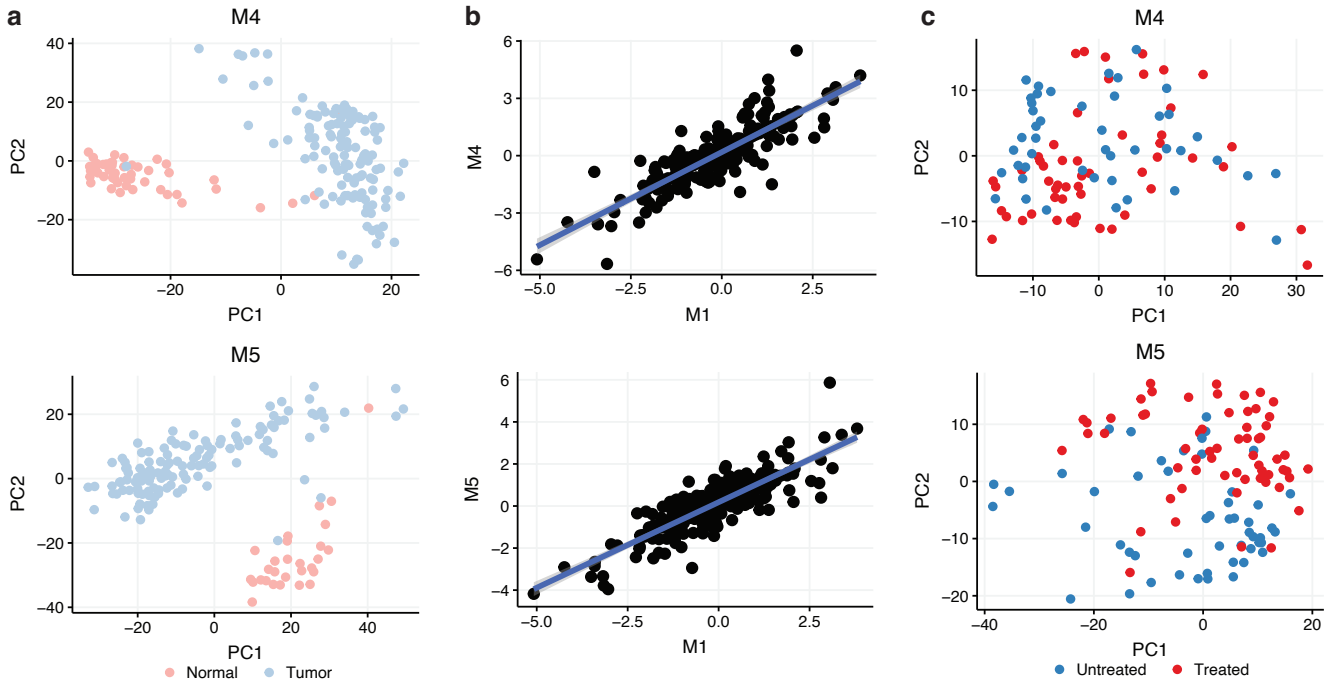

Figure S1. Baseline Metabolic Comparisons, related to Figure 1. a) PCA analysis of all samples colored by tumor and normal. b) Scatterplots comparing log2 fold-change between tumor and normal samples in our metabolomics cohort vs a previously published cohort. c) PCA analysis of all samples colored by treatment exposure.

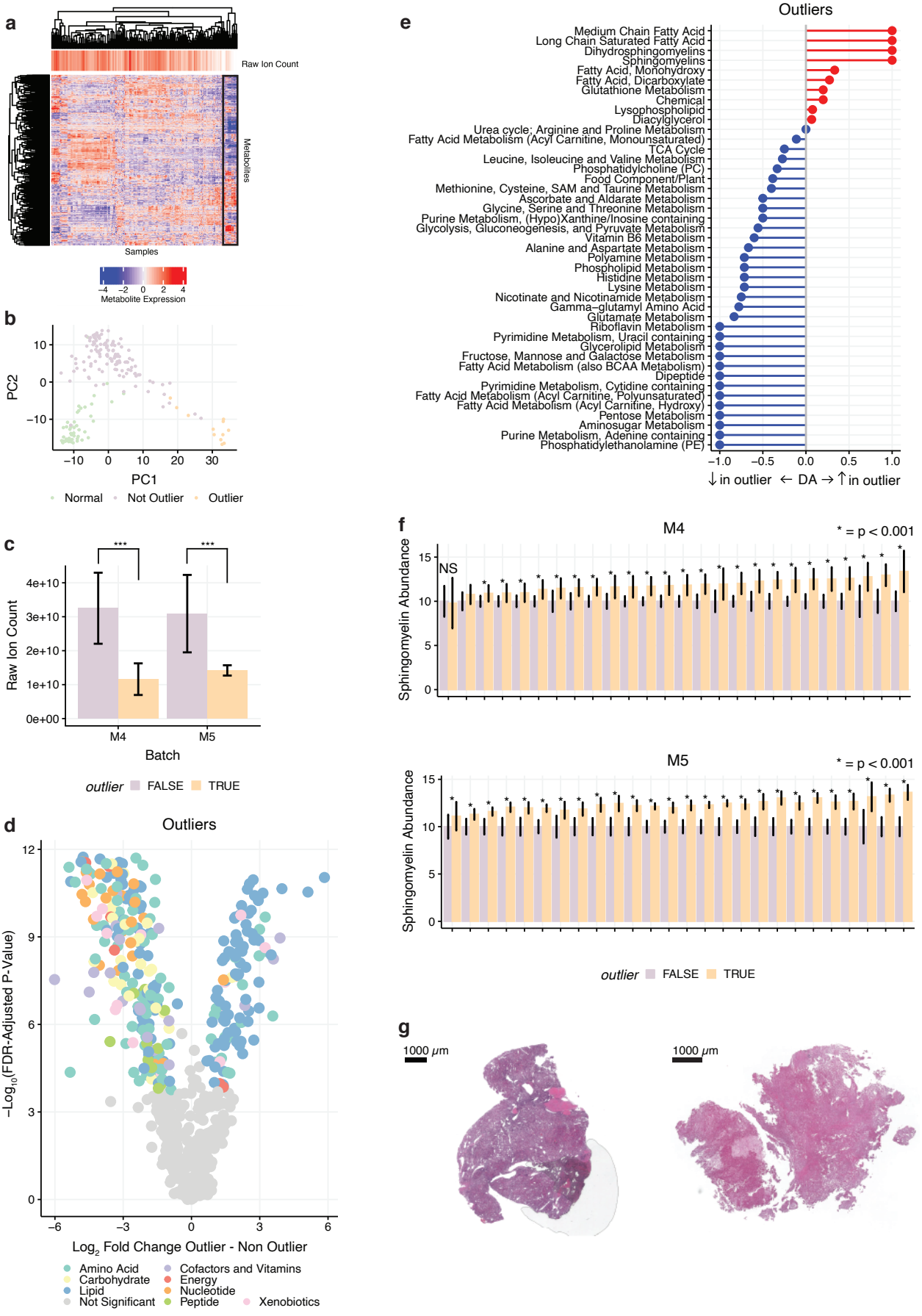
a
Raw Ion Count
Metabolites
Samples
−4 −2 0 2 4
Metabolite Expression
b
PC2
PC1
Normal
Not Outlier
Outlier
c
Raw Ion Count
Batch
M4
M5
outlier FALSE TRUE
d
Outliers
−Log10(FDR-Adjusted P-Value)
Log2 Fold Change Outlier - Non Outlier
Amino Acid
Carbohydrate
Lipid
Not Significant
Cofactors and Vitamins
Energy
Nucleotide
Peptide
Xenobiotics
e
Outliers
Medium Chain Fatty Acid
Long Chain Saturated Fatty Acid
Dihydrosphingomyelins
Sphingomyelins
Fatty Acid, Monohydroxy
Fatty Acid, Dicarboxylate
Glutathione Metabolism
Chemical
Lysophospholipid
Diacylglycerol
Urea cycle; Arginine and Proline Metabolism
Fatty Acid Metabolism (Acyl Carnitine, Monounsaturated)
TCA Cycle
Leucine, Isoleucine and Valine Metabolism
Phosphatidylcholine (PC)
Food Component/Plant
Methionine, Cysteine, SAM and Taurine Metabolism
Ascorbate and Aldarate Metabolism
Glycine, Serine and Threonine Metabolism
Purine Metabolism, (Hypo)Xanthine/Inosine containing
Glycolysis, Gluconeogenesis, and Pyruvate Metabolism
Vitamin B6 Metabolism
Alanine and Aspartate Metabolism
Polyamine Metabolism
Phospholipid Metabolism
Histidine Metabolism
Lysine Metabolism
Nicotinate and Nicotinamide Metabolism
Gamma−glutamyl Amino Acid
Glutamate Metabolism
Riboflavin Metabolism
Pyrimidine Metabolism, Uracil containing
Glycerolipid Metabolism
Fructose, Mannose and Galactose Metabolism
Fatty Acid Metabolism (also BCAA Metabolism)
Dipeptide
Pyrimidine Metabolism, Cytidine containing
Fatty Acid Metabolism (Acyl Carnitine, Polyunsaturated)
Fatty Acid Metabolism (Acyl Carnitine, Hydroxy)
Pentose Metabolism
Aminosugar Metabolism
Purine Metabolism, Adenine containing
Phosphatidylethanolamine (PE)
↓in outlier ← DA → ↑in outlier
f
M4
* = p < 0.001
NS
Sphingomyelin Abundance
M5
* = p < 0.001
Sphingomyelin Abundance
outlier FALSE TRUE
g
1000 μm
1000 μm

Figure S2. Metabolic Features of Necrotic Outliers, related to Figure 1. a) Heatmap of scaled metabolite expression across all samples annotated by raw ion count. Outlier samples are in yellow, normal samples are in green, and non-outlier tumor samples are in purple. b) PCA analysis of samples colored by normal tissue, outliers, and tumor non-outliers). c) Barplot of raw ion counts in outliers compared to non-outliers across both metabolic batches (*** = FDR-adjusted p-value < 0.001. d) Volcano plot of metabolites differentially expressed in outliers vs. non-outlier regions colored by pathway and significance. Metabolites in grey are not significant (FDR-adjusted p-value < 0.05 and/or absolute log2 fold-change > 0.5). e) Pathway-based analysis of metabolic changes in outliers vs. non-outlier regions. f) Barplot of sphingomyelin abundance from metabolomics data in outliers compared to non-outliers across both metabolic batches. (*** = FDR-adjusted p-value < 0.001, NS = not significant). g) Representative images of H&E-stained histological sections. Left: Section with no necrosis (MR03_RC). Bar represents 1000 μm. Right: Section with 100% necrosis and no viable cells (MR05_RB). Bar represents 1000 μm.

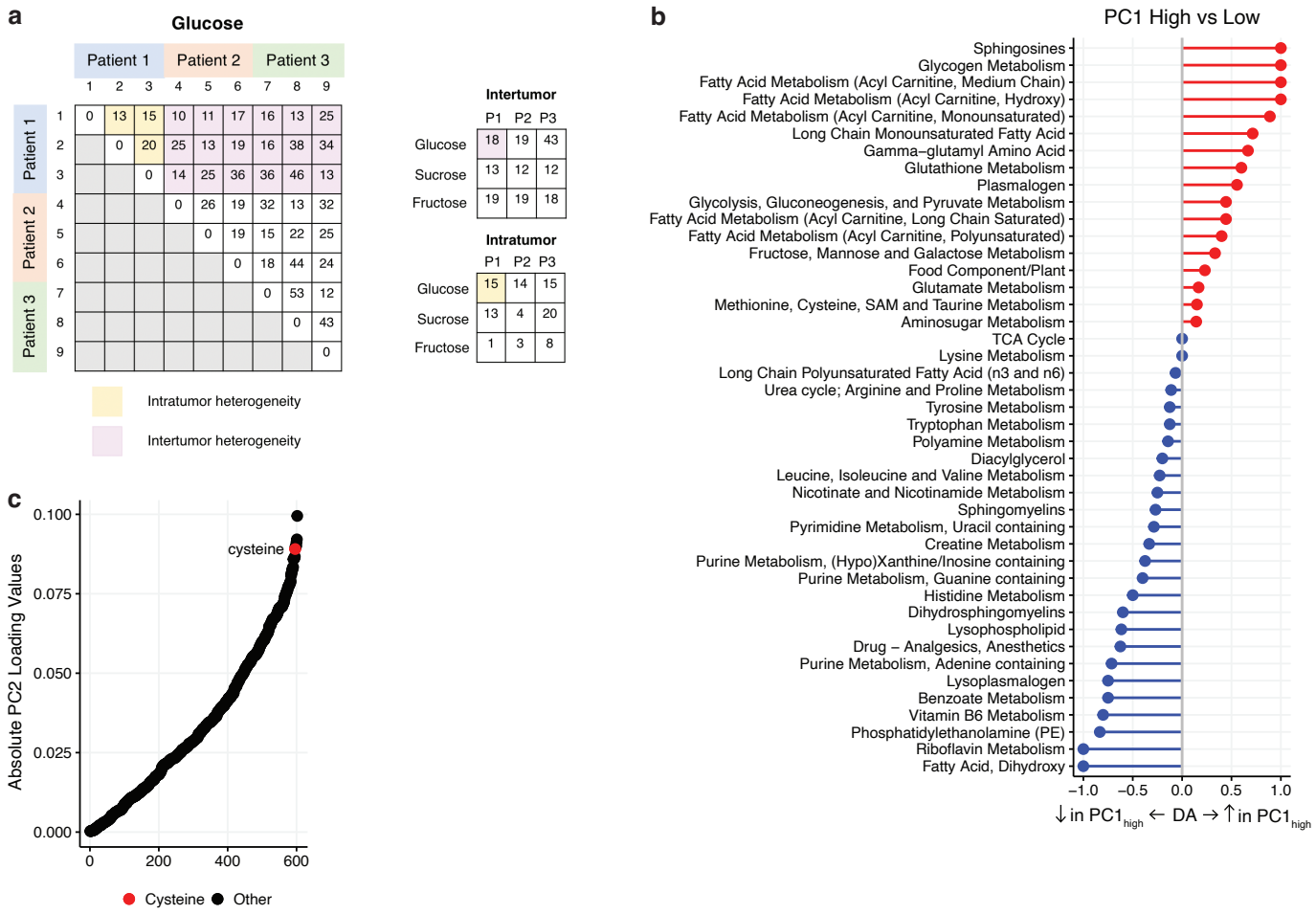

Figure S3. Schematic of mITH score calculations and PCA Features, related to Figures 3 and 4.
a) For each metabolite, pairwise differences are calculated. The median difference between the paired-differences is then taken, yielding a metabolite-specific, patient-specific measure of heterogeneity. Metabolite intratumor heterogeneity values are calculated as the median pairwise difference between samples from the same patient. Intertumoral heterogeneity values are calculated as the median pairwise difference between samples from all other patients. b) Pathway-based analysis of metabolic changes in PC1high vs PC1low regions. c) Absolute value of metabolites in PC2. Cysteine has the seventh highest loading and is shown in red.

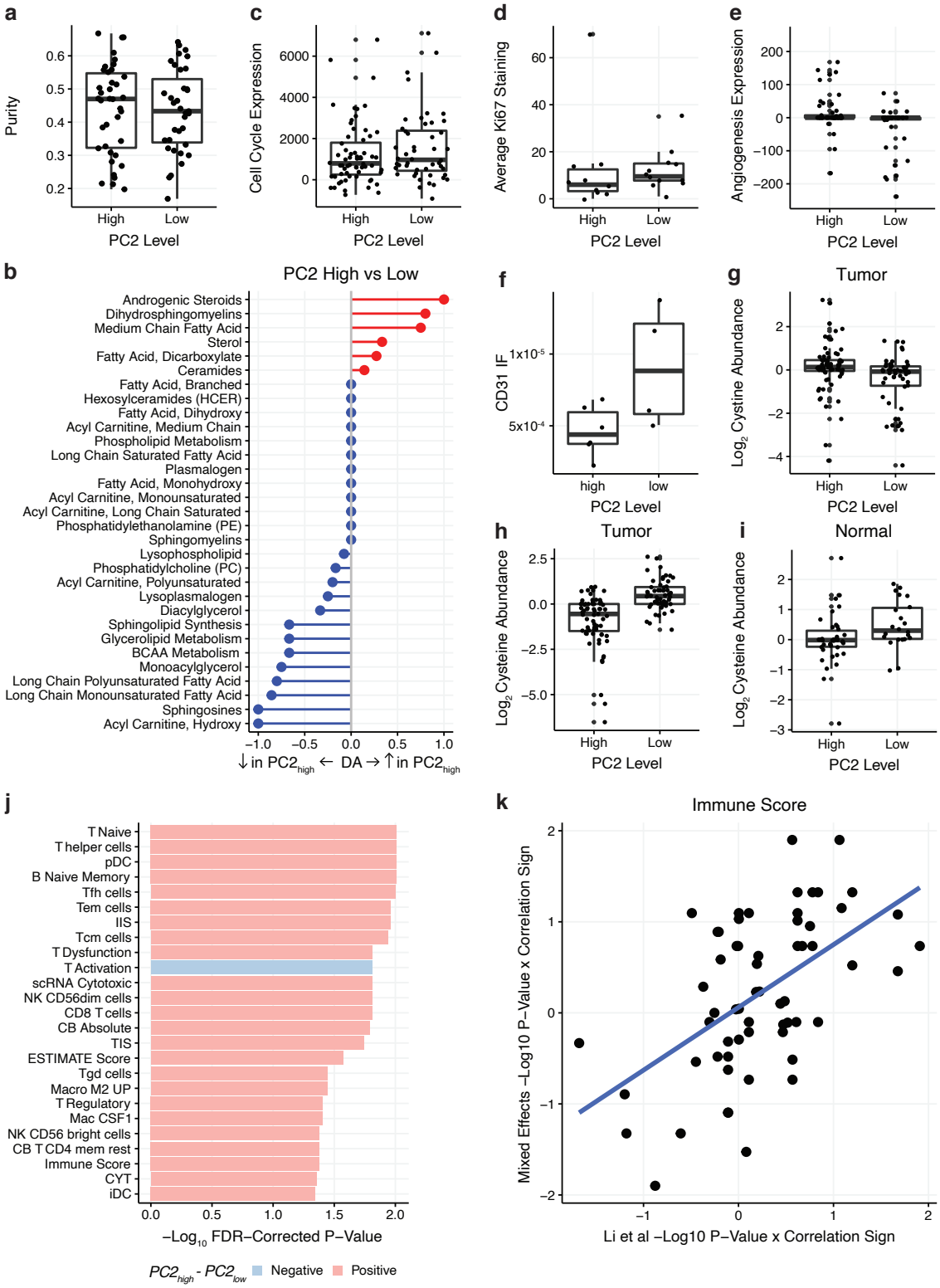

Figure S4. Comparisons Between $PC2_{high}$ vs $PC2_{low}$ regions, related to Figure 3. a) Boxplot comparing tumor purity. b) Differential abundance scores of lipid pathways. c) Boxplot comparing expression of an RNA signature for cell cycle. d) Boxplot comparing average Ki67 staining between regions. e) Boxplot comparing expression of an angiogenesis signature. f) Boxplot comparing CD31 immunofluroescence staining. g) Boxplot comparing cystine abundance in tumor samples. h) Boxplot comparing cysteine abundance in tumor samples. i) Boxplot comparing cysteine abundance in normal samples. j) Barplot of -$\log_{10}$ FDR-adjusted p-values of the 25 immune signatures that are statistically significantly differentially expressed between $PC2_{high}$ and $PC2_{low}$ regions (FDR-adjusted p-value < 0.05). Signatures that are enriched in $PC2_{high}$ regions are in red while signatures that are depleted are in blue. k) Scatterplot comparing the -$\log_{10}$ p-value multiplied by the sign of the Spearman correlation coefficient in our mixed effects multi-regional model vs. a recently published study of ccRCC tumors from Li et al[29].

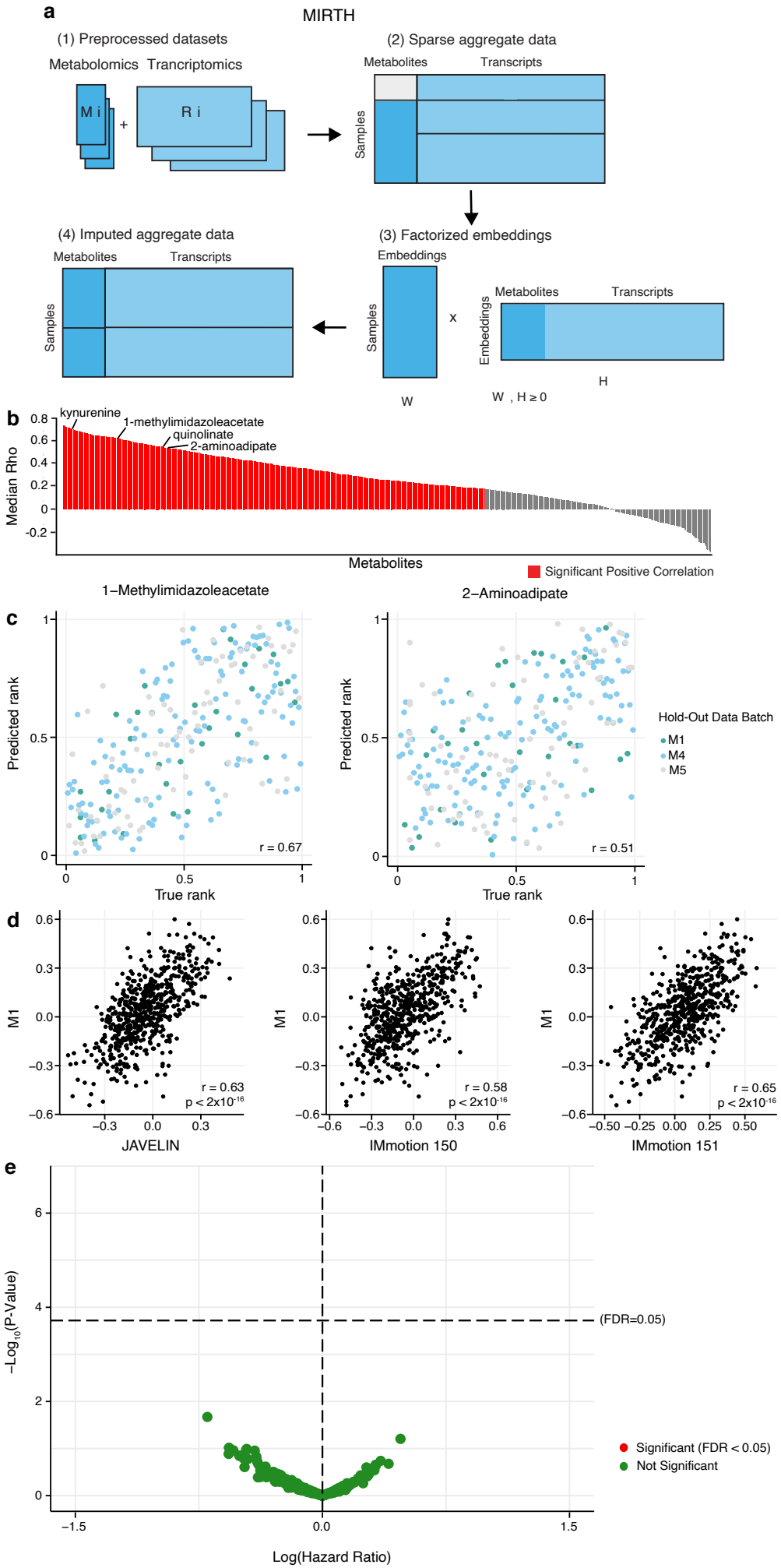
a
MIRTH
(1) Preprocessed datasets
Metabolomics Trancriptomics
M i
R i
(2) Sparse aggregate data
Metabolites
Transcripts
Samples
(3) Factorized embeddings
Embeddings
Samples
W
x
Metabolites
Transcripts
Embeddings
H
W , H ≥ 0
(4) Imputed aggregate data
Metabolites
Transcripts
Samples
b
kynurenine
1-methylimidazoleacetate
quinolinate
2-aminoadipate
Median Rho
Metabolites
Significant Positive Correlation
c
1−Methylimidazoleacetate
2−Aminoadipate
Predicted rank
True rank
r = 0.67
r = 0.51
Hold-Out Data Batch
M1
M4
M5
d
M1
JAVELIN
IMmotion 150
IMmotion 151
r = 0.63
p < 2x10-16
r = 0.58
p < 2x10-16
r = 0.65
p < 2x10-16
e
−Log10(P-Value)
(FDR=0.05)
Log(Hazard Ratio)
Significant (FDR < 0.05)
Not Significant

Figure S5. MIRTH's performance of imputing metabolomics from transcriptomics, related to FIgure 6.
a) Schematic workflow of MIRTH imputation of metabolomics data from paired transcriptomics data. MIRTH solves a non-negative matrix factorization problem using two forms of data: a reference dataset or datasets with matched RNA and metabolite data, and a target dataset with only RNA data. b) Barplot of median Spearman's correlation values for each simulated-missing metabolite in RC18 across 10 MIRTH iterations. Metabolites whose predicted ranks are significantly correlated with ground-truth ranks in >90% of iterations are labeled red. Metabolites with significant negative correlations are discarded. c) Scatterplot comparing true ranks of the metabolite to predicted ranks by MIRTH in each dataset for 1-methylimidazole acetate (left) and 2-aminoadipate (right) d) Correlations between metabolite levels and Immune Score are recapitulated in ground truth data and imputed clinical trial data. e) Volcano plot of 262 reproducibly well-predicted metabolites associated with PFS in immunotherapy arms of ccRCC clinical trials, colored by significance (Cox's proportional-hazards test, adjusted by age and sex). Metabolites in green are not significant (FDR-adjusted p-value > 0.05). Y-axis is the unadjusted p-values of the CPH test.

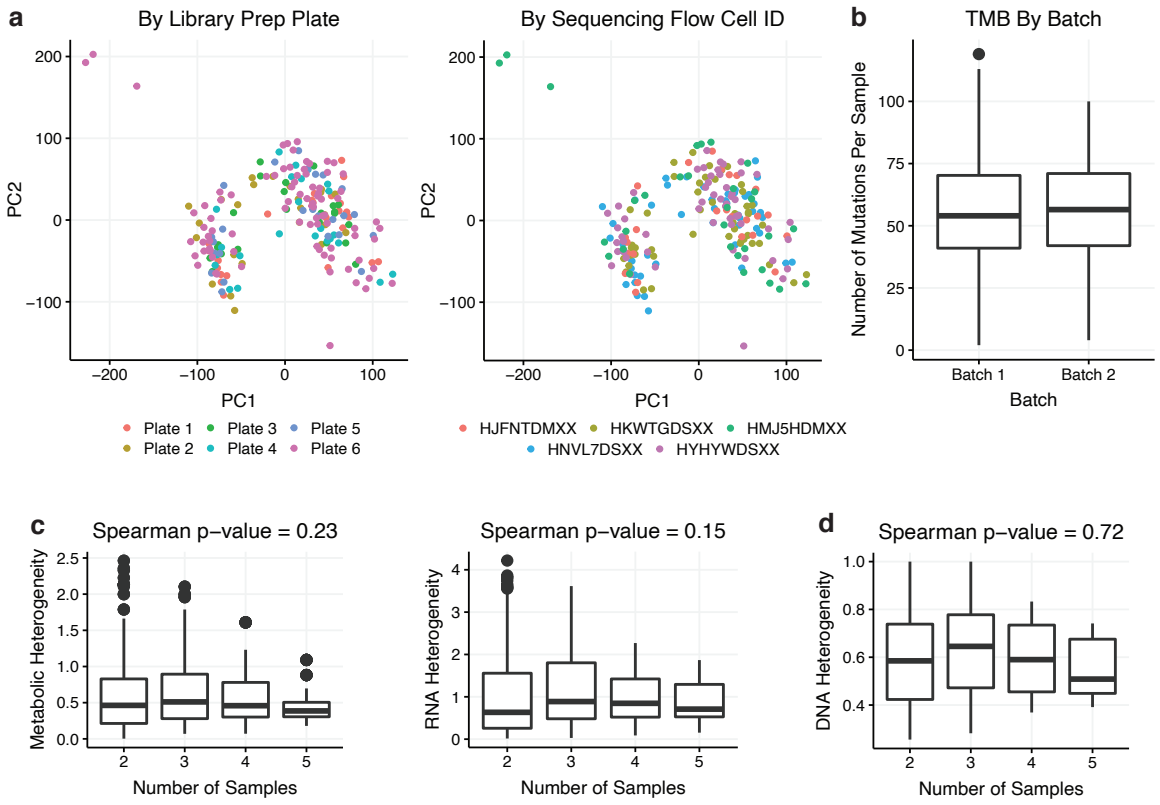

Figure S6. Confirmation of Lack of Bias in Sampling and Heterogeneity, related to STAR Methods.
a) Left: PCA of RNA sequencing samples colored by library prep plate. Right: PCA of RNA sequencing samples colored by flow cell ID. b) Boxplots comparing TMB across whole exome sequencing batches. c) Left: Boxplot showing metabolic intratumor heterogeneity levels based on the number of regions sampled. Downsampling was performed to randomly select 2-5 regions to calculate heterogeneity. Right: Boxplot showing RNA heterogeneity levels based on the number of regions sampled. d) Boxplot showing genetic heterogeneity levels based on the number of regions sampled.