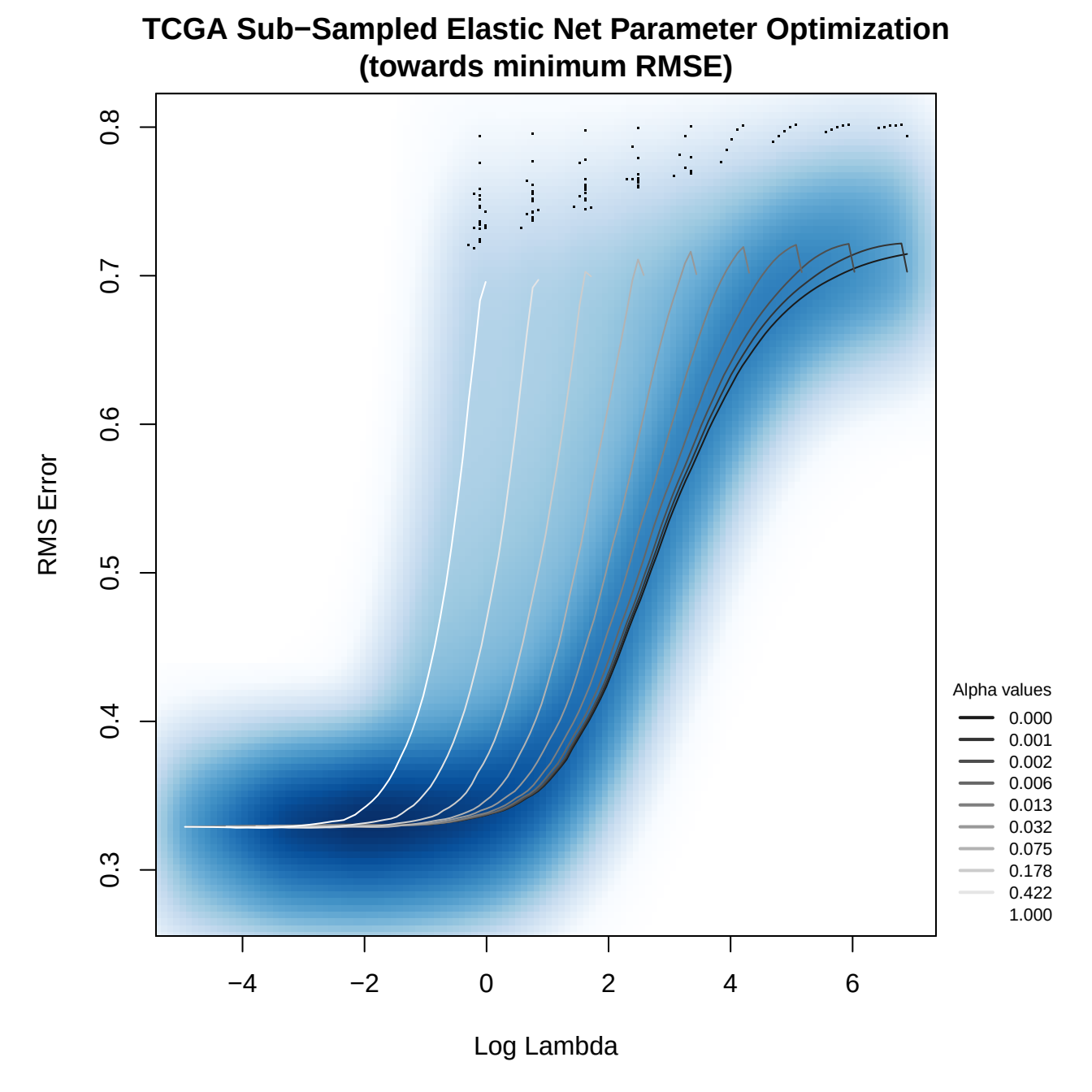
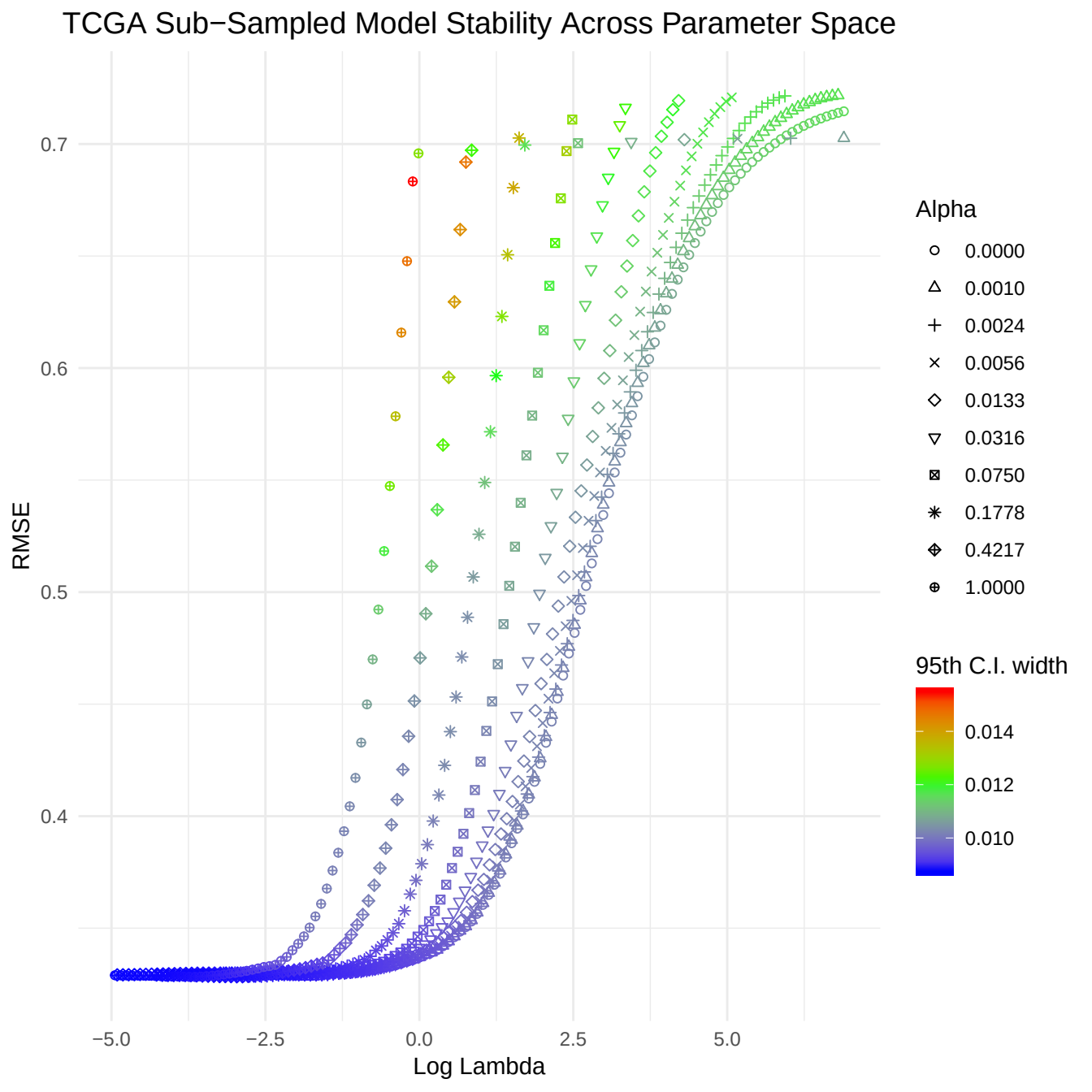
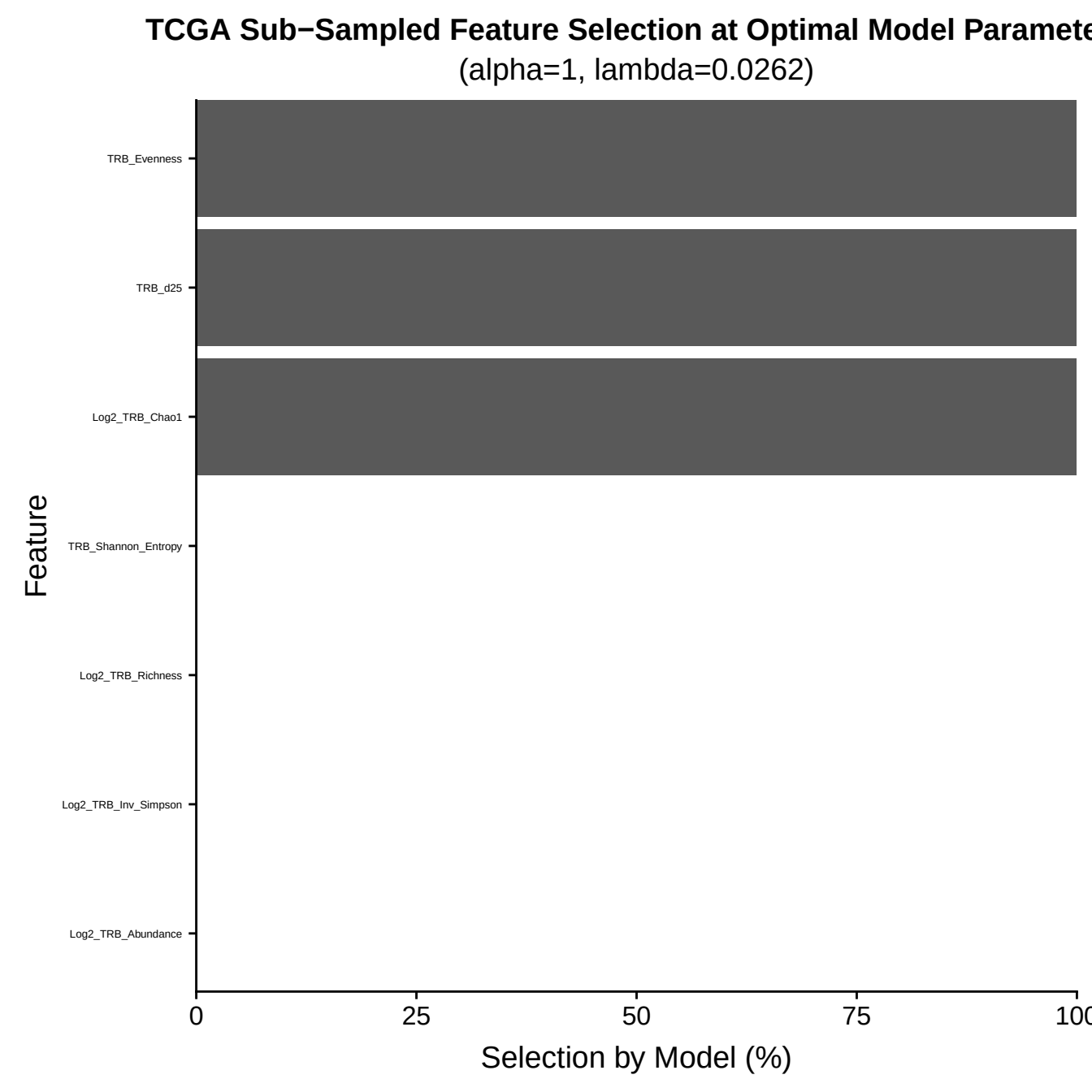
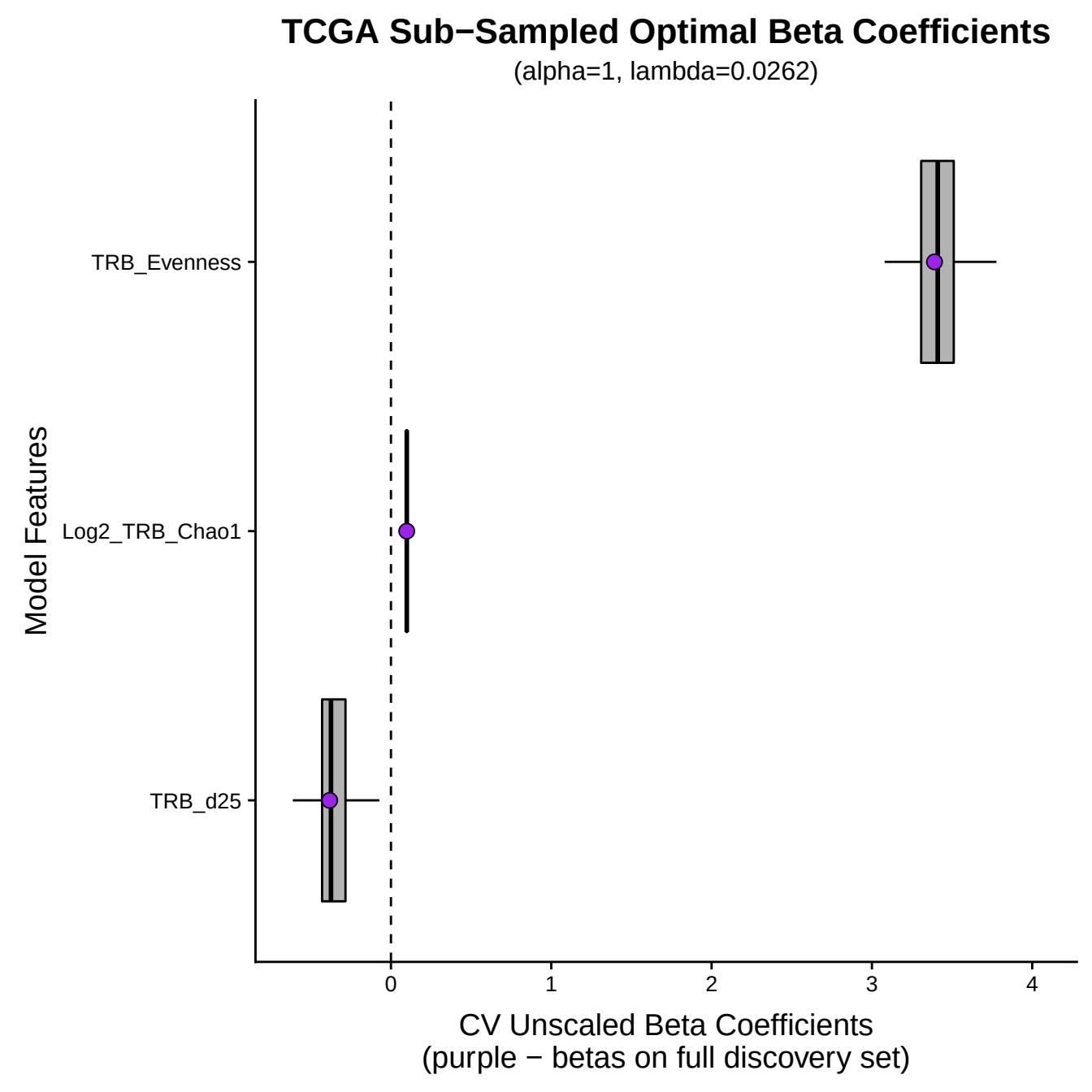
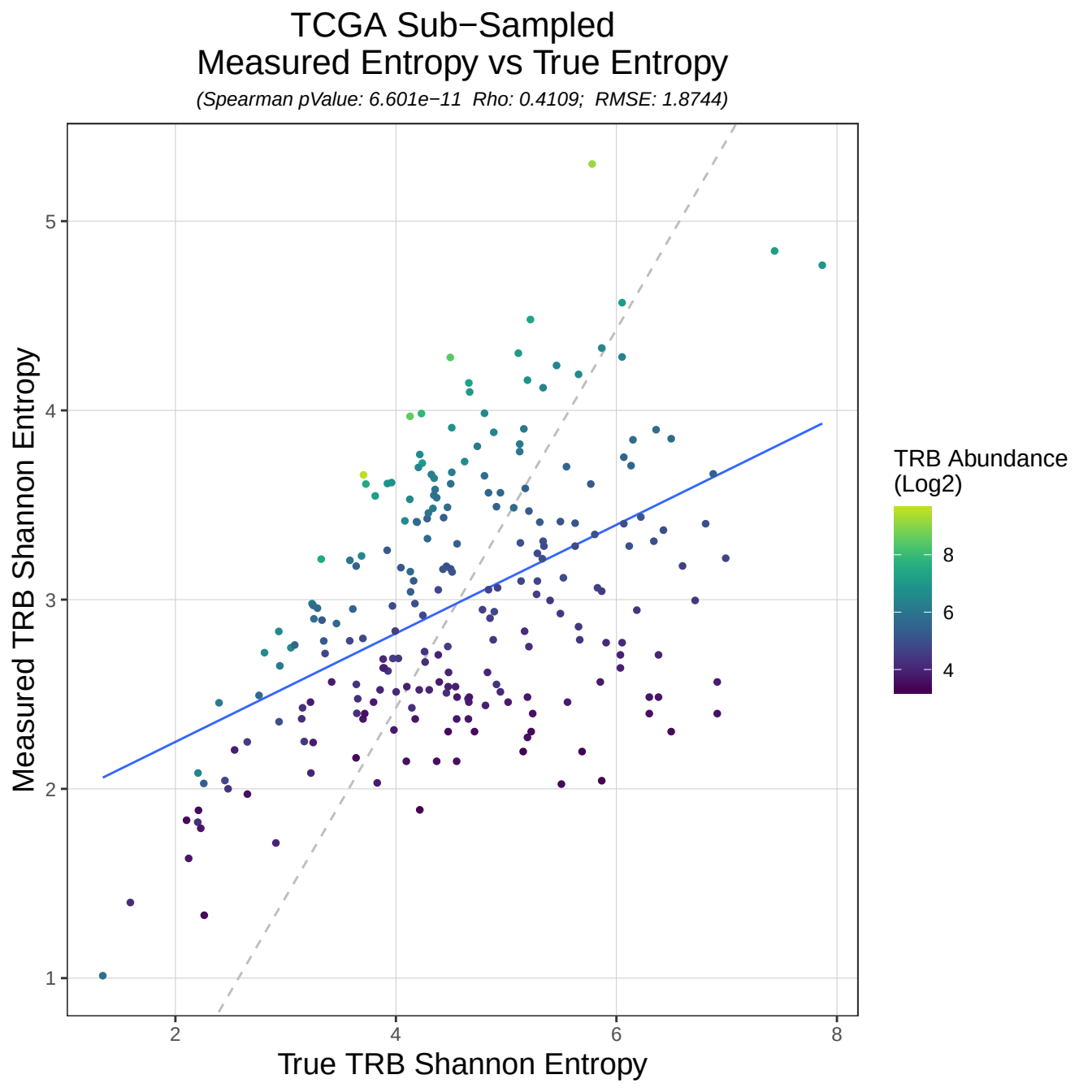
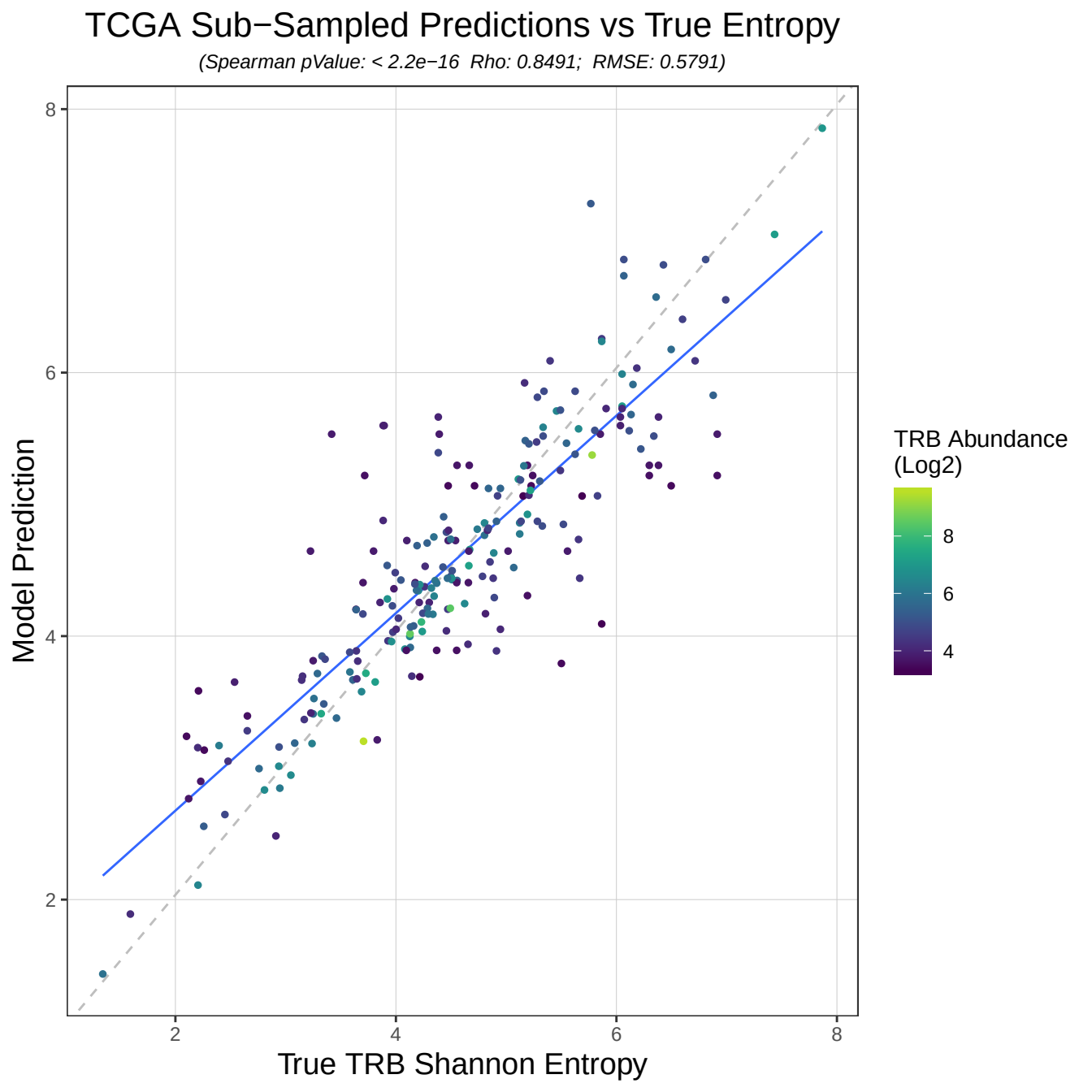

TCGA Sub−Sampled Elastic Net Parameter Optimization
(towards minimum RMSE)
RMS Error
0.3
0.4
0.5
0.6
0.7
0.8
−4
−2
0
2
4
6
Log Lambda
Alpha values
0.000
0.001
0.002
0.006
0.013
0.032
0.075
0.178
0.422
1.000

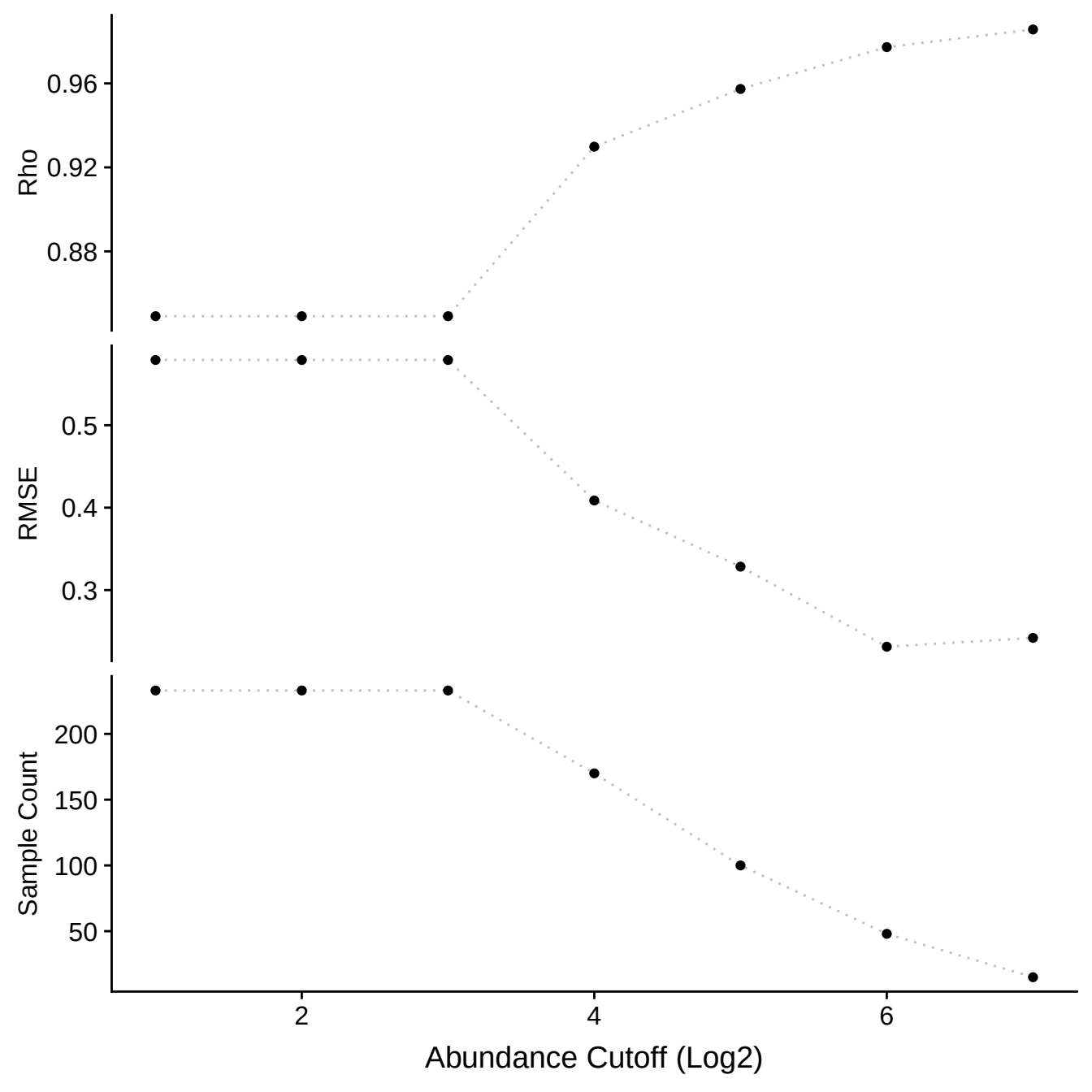

TCGA Sub−Sampled Model Stability Across Parameter Space
RMSE
0.7
0.6
0.5
0.4
−5.0
−2.5
0.0
2.5
5.0
Log Lambda
Alpha
0.0000
0.0010
0.0024
0.0056
0.0133
0.0316
0.0750
0.1778
0.4217
1.0000
95th C.I. width
0.014
0.012
0.010


TCGA Sub−Sampled Feature Selection at Optimal Model Parameter
(alpha=1, lambda=0.0262)
TRB_Evenness
TRB_d25
Log2_TRB_Chao1
TRB_Shannon_Entropy
Log2_TRB_Richness
Log2_TRB_Inv_Simpson
Log2_TRB_Abundance
Feature
0
25
50
75
100
Selection by Model (%)


TCGA Sub−Sampled Optimal Beta Coefficients
(alpha=1, lambda=0.0262)
TRB_Evenness
Log2_TRB_Chao1
TRB_d25
Model Features
0
1
2
3
4
CV Unscaled Beta Coefficients
(purple − betas on full discovery set)


TCGA Sub−Sampled
Measured Entropy vs True Entropy
(Spearman pValue: 6.601e−11 Rho: 0.4109; RMSE: 1.8744)
Measured TRB Shannon Entropy
True TRB Shannon Entropy
1
2
3
4
5
2
4
6
8
TRB Abundance
(Log2)
8
6
4


TCGA Sub−Sampled Predictions vs True Entropy
(Spearman pValue: < 2.2e−16 Rho: 0.8491; RMSE: 0.5791)
Model Prediction
True TRB Shannon Entropy
TRB Abundance (Log2)
2
4
6
8


Rho
0.96
0.92
0.88
RMSE
0.5
0.4
0.3
Sample Count
200
150
100
50
2
4
6
Abundance Cutoff (Log2)