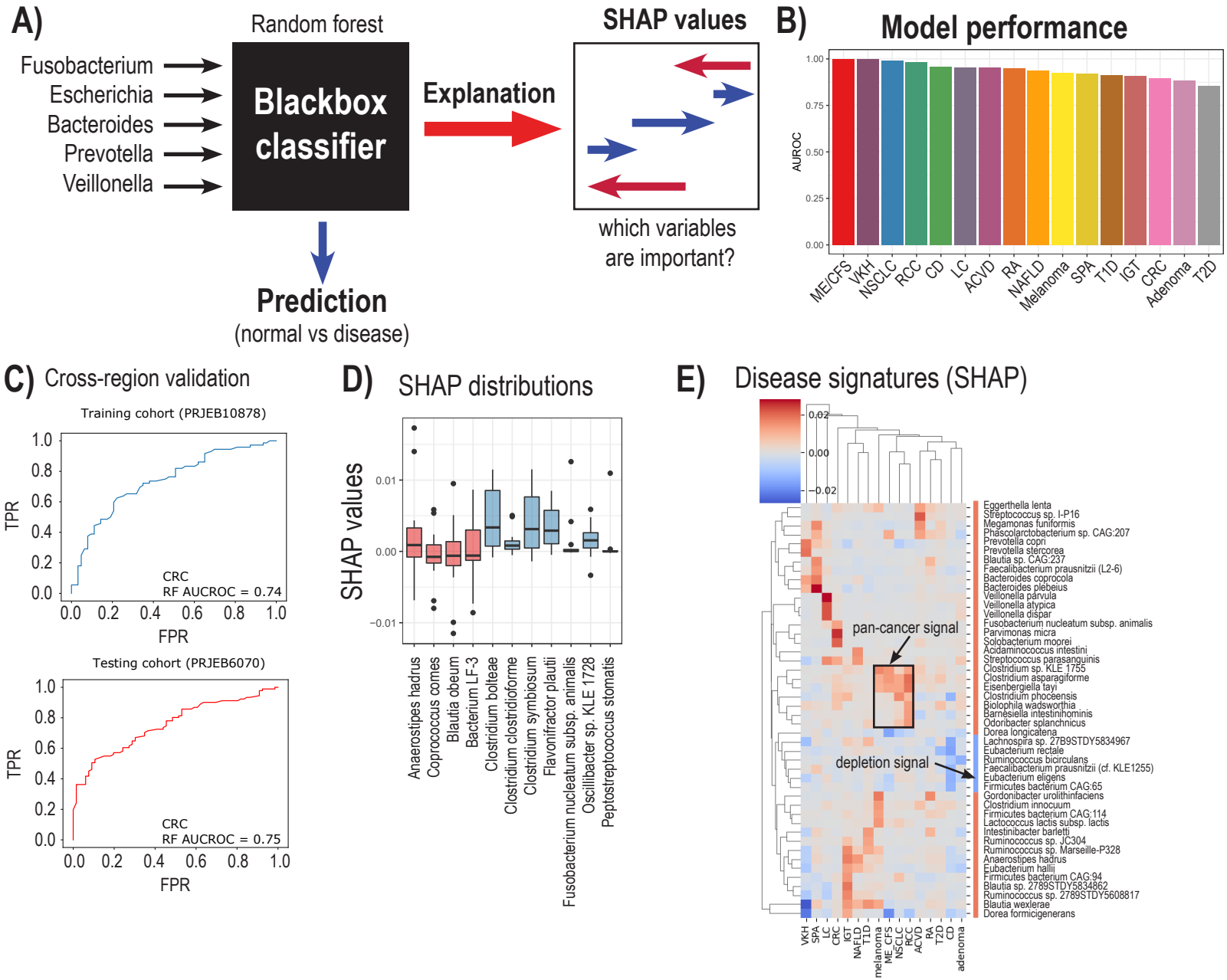
A)
Random forest
Fusobacterium
Escherichia
Bacteroides
Prevotella
Veillonella
Blackbox classifier
Explanation
SHAP values
which variables are important?
Prediction
(normal vs disease)
B)
Model performance
AUROC
1.00
0.75
0.50
0.25
0.00
ME/CFS
VKH
NSCLC
RCC
CD
LC
ACVD
RA
NAFLD
Melanoma
SPA
T1D
IGT
CRC
Adenoma
T2D
C)
Cross-region validation
Training cohort (PRJEB10878)
TPR
FPR
CRC
RF AUCROC = 0.74
Testing cohort (PRJEB6070)
CRC
RF AUCROC = 0.75
D)
SHAP distributions
SHAP values
0.01
0.00
−0.01
Anaerostipes hadrus
Coprococcus comes
Blautia obeum
Bacterium LF-3
Clostridium bolteae
Clostridium clostridioforme
Clostridium symbiosum
Flavonifractor plautii
Fusobacterium nucleatum subsp. animalis
Oscillibacter sp. KLE 1728
Peptostreptococcus stomatis
E)
Disease signatures (SHAP)
0.02
0.00
−0.02
pan-cancer signal
depletion signal
Eggerthella lenta
Streptococcus sp. I-P16
Megamonas funiformis
Phascolarctobacterium sp. CAG:207
Prevotella copri
Prevotella stercorea
Blautia sp. CAG:237
Faecalibacterium prausnitzii (L2-6)
Bacteroides coprocola
Bacteroides plebeius
Veillonella parvula
Veillonella atypica
Veillonella dispar
Fusobacterium nucleatum subsp. animalis
Parvimonas micra
Solobacterium moorei
Acidaminococcus intestini
Streptococcus parasanguinis
Clostridium sp. KLE 1755
Clostridium asparagiforme
Eisenbergiella tayi
Clostridium phoceensis
Biolophila wadsworthia
Barnesiella intestinihominis
Odoribacter splanchnicus
Dorea longicatena
Lachnospira sp. 27B9STDY5834967
Eubacterium rectale
Ruminococcus bicirculans
Faecalibacterium prausnitzii (cf. KLE1255)
Eubacterium eligens
Firmicutes bacterium CAG:65
Gordonibacter urolithinfaciens
Clostridium innocuum
Firmicutes bacterium CAG:114
Lactococcus lactis subsp. lactis
Intestinibacter bartlettii
Ruminococcus sp. JC304
Ruminococcus sp. Marseille-P328
Anaerostipes hadrus
Eubacterium hallii
Firmicutes bacterium CAG:94
Blautia sp. 2789STDY5834862
Ruminococcus sp. 2789STDY5608817
Blautia wexlerae
Dorea formicigenerans
VKH
SPA
LC
CRC
IGT
NAFLD
T1D
melanoma
ME_CFS
NSCLC
RCC
ACVD
RA
T2D
CD
adenoma

Figure 5