Supporting Information

Deep learning accurately predicts estrogen receptor status in breast cancer metabolomics data

Fadhl M Alakwaa[1], Kumardeep Chaudhary[1], Lana X Garmire[1,2*]

[1]Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI 96813, USA.

[2]Molecular Biosciences and Bioengineering Graduate Program, the University of Hawaii at Manoa, Honolulu, HI 96822, USA.

*To whom correspondence should be addressed.

Lana X Garmire, PhD

Associate Professor,

Email address: lgarmire@cc.hawaii.edu

Phone: +1 (808) 441-8193

Table of contents:

Figure S1

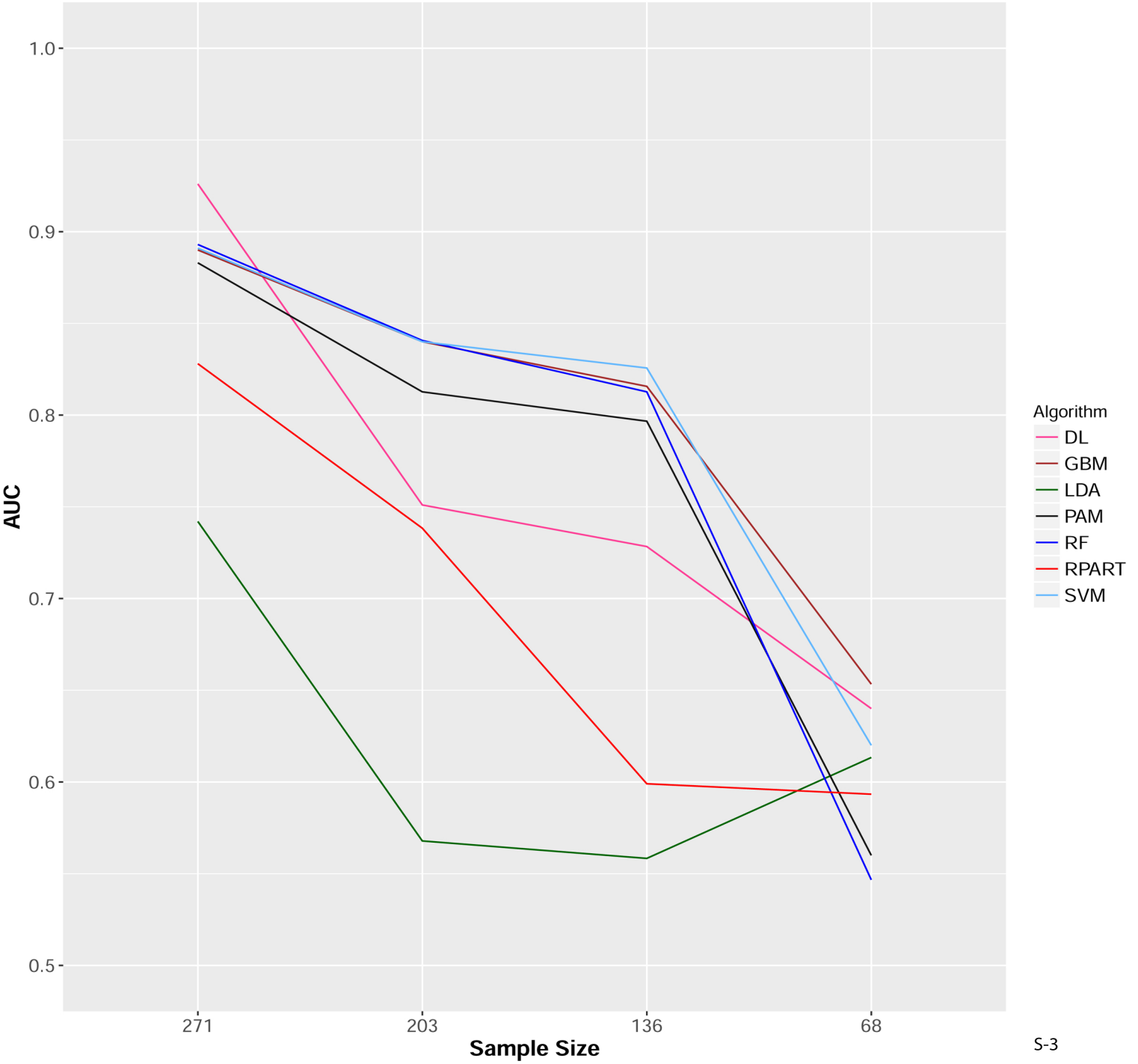

1.0
0.9
0.8
0.7
0.6
0.5
AUC
271
203
136
68
Sample Size
Algorithm
DL
GBM
LDA
PAM
RF
RPART
SVM

Figure S2

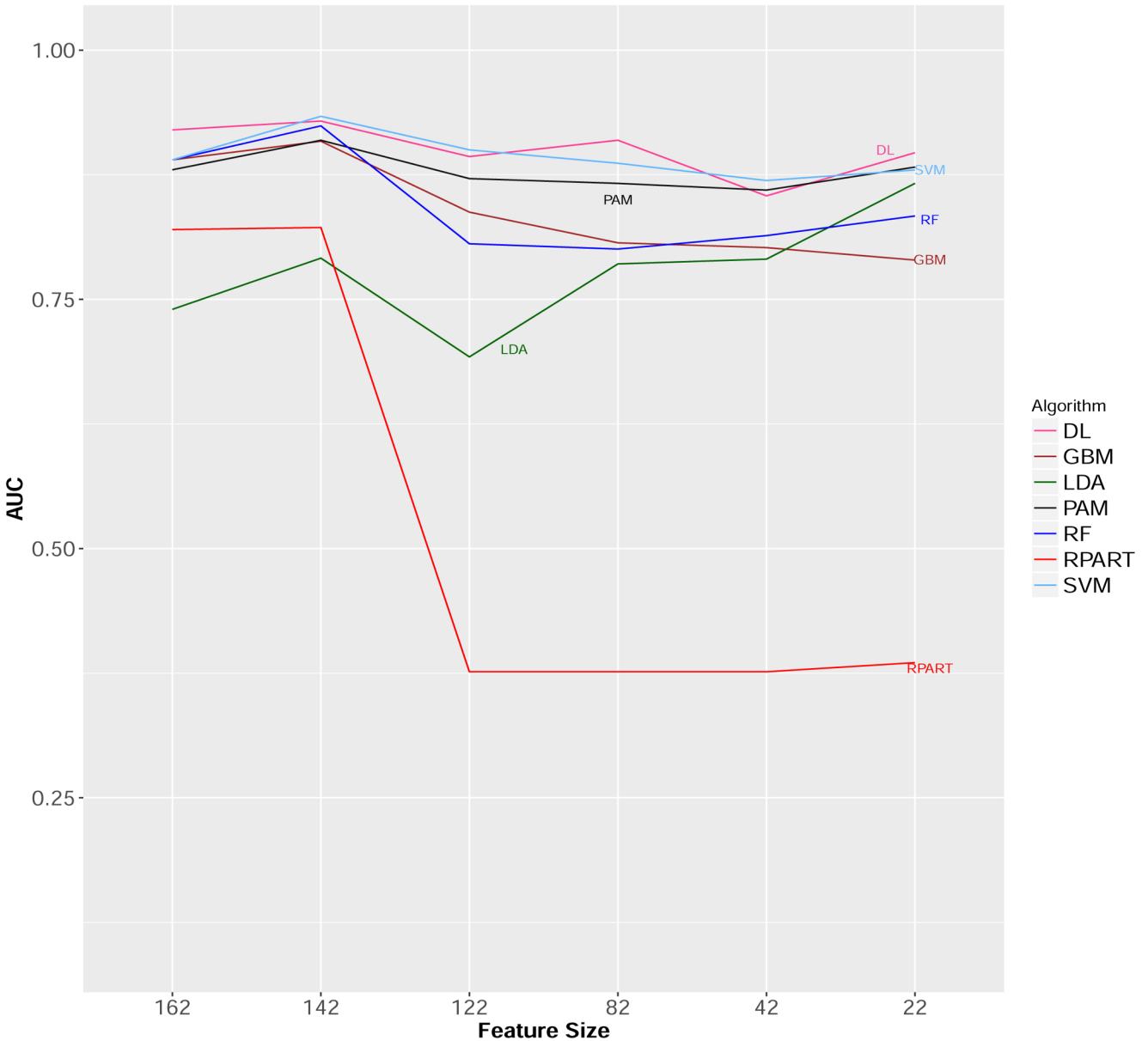
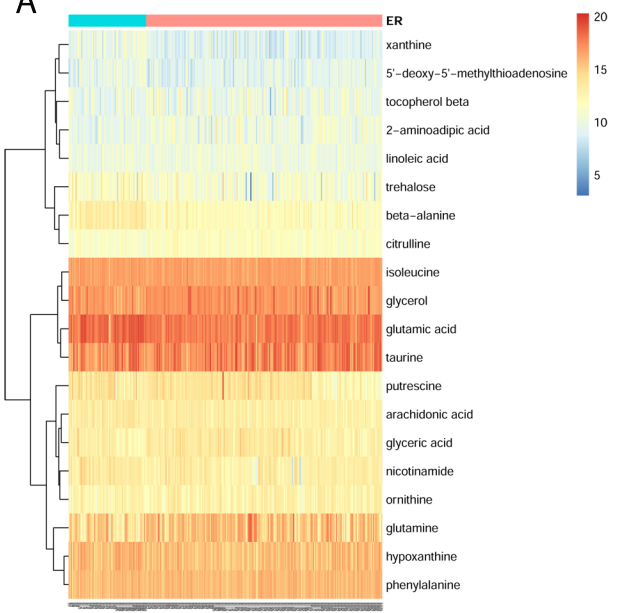

1.00
0.75
0.50
0.25
AUC
162
142
122
82
42
22
Feature Size
DL
SVM
PAM
RF
GBM
LDA
RPART
Algorithm
DL
GBM
LDA
PAM
RF
RPART
SVM

Figure S3


A
ER
xanthine
5'-deoxy-5'-methylthioadenosine
tocopherol beta
2-aminoadipic acid
linoleic acid
trehalose
beta-alanine
citrulline
isoleucine
glycerol
glutamic acid
taurine
putrescine
arachidonic acid
glyceric acid
nicotinamide
ornithine
glutamine
hypoxanthine
phenylalanine
20
15
10
5

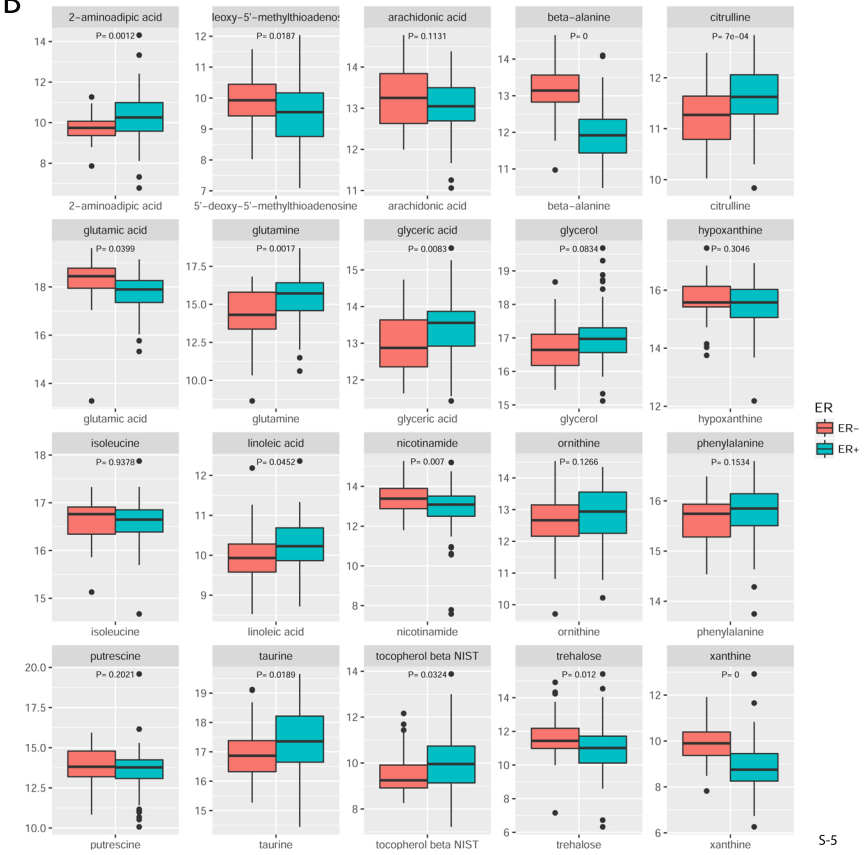

B
2-aminoadipic acid
P= 0.0012
5'-deoxy-5'-methylthioadenosine
P= 0.0187
arachidonic acid
P= 0.1131
beta-alanine
P= 0
citrulline
P= 7e-04
glutamic acid
P= 0.0399
glutamine
P= 0.0017
glyceric acid
P= 0.0083
glycerol
P= 0.0834
hypoxanthine
P= 0.3046
isoleucine
P= 0.9378
linoleic acid
P= 0.0452
nicotinamide
P= 0.007
ornithine
P= 0.1266
phenylalanine
P= 0.1534
putrescine
P= 0.2021
taurine
P= 0.0189
tocopherol beta NIST
P= 0.0324
trehalose
P= 0.012
xanthine
P= 0
ER
ER-
ER+

Figure S4