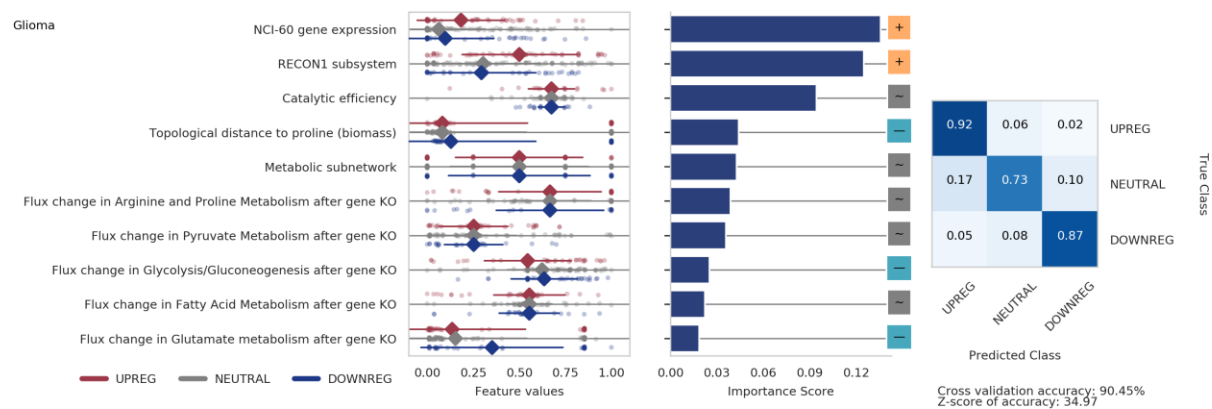
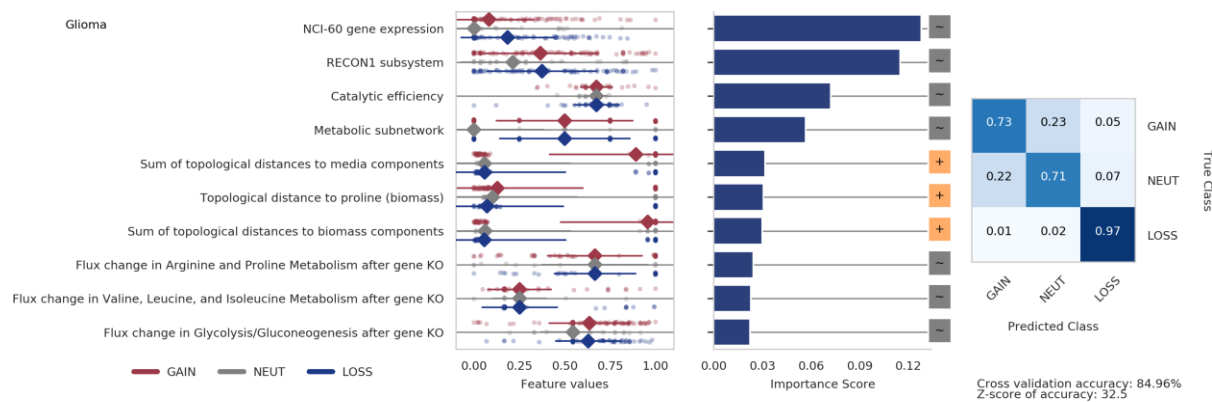
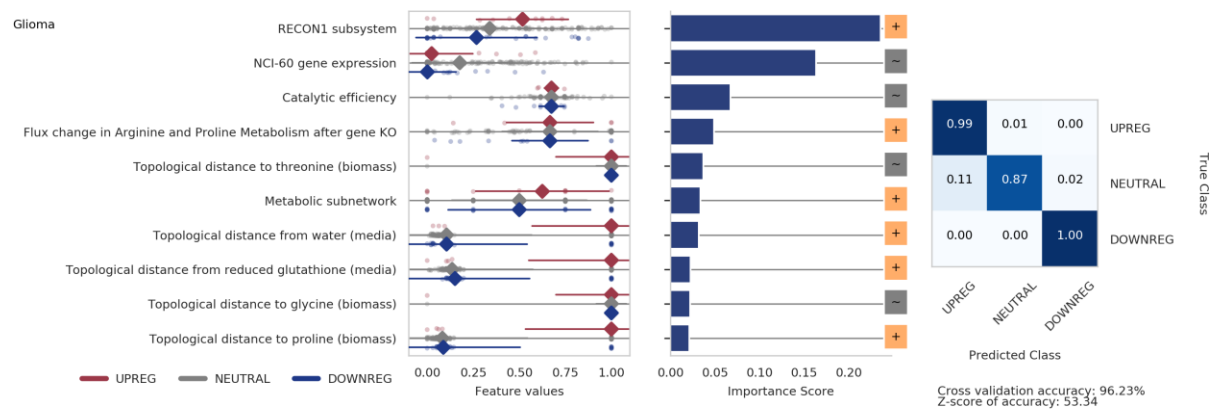


Supplementary Information

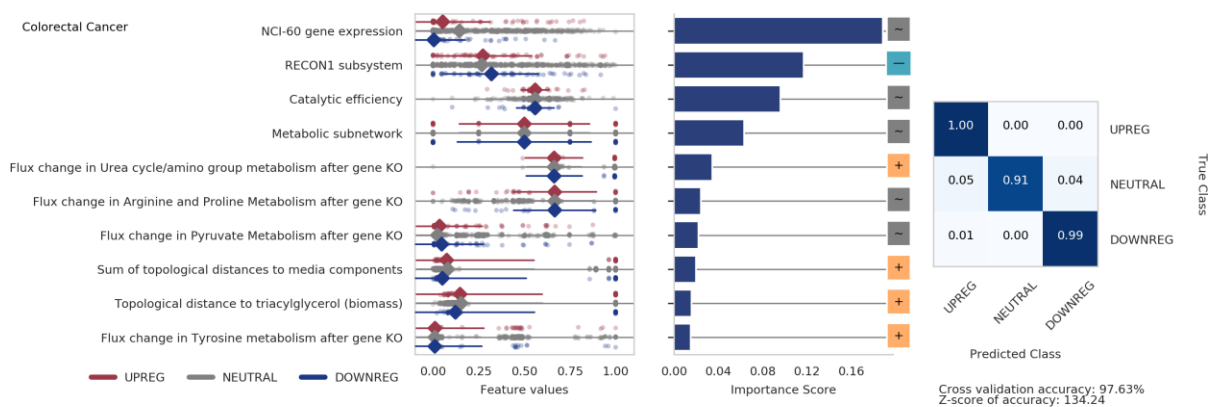
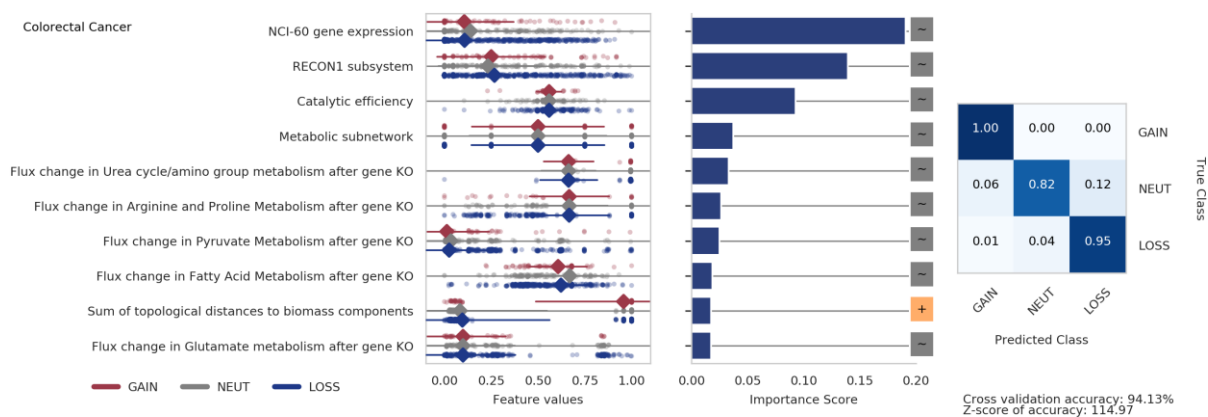
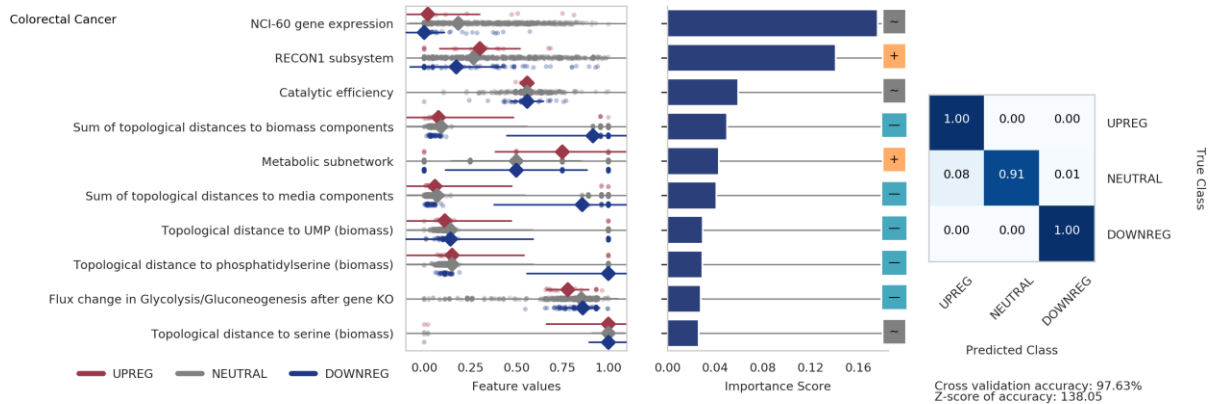


Supplementary Figure 1. MetOncoFit predictions for gliomas

Top | Differential expression: Several topological and metabolic flux features are predictive of dysregulated gene expression. Genes with high topological distances to biomass and medium components - proline and glutathione, were upregulated and those that are closer to these metabolites are downregulated. The 10-fold cross validation accuracy is 96%.

Middle | Copy number variation: Topological distances to biomass and medium components such as the amino acid - proline, are associated with a gain in copy number. 10-fold cross validation accuracy is 85%.

Bottom | Patient survival: Upregulation of genes topologically closer to proline are associated with increased patient survival. Flux through glycolysis and arginine/proline metabolism are also predictive of patient survival. 10-fold cross validation accuracy is 90%.

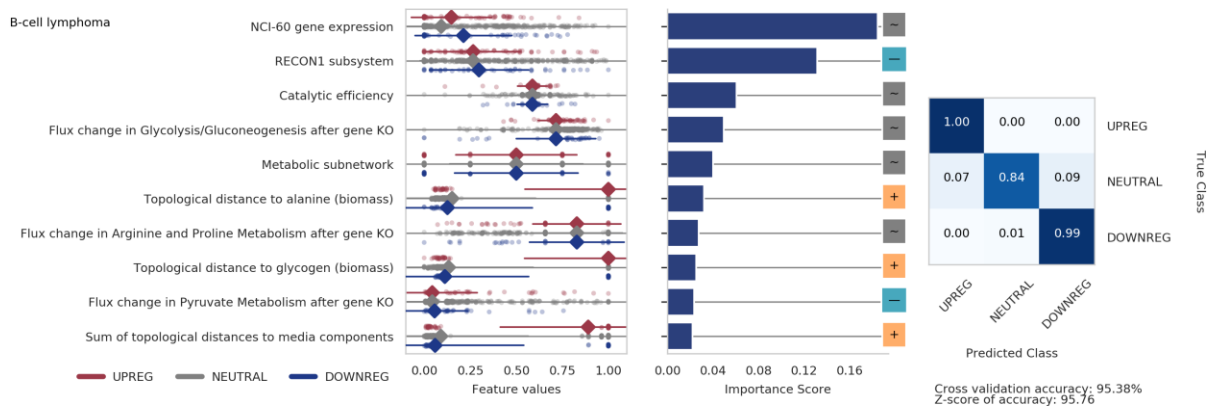
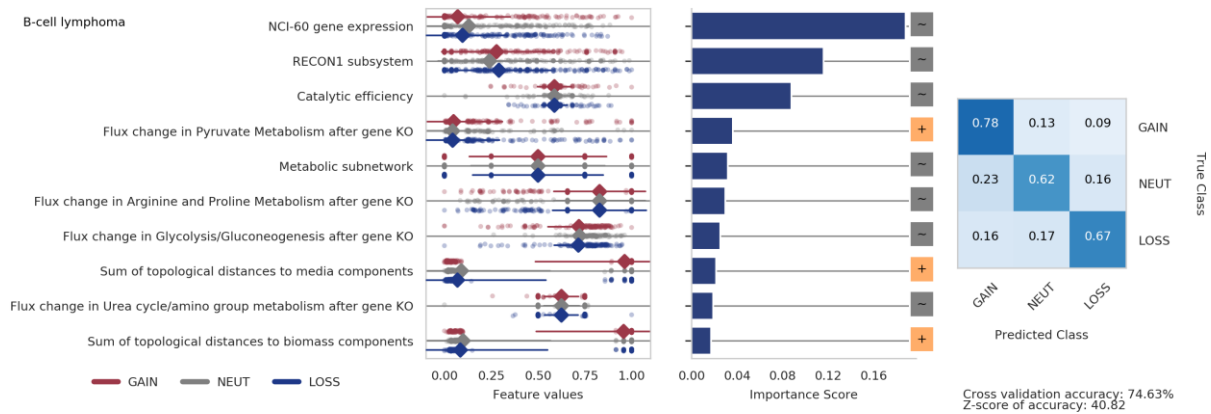
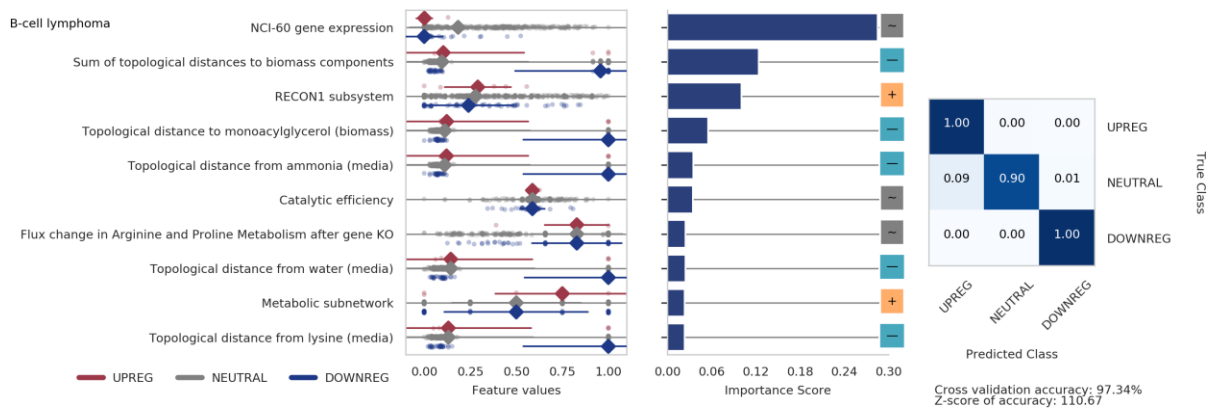


Supplemental Figure 2. MetOncoFit predictions for colorectal cancer

Top | Differential expression: Similar to other cancers, several topological features are associated with differential expression in colon cancer. Genes with low topological distances to biomass components – UMP and phosphatidyl serine, were upregulated. 10-fold cross validation accuracy is 98%.

Middle | Copy number variation: The biomass epicenter score is associated with a gain in copy number, genes that were farther from the network center showed a significant gain in copy number. Fatty acid metabolism showed a non-linear association with copy number, wherein genes that decreased flux through this pathway displayed copy number alterations, which can be either a loss or gain in copy number. 10-fold cross validation accuracy is 94%.

Bottom | Patient survival: Flux change in several metabolic pathways, including urea cycle, arginine and proline metabolism are predictive of patient survival. The sum of topological distances to biomass components are associated with negative patient outcomes; down regulation of genes closer to the network center was associated with positive patient survival. 10-fold cross validation accuracy is 98%.

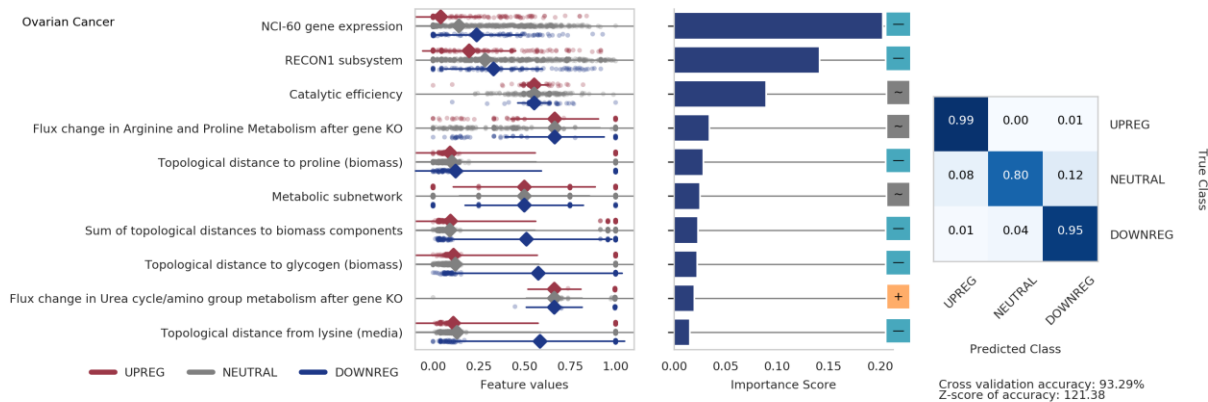
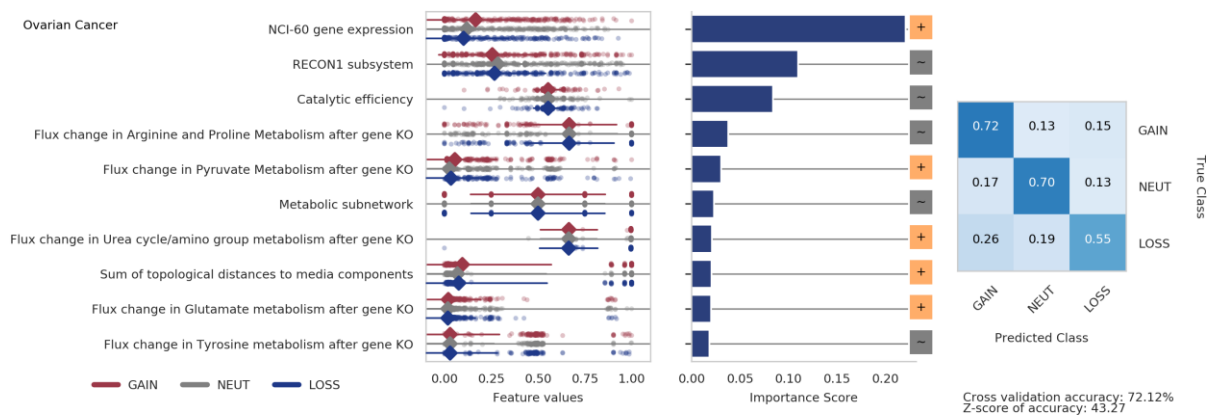
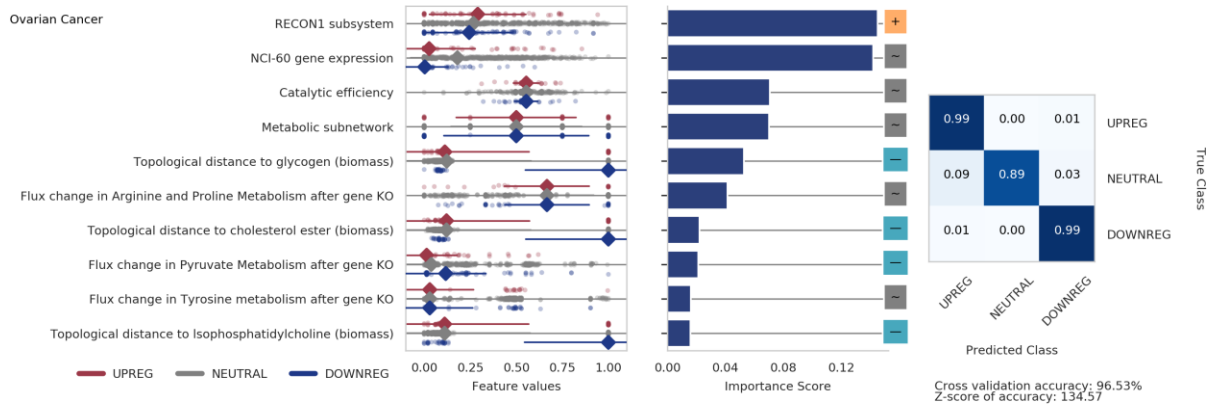


Supplemental Figure 3. MetOncoFit predictions for B-cell Lymphoma

Top | Differential expression: Several topological features are associated with differential expression in B-cell Lymphoma cancer. Genes topologically closer to biomass and media components – lysine, ammonia and monoacyl glycerol, were upregulated. 10-fold cross validation accuracy is 97%.

Middle | Copy number variation: Increased flux through pyruvate metabolism and high biomass and media epicenter scores are associated with a gain in copy number. 10-fold cross validation accuracy is 75%.

Bottom | Patient survival: Flux through pyruvate metabolism is associated with patient mortality; further downregulation of genes topologically closer to alanine and glycogen are associated with positive patient survival. 10-fold cross validation accuracy is 95%.

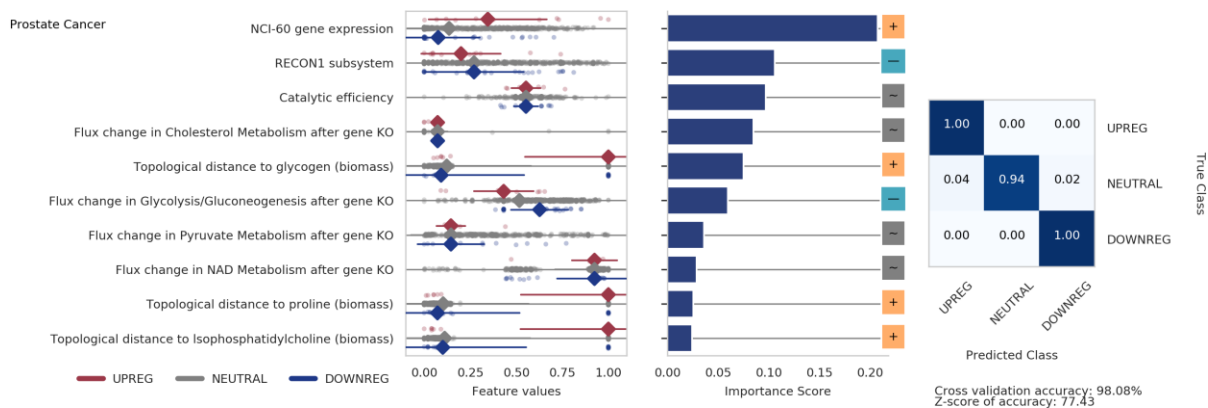
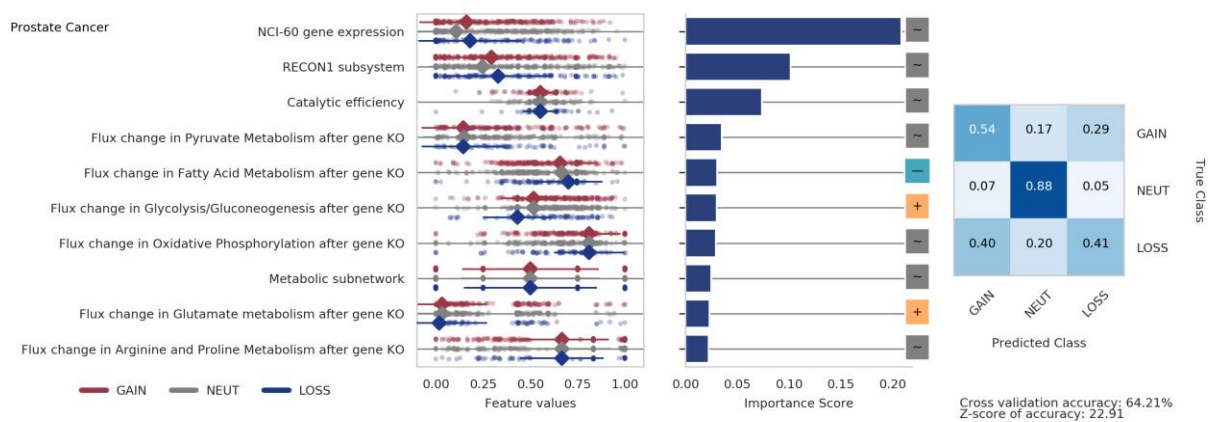
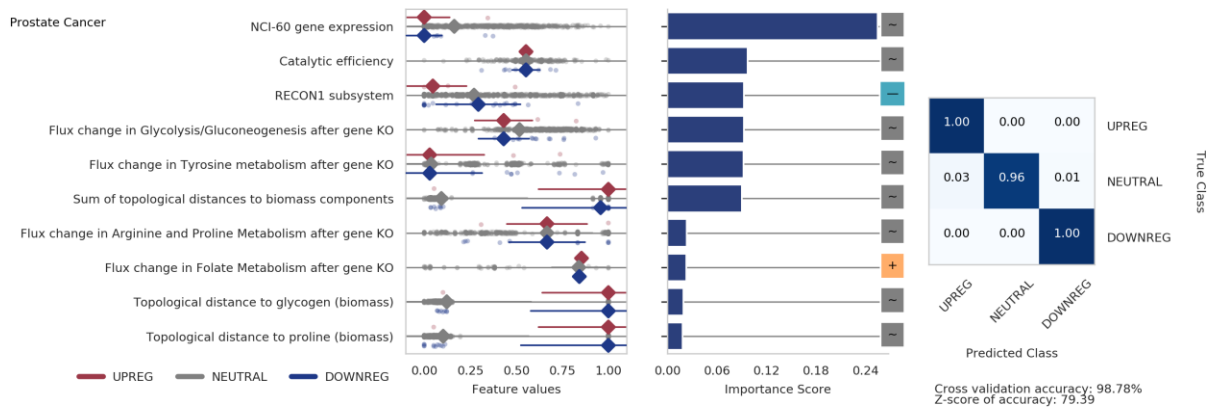


Supplemental Figure 4. MetOncoFit predictions for ovarian cancer

Top | Differential expression: Similar to other cancers, several topological features are associated with differential expression in colon cancer. Genes with low topological distances to biomass components – cholesterol, glycogen and Isophosphatidyl choline, were upregulated. 10-fold cross validation accuracy is 97%.

Middle | Copy number variation: Genes with high media epicenter scores are associated with a gain in copy number. 10-fold cross validation accuracy is 72%.

Bottom | Patient survival: Metabolic flux through arginine and proline metabolism, and urea cycle are predictive of patient survival. 10-fold cross validation accuracy is 93%.

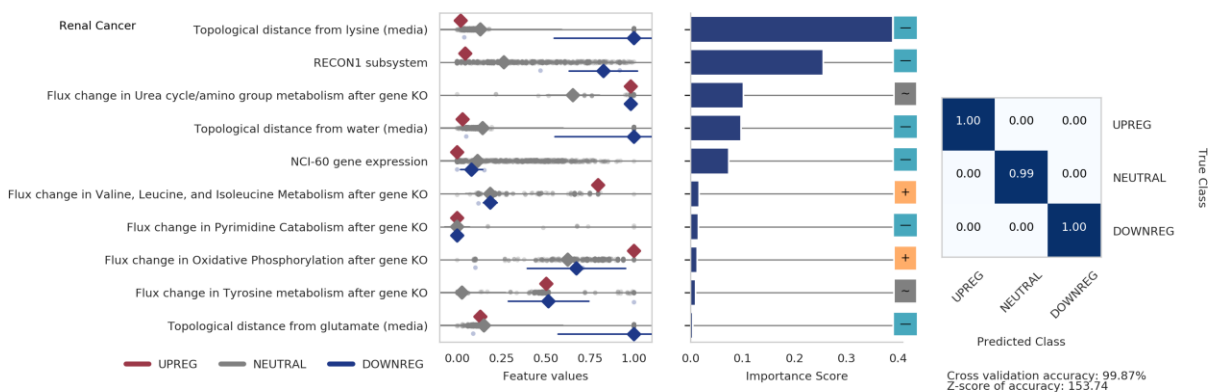
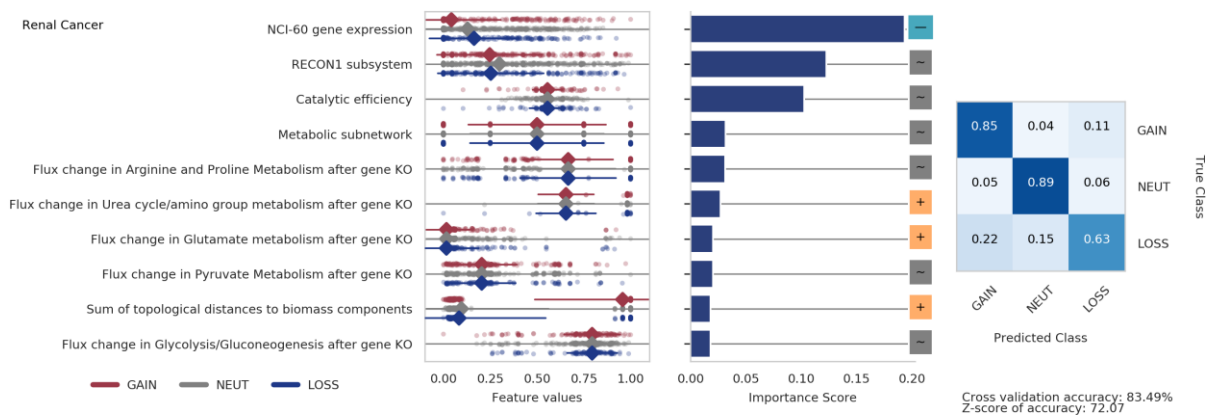
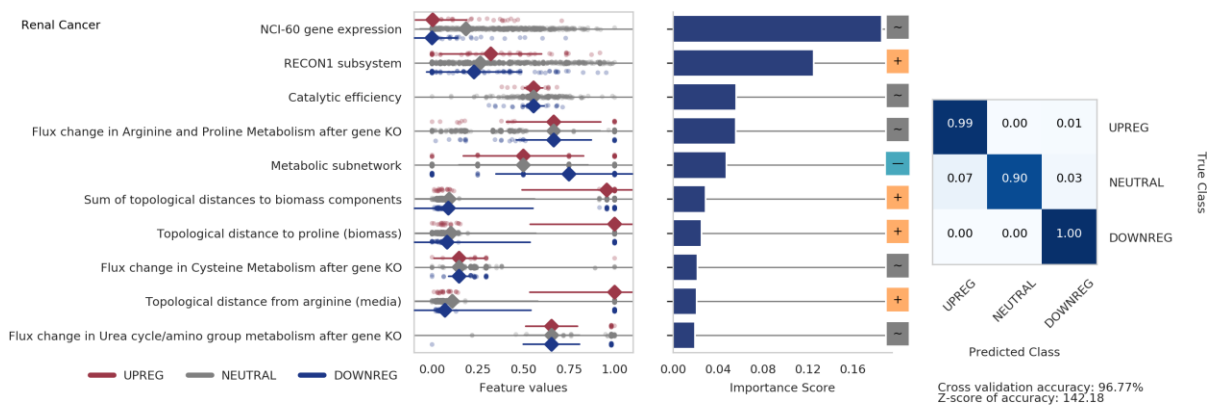


Supplemental Figure 5. MetOncoFit predictions for prostate cancer

Top | Differential expression: Flux through folate metabolism is associated with gene upregulation. Glycolytic flux showed a non-linear association with differential gene expression, wherein genes that decreased flux through this pathway were either up- and down-regulated. Hence flux through this pathway is predictive of dysregulation but not in a direct linear fashion. 10-fold cross validation accuracy is 99%.

Middle | Copy number variation: Increased flux through glutamate metabolism and glycolysis is associated with a gain in copy number. Metabolic flux through fatty acid metabolism is associated with copy number loss. 10-fold cross validation accuracy is 64%.

Bottom | Patient survival: Several topological features are associated with both positive patient outcomes and patient mortality. Downregulation of genes with low topological distances to biomass components – proline, glycogen and Isophosphatidyl choline, is associated with better patient survival. 10-fold cross validation accuracy is 98%.

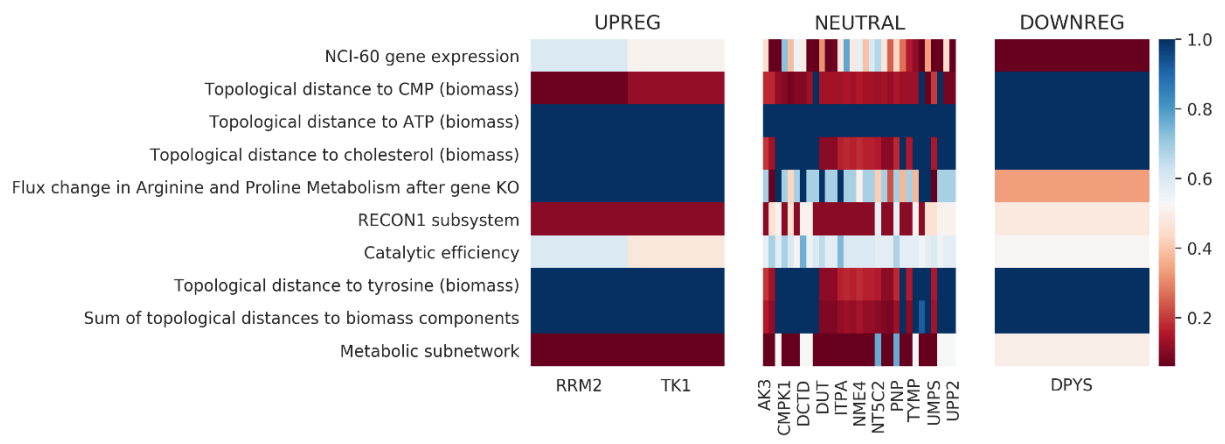


Supplemental Figure 6. MetOncoFit predictions for renal cancer

Top | Differential expression: Several topological and metabolic flux features are associated with gene upregulation. For example, genes away from the network center (i.e. high topological distance to all biomass components) were likely to be upregulated, while those closer to the amino-acid - proline are downregulated. 10-fold cross validation accuracy is 97%.

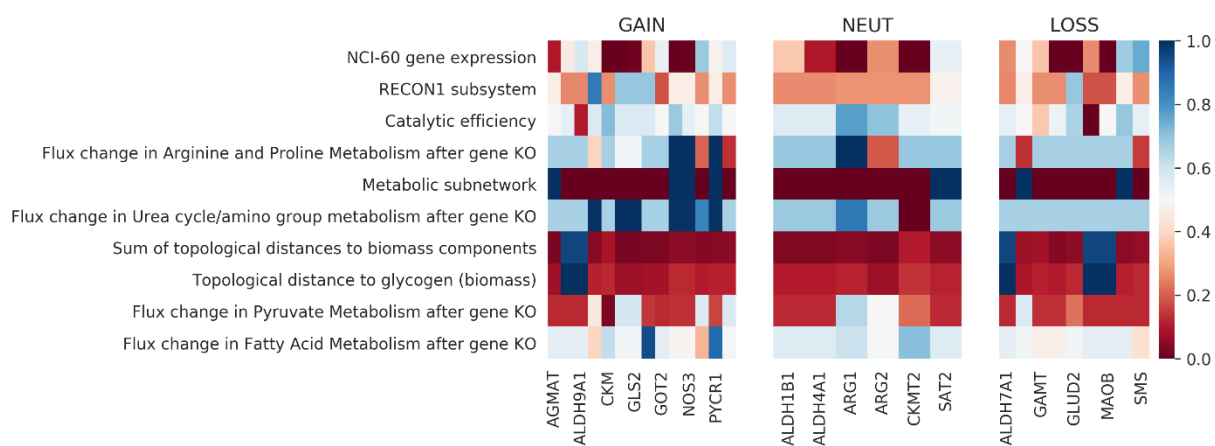
Middle | Copy number variation: Flux through arginine and proline metabolism, urea cycle and glutamate pathway are predictive of copy number alterations. Similar to other cancers, genes away from the network center were likely to show a gain in copy number. 10-fold cross validation accuracy is 83%.

Bottom | Patient survival: Upregulation of genes with low topological distances to biomass components – lysine, water and glutamate, is associated with better patient survival. Increased flux through oxidative phosphorylation pathway is also predictive of increased patient survival. 10-fold cross validation accuracy is 99%.



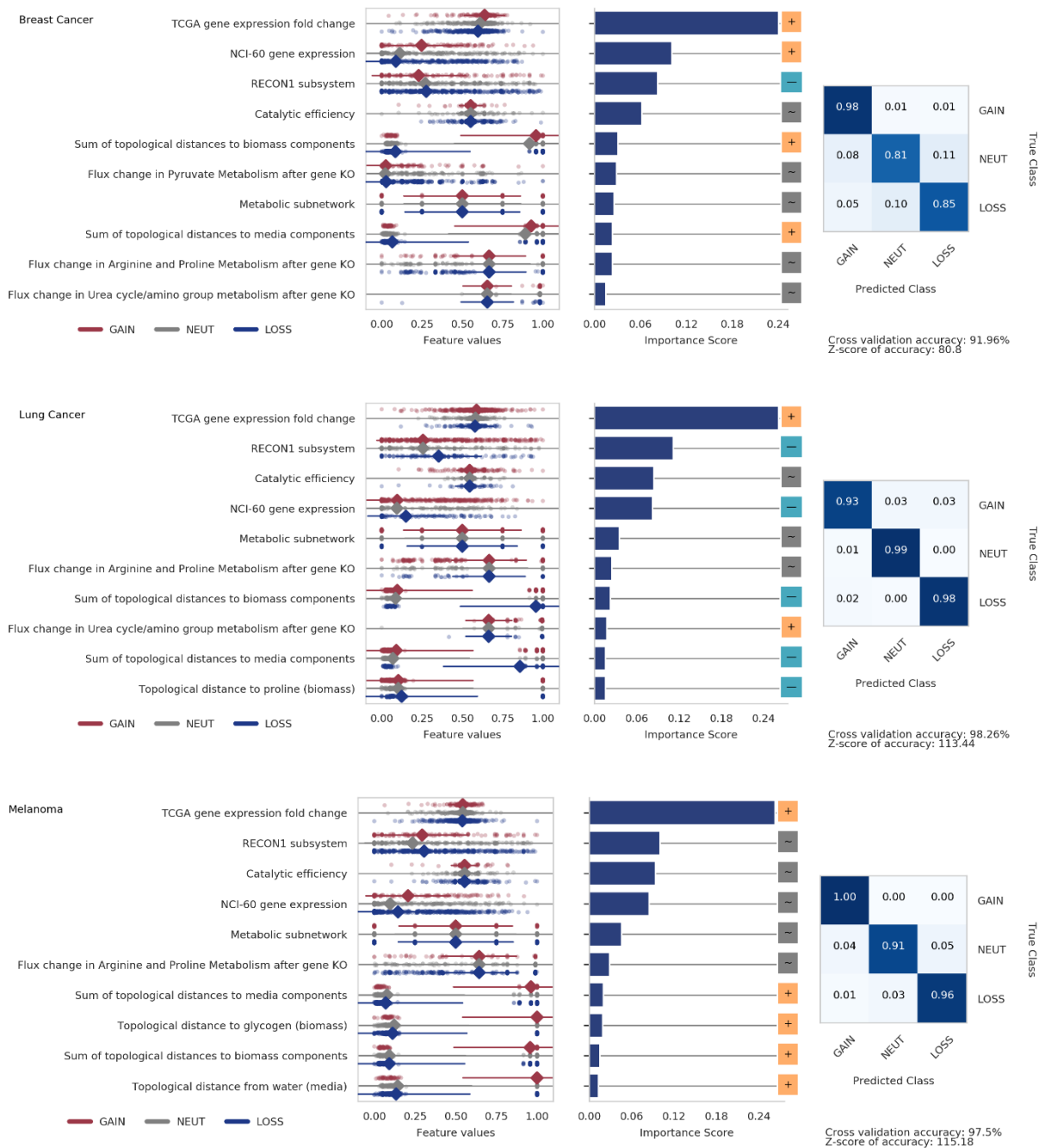
Supplemental Figure 7. MetOncoFit correctly predicts upregulation of RRM2

Genes selected in the heatmaps are from the KEGG gene set for Pyrimidine metabolism. MetOncoFit correctly predicted that RRM2 is dysregulated based on **a)** its short topological distance to CMP, **b)** upregulation in breast cancer cell lines, and **c)** high catalytic efficiency.



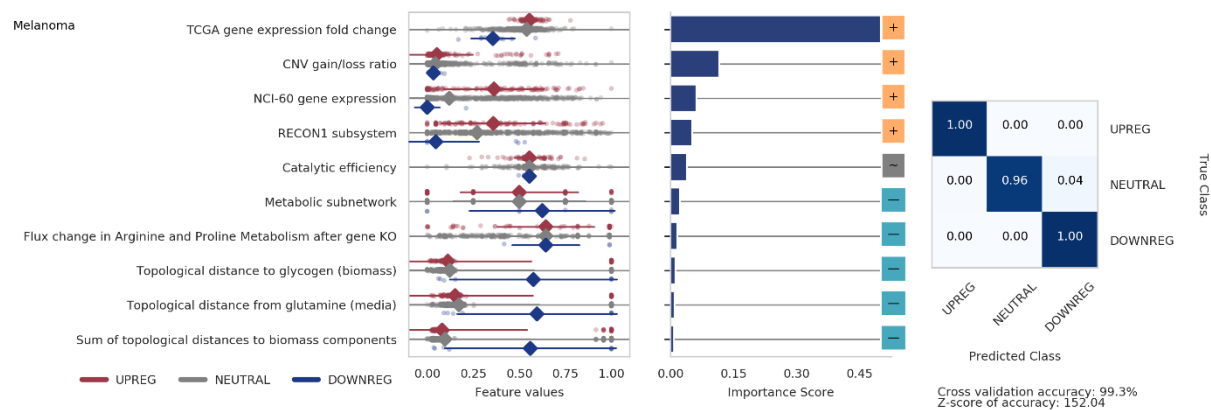
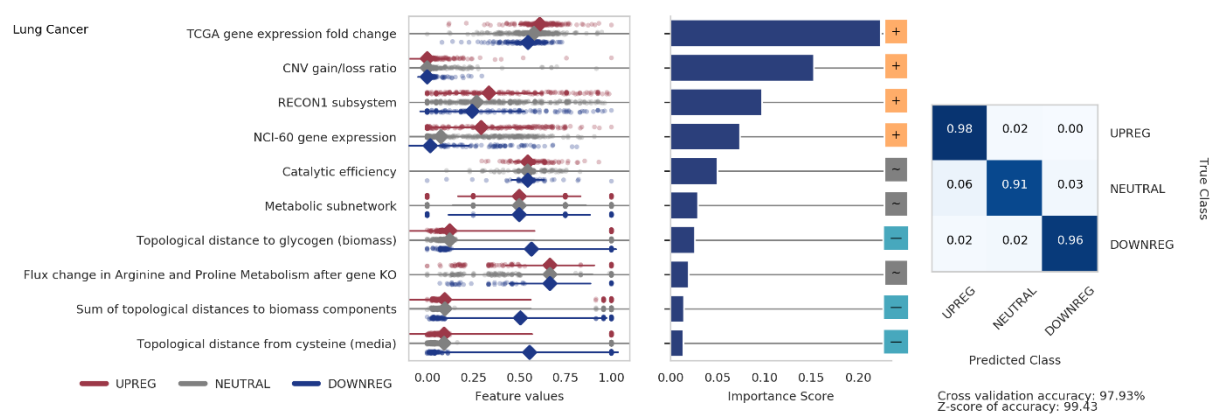
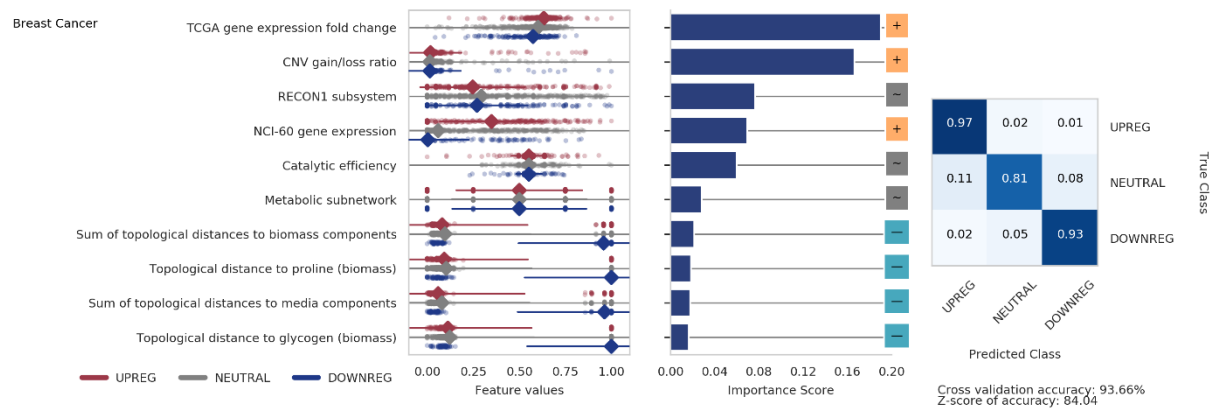
Supplemental Figure 8. Gain in copy number for metabolic genes in the urea cycle is a recurring metabolic rewiring strategy in NSCLC

Genes selected in the heatmaps are from the KEGG gene set for arginine and proline metabolism, which contains the urea cycle. Several genes that directly participate in or are proximal to the urea cycle, including GLS2, GOT2, and NOS3 display a gain in copy number and have high metabolic flux through the urea cycle.



Supplementary Figure 9. Incorporating TCGA gene expression fold change into the cancer models improves MetOncoFit's predictive performance for copy number variation.

High TCGA gene fold change was correlated with a gain in metabolic gene copy number for breast cancer, NSCLC, and melanoma. This analysis revealed several topological features including topological biomass epicenter scores, and metabolic flux through several pathways including arginine and proline metabolism as predictive of copy number variation in these cancers.



Supplementary Figure 10. Integrating TCGA fold change expression and copy number gain/loss ratios improve MetOncoFit predictions for cancer patient survival

TCGA gene expression fold change and copy number values are the top predictors of cancer patient survival in breast cancer, NSCLC, and melanoma. NCI-60 gene expression, The topological biomass epicenter score, catalytic efficiency and flux through arginine and proline metabolism are also top predictors across all cancers.