ΕΘΝΙΚΟ ΙΔΡΥΜΑ ΕΡΕΥΝΩΝ
National Hellenic Research Foundation

e-NiOS
e-Noesis Inspired Operational Systems

Radiogenomics and Colon Cancer: Results of a Collaborative DKFZ-NHRF Study

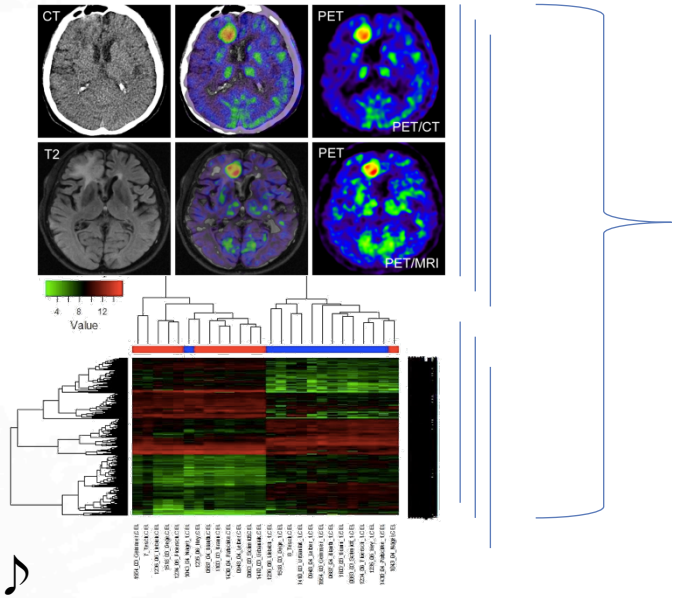
Aristotelis Chatziioannou
Principal Investigator
Metabolic Engineering & Bionformatics Group, IBMCB/NHRF
CEO e-NIOS Applications PC

PJTECH CATALYST

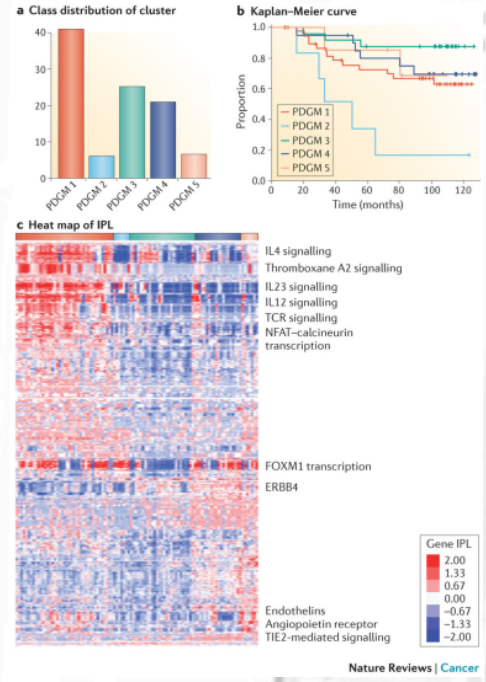
ACCC Scientific Symposium, Athens, GR, Dec, 8th 2017

iqbility
Driving Ideas ahead

Radiogenomics:A novel path for efficient disease characterization



Artificial Intelligence/ Big Data Computing



♪

- Digitized images (CT,MR, PET) more than pictures, they are quantitative data♪

♪

- Oncology :key application area, as patients undergo routinely either imaging and/or neat molecular characterization (tumor panels, sequencing of tumor samples)♪

- Integration of both layers, aids robust patient stratification, reveals causal phenotype–genotype associations♪

Composite signature approach:♪

♪

- Interprets patients' high–throughput omic data, through advanced AI techniques, so as to derive highly descriptive, compact signatures♪

♪

- Optimization of the initial molecular signatures, through epidemiological evaluation♪

- Signature–oriented, association with imaging/ clinical features for robust phenotypic analysis of patient cohorts, for better diagnostic / therapeutic stratification♪

CRC Dataset profile

- **30 patients** diagnosed with CRC provided by the German Cancer Research Center, Heidelberg
- 2 different Affymetrix microarray platforms: **HG-U133A** (**13** patients-**26** matched samples) & **HG-U133plus2** (**17** patients-**34** matched samples)-60 total samples.
- Examined samples acquired from frozen tissue sections of surgically removed primary colon tumor from each patient and adjacent normal mucosa
- 6 of the patients had synchronous metastases.
- Molecular data complemented by kinetic measurements of glucose uptake rate by cancer cells, as extrapolated by PET measurements.

Statistical Selection

- Non-metastatic cancer samples vs. control samples (**1853 DE genes**)
- "metastatic" vs. adjacent controls (**1166 DE genes**)
- "total" CRCs versus controls (**1760 DE**)

↓

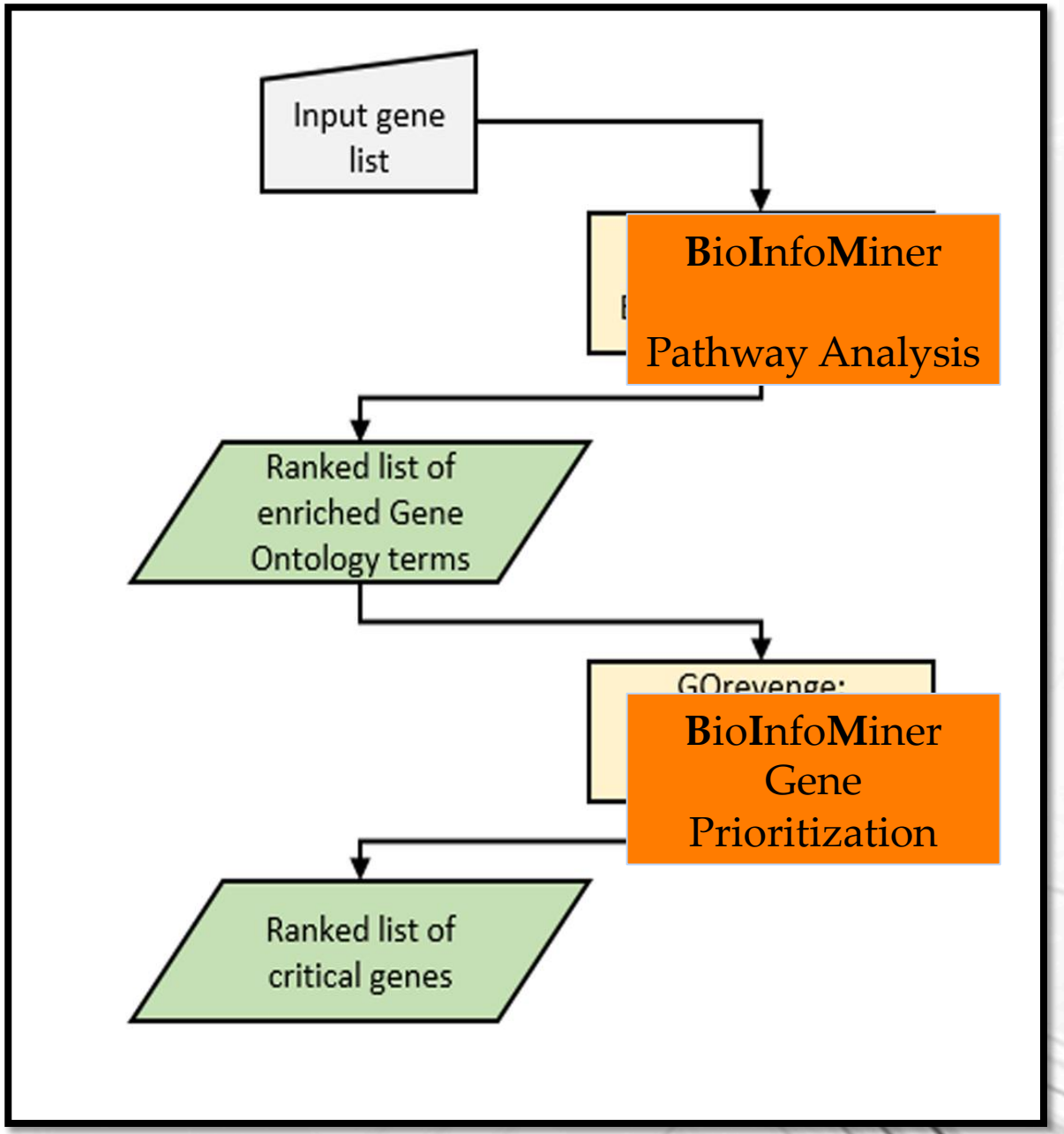
- Common DE genes among all 3 comparisons: **911 genes**

↓

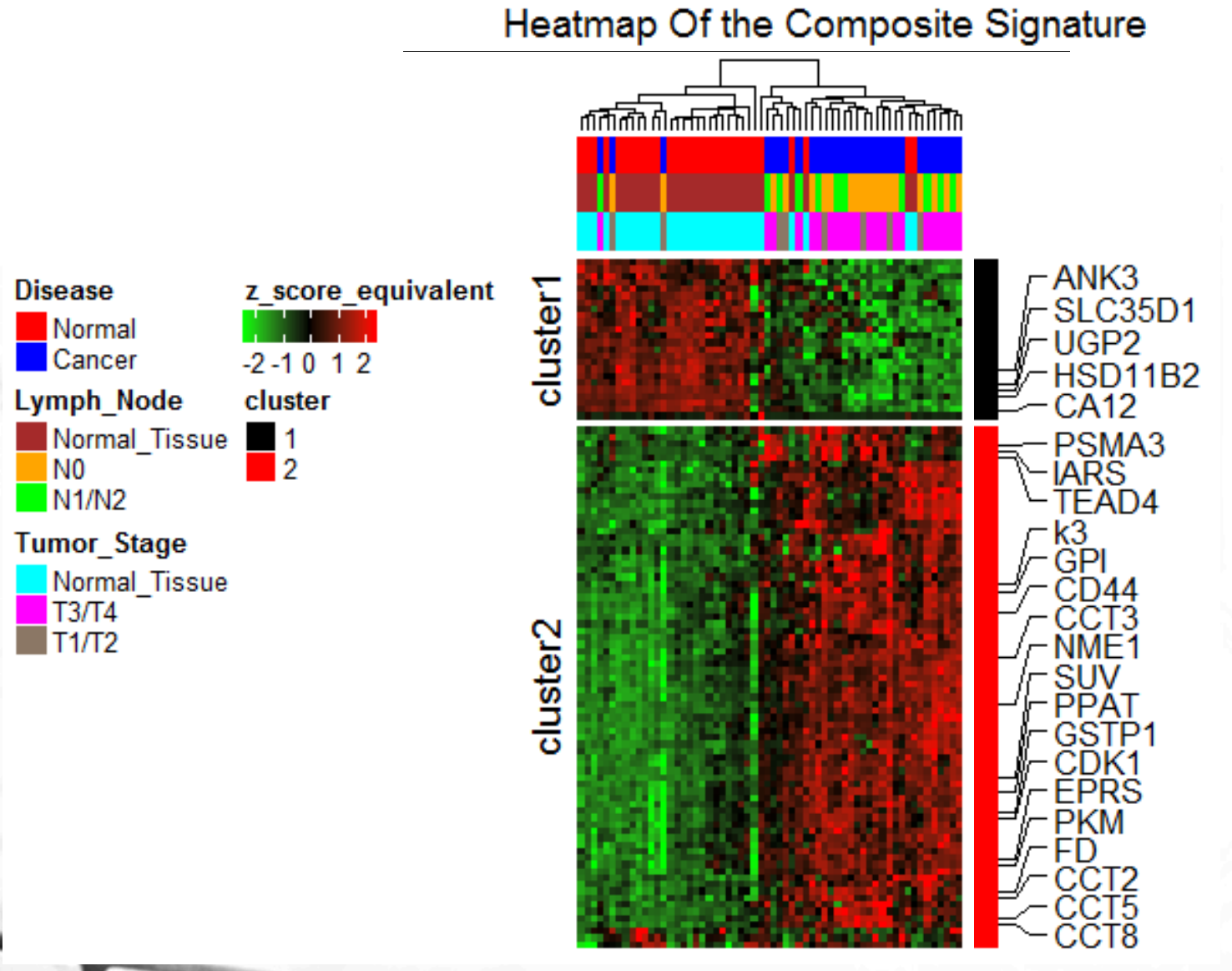
- Exploitation of 3 distinct Ontological vocabularies (Gene Ontology,REACTOME, MGI)

↓

- Final list of hub genes from all 3 topological nets: **94 genes**



Integration with PET kinetic data
Composite signature of 102 features



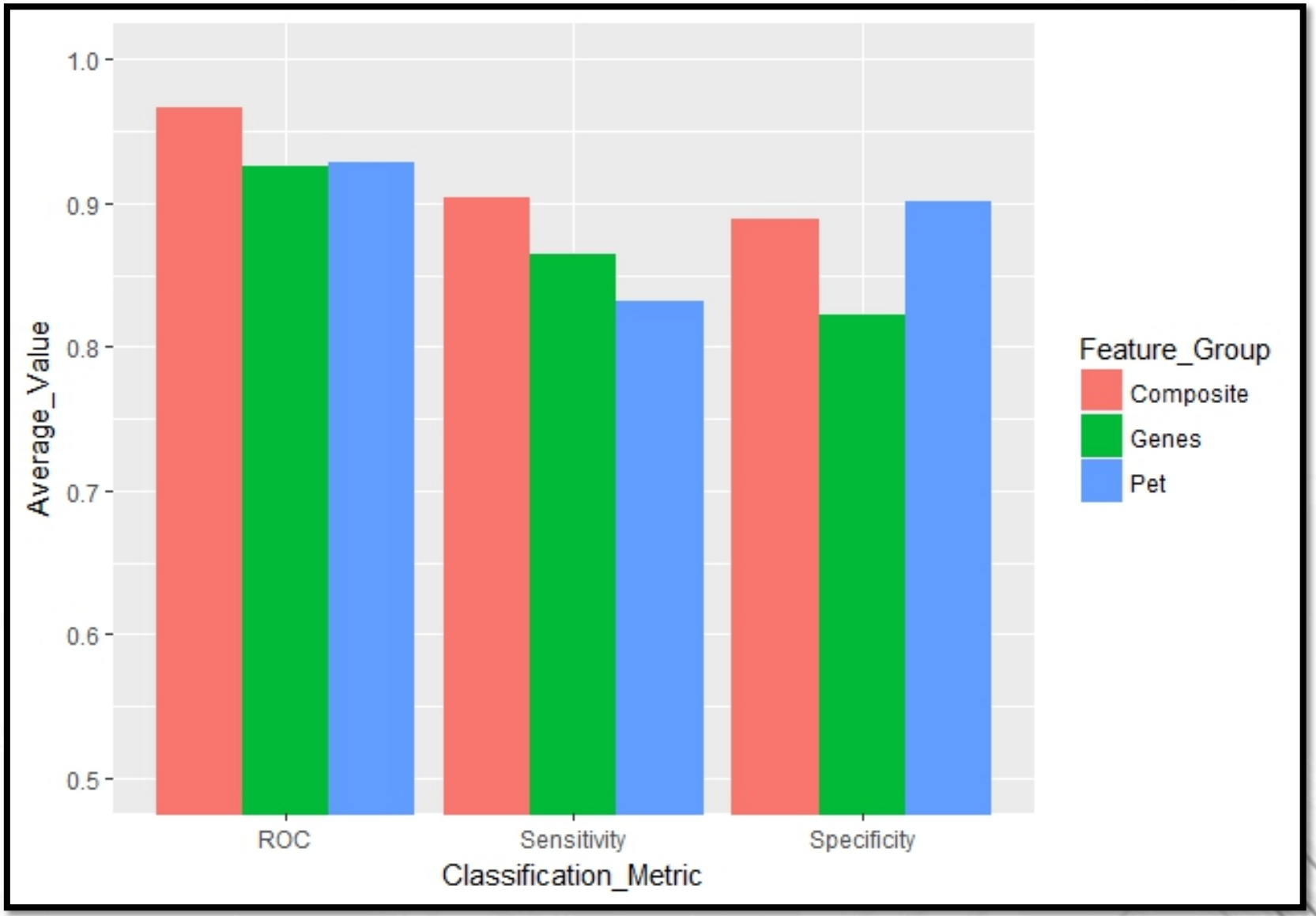
Genes showing negative correlation with PET variables:

- ✓ UDP-glucuronate biosynthetic process
- ✓ carbohydrate derivative biosynthetic process

Genes showing positive correlation with PET variables:

- ✓ positive regulation of telomere maintenance via telomerase
- ✓ positive regulation of telomere maintenance via recombination
- ✓ cell-cell recognition

Evaluation of the relative performance of the three initial signatures (102-Composite, 94-Gene, 8-PET-kinetic)



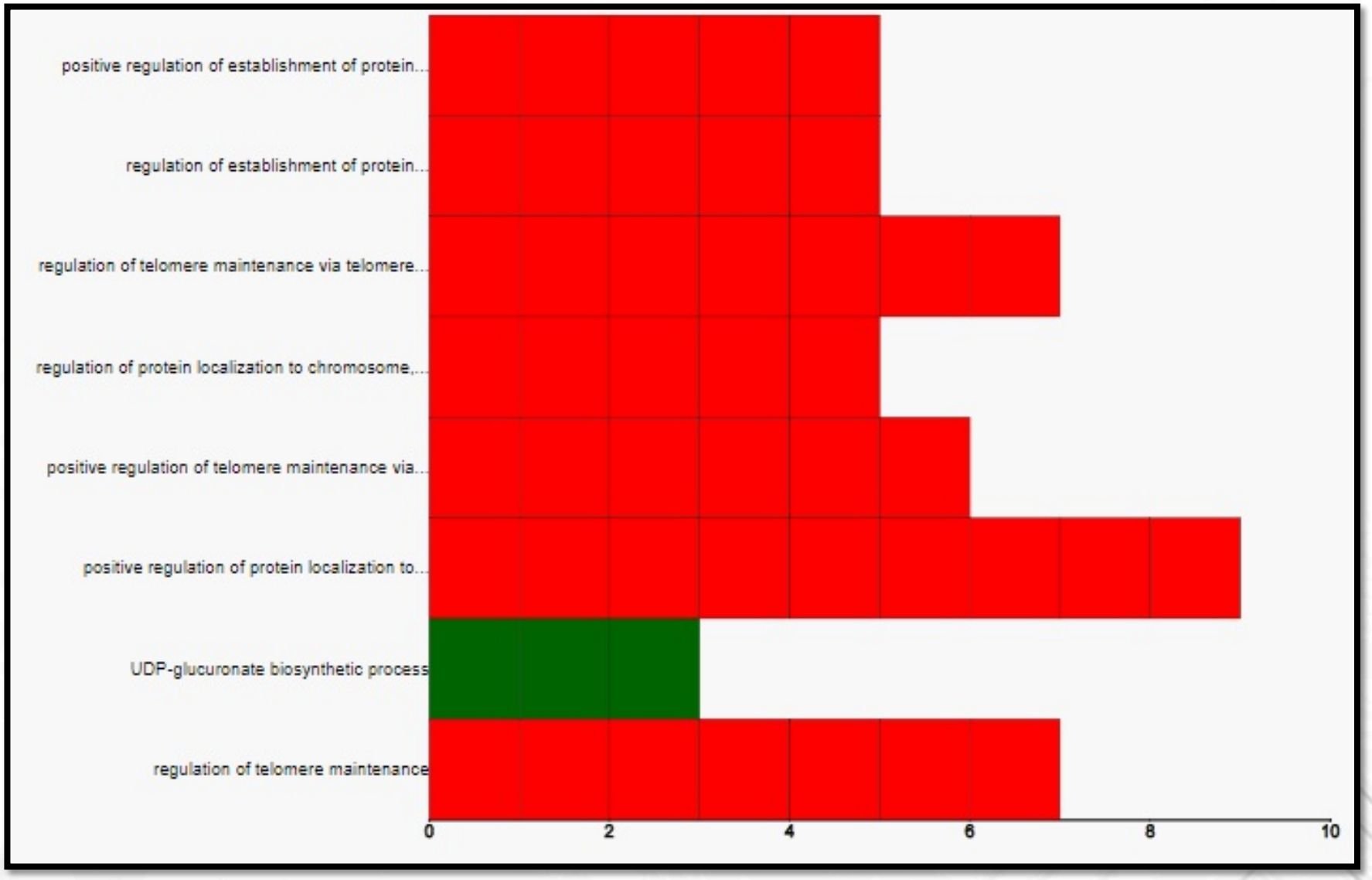
Association of gene expression and clinical data

GPI: Pluripotent functionality, glycolysis,, tumor-motility, angiogenesis

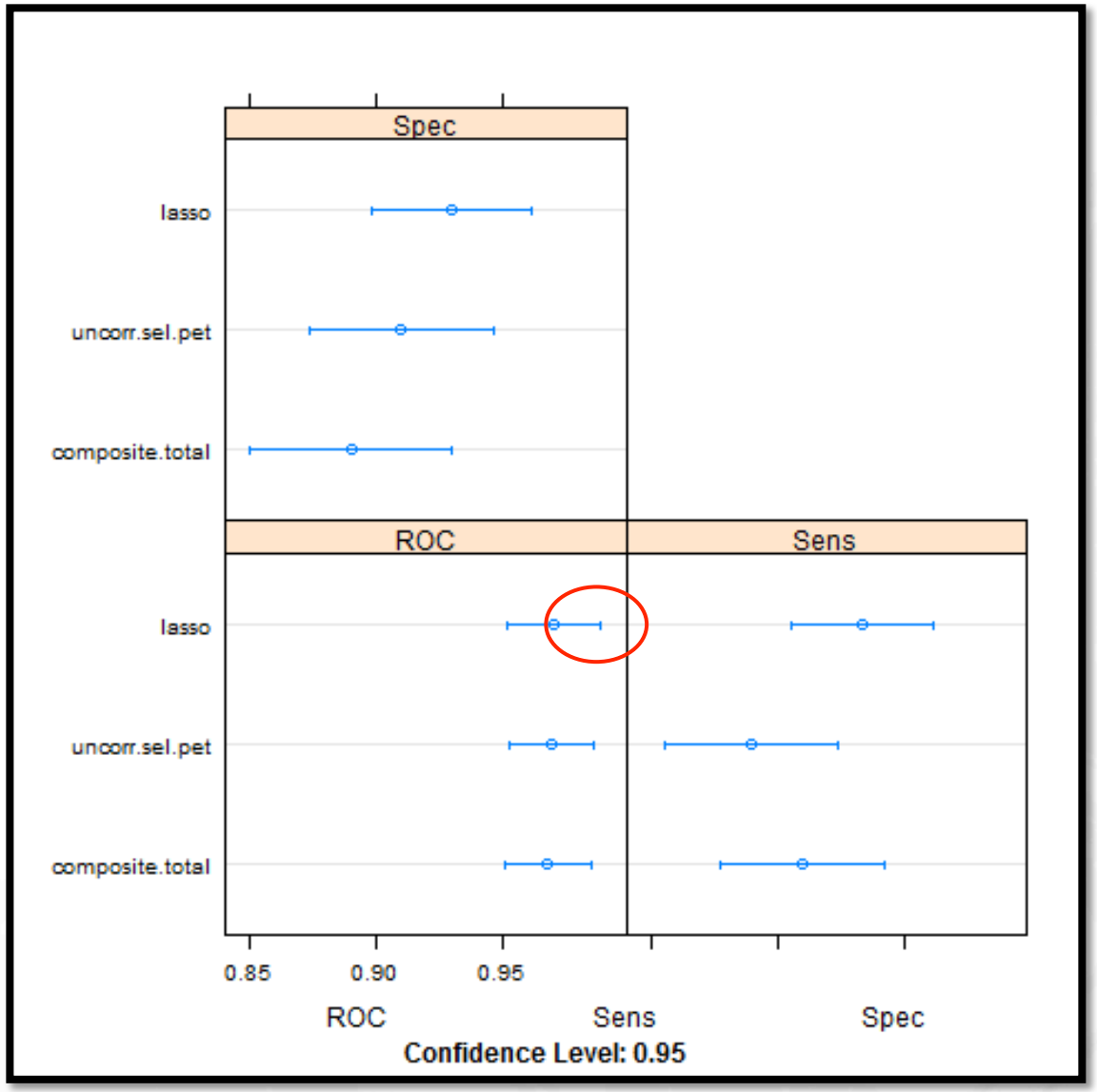
CD44: Antigen presentation, immunomodulation, Stem cell marker

UGP2: Important intermediary in carbohydrate metabolism, involved in gluconeogenesis, lactogenesis

Functional implication of 4 key PET parameters (FD, SUV, k3, INF) -- highly correlated with 46 out of the 94 genes



Optimization of the CRC Composite Signature

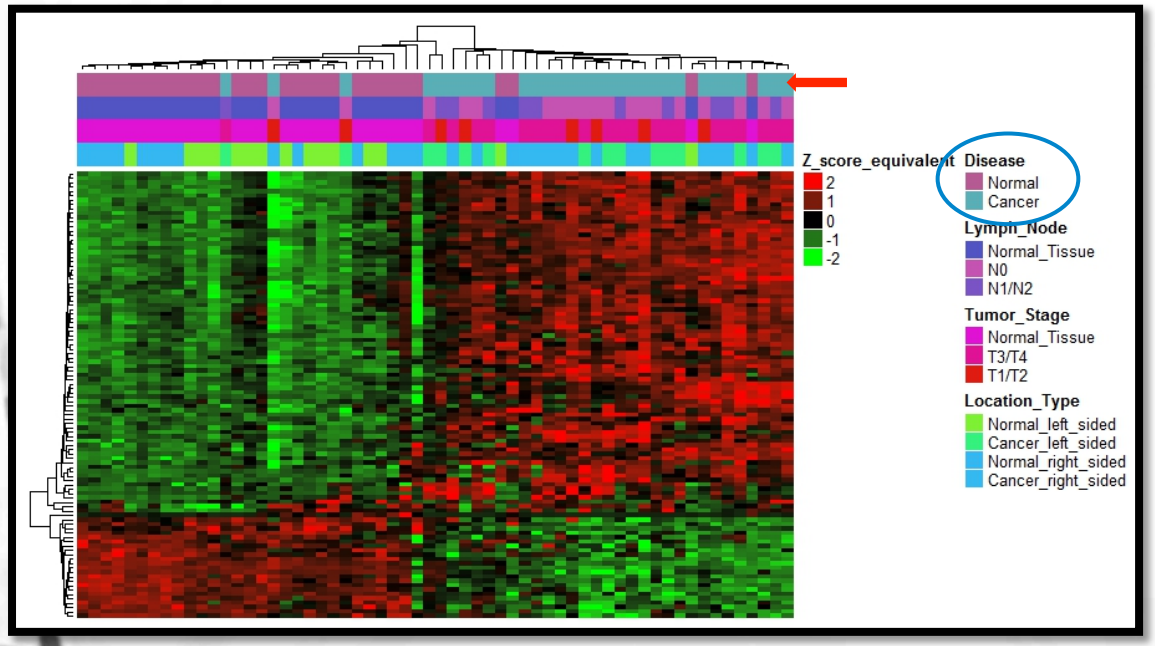


- Dimensionality reduction of the initial composite signature
- Regularization of the signature in order to select features with distinct role
- Important for application of targeted therapeutic approaches

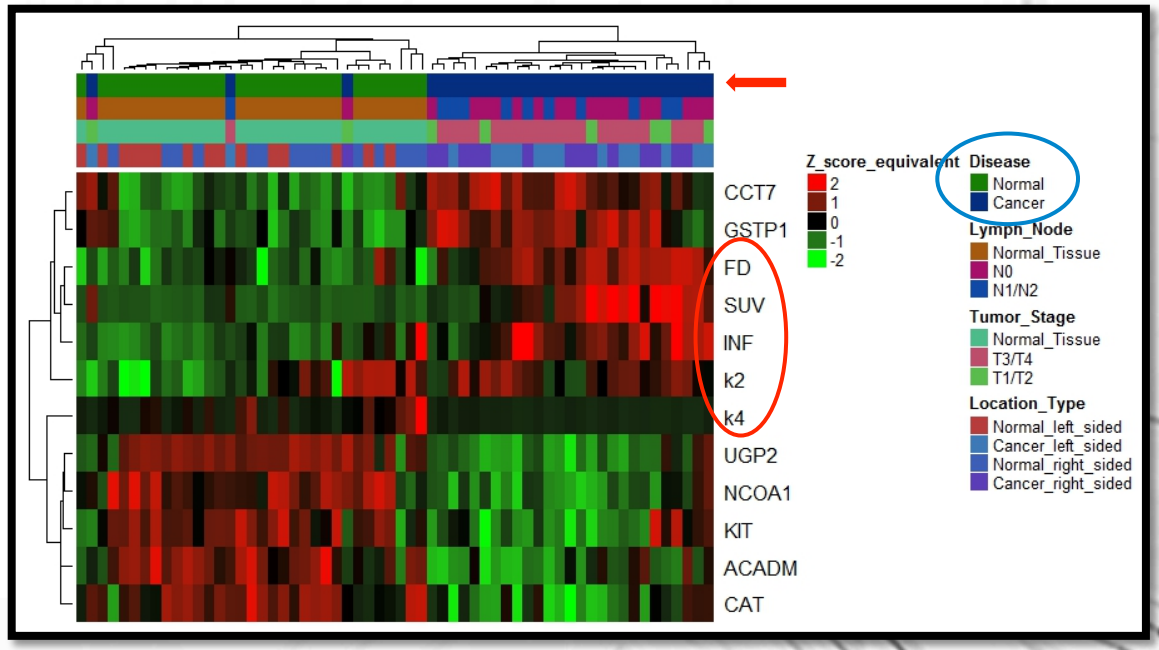
ACCC Scientific Symposium, Athens, GR, Dec, 8th 2017

Comparative evaluation of initial and compact composite signatures

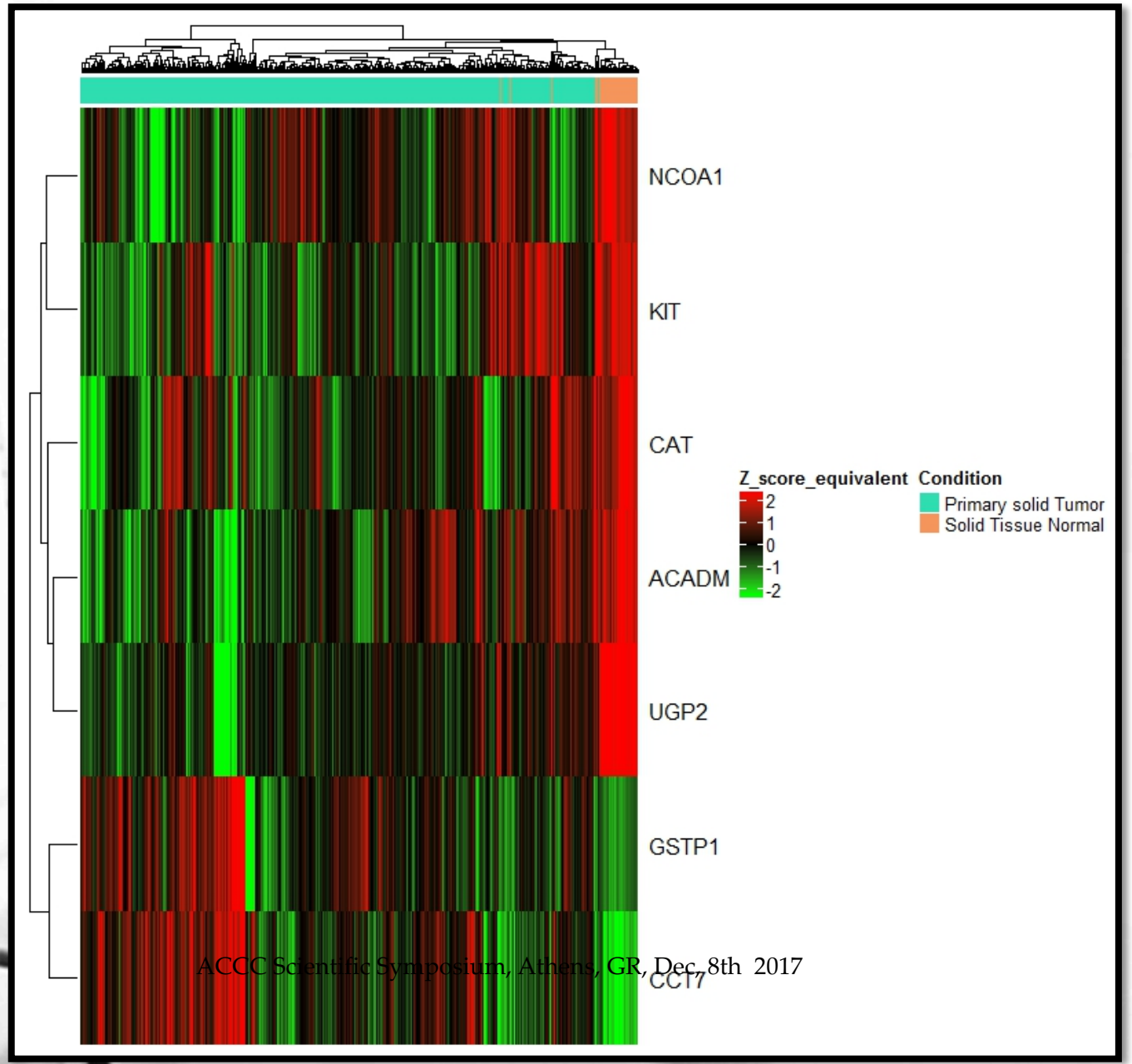
Composite signature of 102 features



Small final compact signature of 12 features

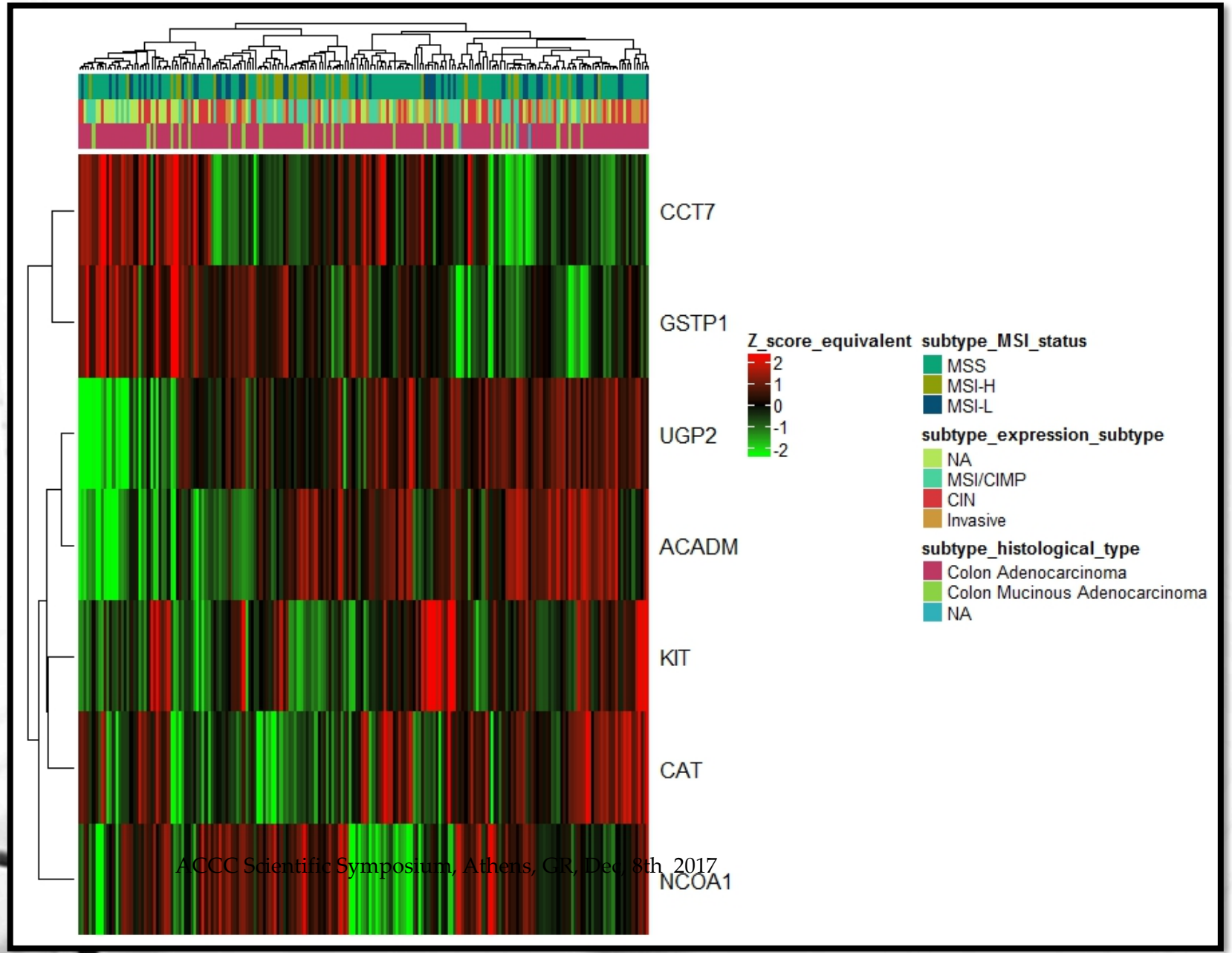


TCGA COAD RNA-Seq dataset-7 genes from lasso selection



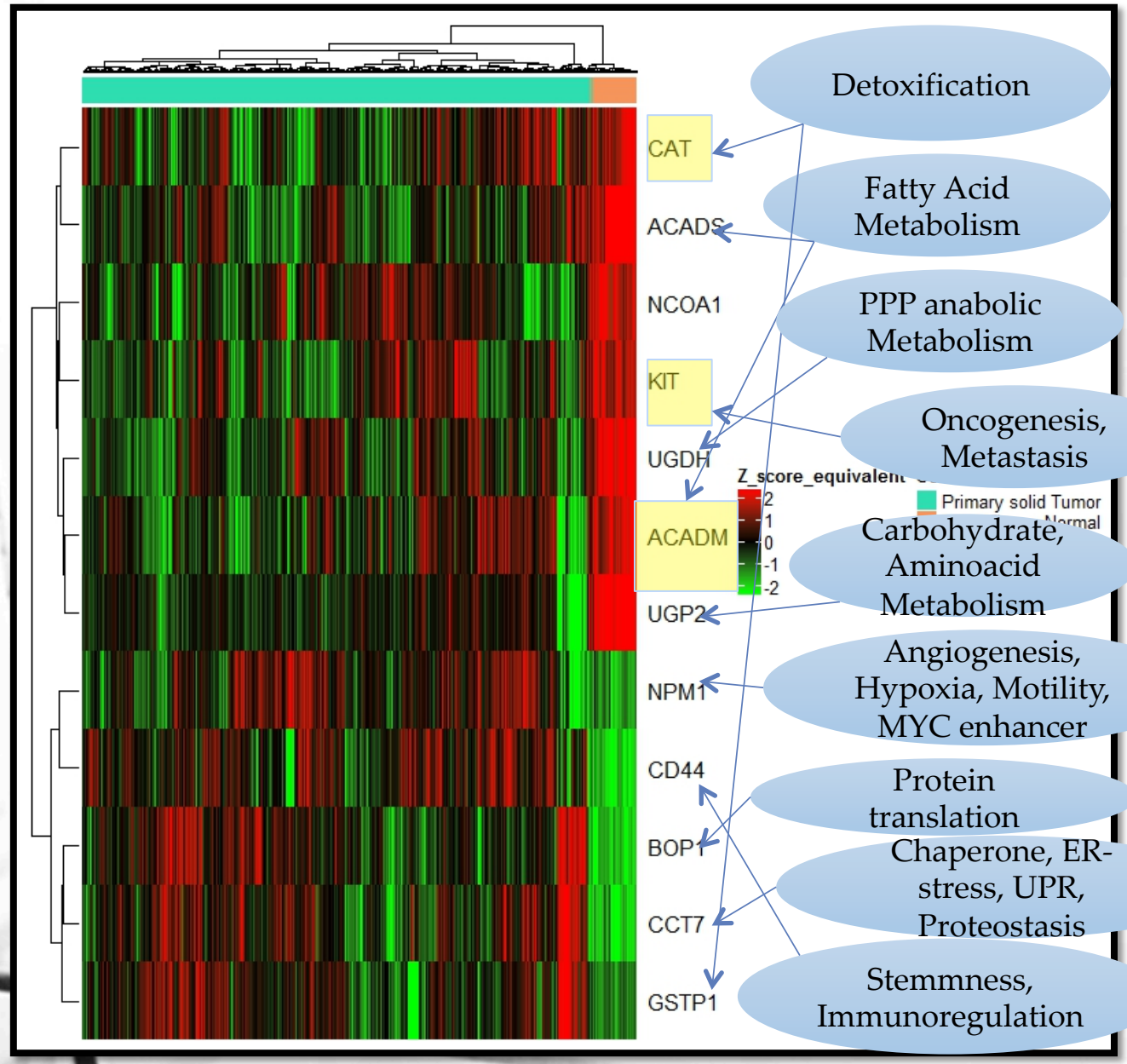
Hierarchical clustering with 7 genes from the 12 selected microarray lasso features (distance: pearson, linkage: average)-scaled HTSeq-counts (vst-transformed)-506 samples: 465 primary solid Tumors and 41 Solid Tissue Normal samples

TCGA COAD RNA-Seq dataset-7 genes-COAD subtypes



Hierarchical clustering with 7 genes from the 12 selected microarray lasso features (distance: pearson, linkage: average)-scaled HTSeq-counts (vst-transformed)-198 samples: only cancer samples with available subtype information

Epidemiological Evaluation of the compact signature in high-quality CRC RNA-Seq samples of TCGA Repository-



Hierarchical clustering with 12 genes (distance: euclidean, linkage: ward)-scaled HTSeq-counts (vst-transformed)-506 samples: 465 primary solid Tumors and 41 Solid Tissue Normal samples

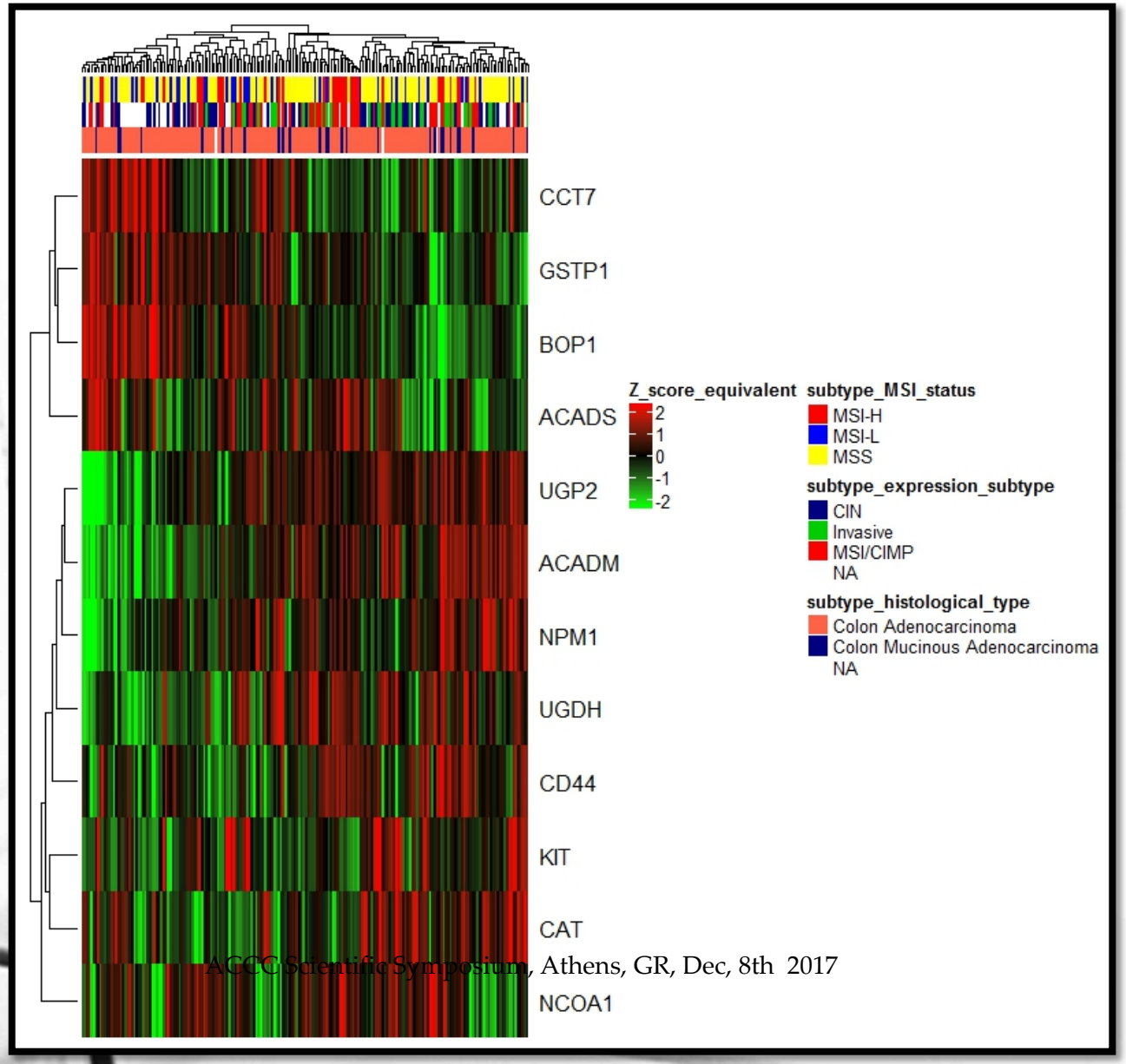
Genes totally uncorrelated with PET variables

CAT:key antioxidant heme enzyme, active against toxic hydrogen peroxide

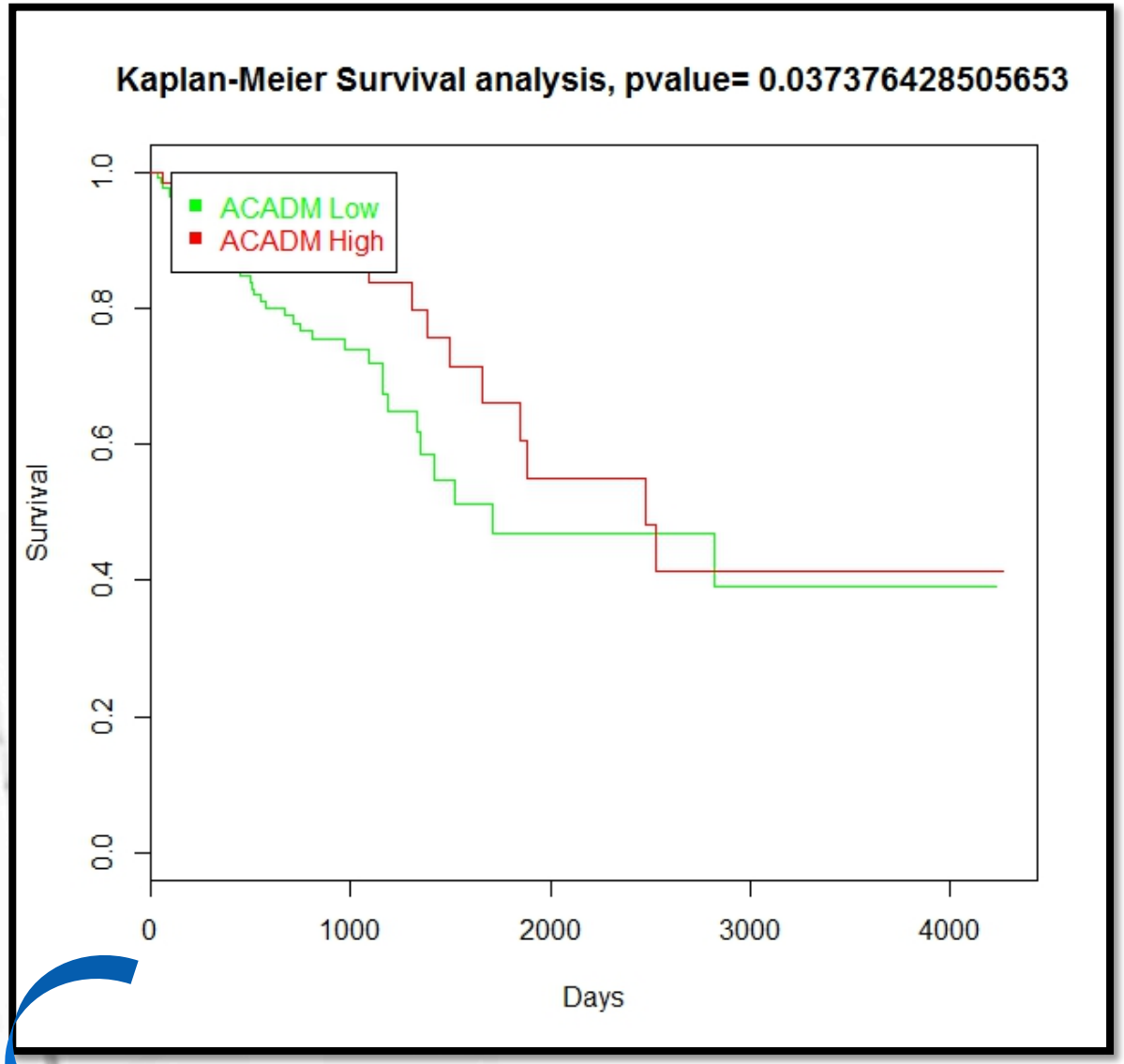
KIT:c-KIT proto-oncogene,tyrosine kinase cell-surface receptor, active in gastrointestinal tumors, metastatic marker

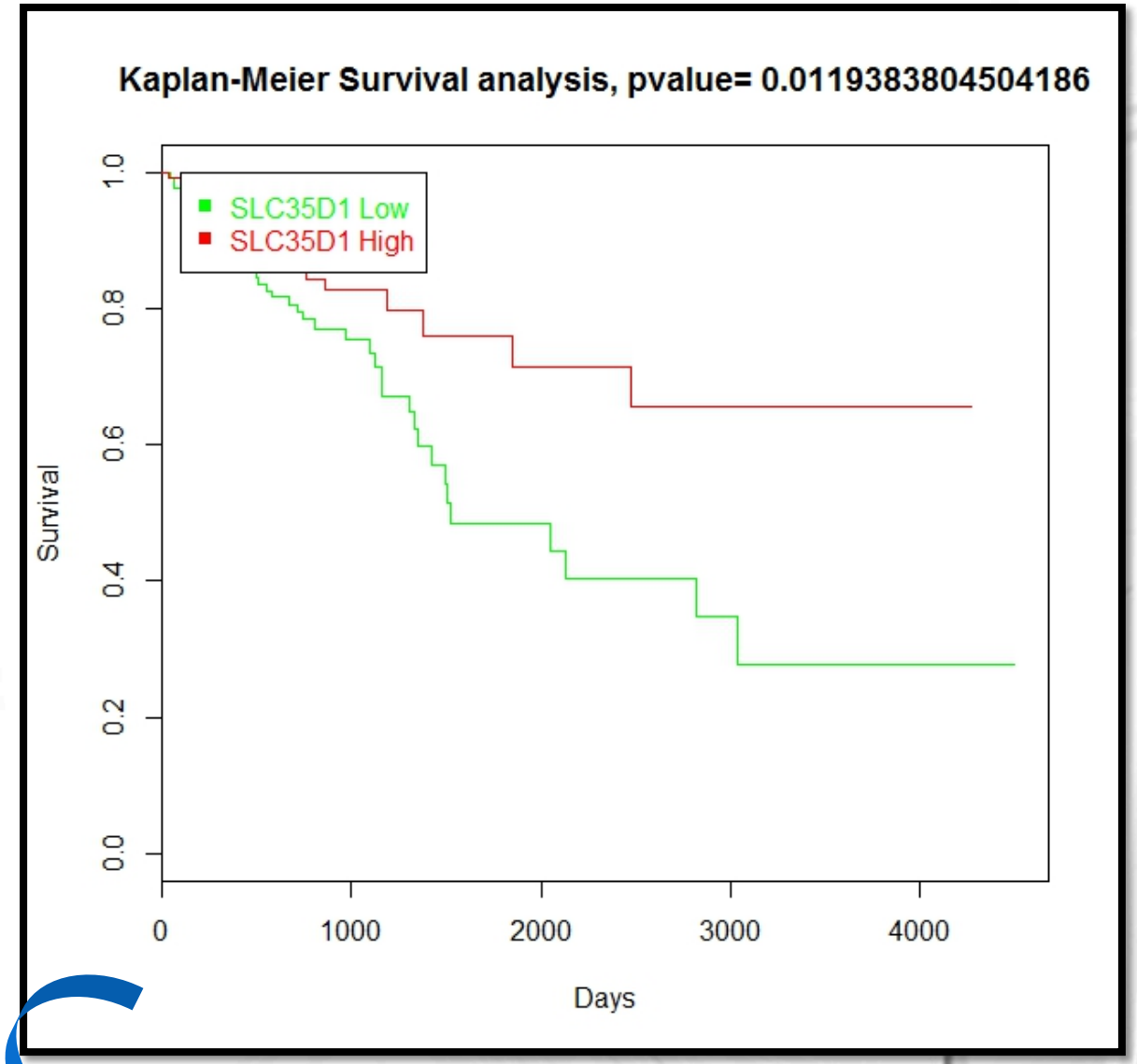
ACADM:acyl-CoA dehydrogenase medium chain, initial step of fatty acid beta-oxidation

TCGA COAD RNA-Seq dataset-12 genes-COAD subtypes



Hierarchical clustering with 12 genes (distance: pearson, linkage: average)-scaled HTSeq-counts (vst-transformed)-198 samples: only cancer samples with available subtype information





Included also in the 12 lasso selected features-main orchestrator of the fatty-acid metabolism pathway

Negative correlation with SUV & FD clinical variables-participates in UDP-glucuronate metabolic process

Summary

- A powerful generic methodology for integration, interpretation and stratification of complex clinical data in a real CRC dataset
- Derivation of an initial highly informative set of 94 differentiated genes to be associated with the respective 8 kinetic PET variables fpr each patient.
- Inference of a compact CRC composite signature associated with distinct modes of tumor physiology (carbohydrate/aminoacid/ PPP anabolic/fatty acid metabolism, stemness, immunomodulation, ER,telomere biology etc)
- This CRC signature was further validated by an independent collection of high quality RNA-Seq samples from The Cancer Genome Atlas (COAD-521 samples), detailing those molecular pathways, represented efficiently by the given PET kinetic parameters
- More importantly provides concrete target genes, with important discriminatory potential that could be targeted by novel contrast agents

Acknowledgements

NHRF MEBP

Olga Papadodima (opapadod@eie.gr) Funct Res

Marianthi Logotheti (mlogotheti@eie.gr)

Post-grad

Irene Liampa(eliampa@eie.gr) MSc

Georgia Kontogianni (gkontogianni@eie.gr) MSc

E-NIOS Applications PC

Efstathios- Jason Vlachavas MSc

Eleftherios Pilalis (epilalis@e-nios.com) PhD

Thodoris Koutsandreas MSc

Ilona Binenbaum, MSc

Surgical Clinic A, Klinikum Ludwigshafen

Stefan Willis
Sven Klippel

Center of Medical Research, University of Rostock,

Dirk Koczan

DKFZ, Heidelberg

Caixia Cheng
Leyun Pan
Christos Sachpekidis
Antonia Dimitrakopoulou- Strauss

NHRF

Alexandros Pintzas
Vassilios Greogoriou