Prognostic signatures of oxygen-sensing genes predict patient survival in ten cancers including those of the liver, pancreas and stomach

Wai Hoong Chang, Donall Forde and Alvina G. Lai

Nuffield Department of Medicine, University of Oxford, Old Road Campus, OX3 7FZ, United Kingdom

For correspondence: Alvina.Lai@ndm.ox.ac.uk; alvinagracelai@gmail.com

Abstract

Objectives: Tumor hypoxia is associated with metastasis and resistance to chemotherapy and radiotherapy. Genes involved in oxygen-sensing are clinically relevant and have significant implications on prognosis.

Methods: We identified of two signatures, signature 1 (good prognosis) and signature 2 (adverse prognosis), each consisting of 5 genes using three pancreatic cancer cohorts (n=681). We validated the signatures' performance in predicting survival in ten cancers using Cox regression and receiver operating characteristic (ROC) analyses.

Results: Signature 1 and signature 2 were associated with good and poor overall survival respectively. Prognosis of signature 1 in 8 cohorts representing 6 cancers (n=2,627): bladder (hazard ratio [HR]=0.68, P=0.039), papillary renal cell (HR=0.35, P=0.013), liver (HR=0.64, P=0.033 and HR=0.49, P=0.025), lung (HR=0.66, P=0.014) and pancreatic (HR=0.42, P<0.001 and HR=0.64, P=0.04) and endometrial (HR=0.40, P<0.001). Prognosis of signature 2 in 12 cohorts representing 9 cancers (n=4,134): bladder (HR=1.46, P=0.039), cervical (HR=1.97, P=0.035), head and neck (HR=1.39, P=0.038), renal clear cell (HR=1.47, P=0.012), papillary renal cell (HR=3.89, P=0.0015), liver (HR=5.10, P<0.0001 and HR=2.26, P<0.001), lung (HR=1.54, P=0.011), pancreatic (HR=2.09, P=0.002, HR=1.46, P=0.018, and HR=1.99, P<0.0001) and stomach (HR=1.78, P=0.004). Multivariate Cox regression confirmed independent clinical relevance of signatures in these cancers. ROC analyses confirmed superior performance of signatures to current tumor staging benchmarks. *KDM8* is a potential tumor suppressor downregulated in liver and pancreatic cancers and is an independent prognostic factor. *KDM8* expression negatively correlated with cell cycle regulators. Low *KDM8* in tumors was associated with loss of cell adhesion phenotype through *HNF4A* signaling.

Conclusions: Pan-cancer signatures of oxygen-sensing genes used for risk assessment in 10 cancers (n=6,761) could guide individualized treatment plans.

Introduction

Solid tumors face considerable demands in maintaining oxygen availability due to their unique vasculature systems[1]. Rapid neoplastic cell proliferation and overexpression of angiogenic factors leading to the formation of disorganized blood vessels result in insufficient oxygen supply to tumor cells[2,3]. Hence, there is a requirement for tumors to evolve systems that detect changes in oxygen homeostasis. The discovery of hypoxia-inducible factor (HIF) as one of the key oxygen-sensing gene represents a quantum leap forward in tumor biology[4,5]. HIF has co-evolved with the need for more sophisticated oxygen delivery system but intriguingly, it is found in primitive animals that lack oxygen-delivery structures altogether[6]. Its discovery has led to the development of drugs used to treat anemia and cancer[7,8].

In addition to HIF, 2-oxoglutarate (2OG) oxygenases represent another family of oxygen-sensing proteins. As suggested by its name, this group of enzymes has an absolute requirement for molecular oxygen. They catalyze a range of oxidative modifications and their activities are affected by nutrient and oxygen availability[9], both of which are altered within the tumor microenvironment. Several members from this gene family have been implicated in cancer. *Ten-Eleven Translocation 2* (*TET2*) is frequently found to be mutated in leukemia[10] and other solid malignancies[11]. JmjC lysine demethylases (*KDMs*) epigenetically regulate cancer-related genes[12] and inactivating mutations are observed in multiple cancers; glioblastoma/neural, breast, colorectal and renal[13,14].

We hypothesize that 2OG oxygenases expression could help predict prognosis in solid malignancies that are characteristically oxygen-deprived. Additionally, we hypothesize this would be applicable to different types of cancers as they share a uniform need to overcome hypoxia for survival. Starting from an initial set of 61 2OG oxygenases, we developed two prognostic gene signatures, each consisting of a minimal 5 genes, that could facilitate risk stratification and predict overall survival in cancer patients. We confirmed their prognostic performance through a multi-cohort pan-cancer validation process.

Methods

Detailed methods are available in supplementary information.

Results

Multi-cohort analyses revealed two prognostic 2OG oxygenases signatures

We analyzed the prognostic significance of 61 2OG oxygenase genes in 25 cancers (n = 19,781) from The Cancer Genome Atlas[15](Table S1 & S2). Prognostic genes were defined as those whose expression levels were significantly correlated with patients' overall survival. Since the highest number of prognostic genes (29/61) was observed in pancreatic cancer (n=178) and it is one of the most difficult to treat cancers, two additional pancreatic cancer cohorts from the International Cancer Genome Consortium [ICGC][16] (n=269 and n=234) were used in combination as training cohorts (Fig.1A). We defined two gene signatures; signature 1 and signature 2 as favorable and unfavorable prognostic factors respectively, by taking into consideration genes that were significant in univariate Cox analyses in two out of three pancreatic cancer cohorts (Fig.1A,1B). Signature 1 constitutes 5 genes: *KDM8, KDM6B, P4HTM, ALKBH4* and *ALKBH7*. Likewise, 5 genes make up signature 2: *KDM3A, P4HA1, ASPH, PLOD1* and *PLOD2* (Fig.1A).

Patients were median-dichotomized based on mean expression scores of signatures 1 and 2 genes. Patients with high expression of signature 1 genes had significantly better survival in 6 cancers: bladder (hazard ratio [HR]=0.68, P=0.039), papillary renal cell (HR=0.35, P=0.013), liver (HR=0.64, P=0.033 and HR=0.49, P=0.025), lung (HR=0.66, P=0.014), pancreatic (HR=0.42, P<0.001 and HR=0.64, P=0.04) and endometrial (HR=0.40, P<0.001); consistent with the fact that signature 1 is a marker of good prognosis (Fig.2A; Table S3). In contrast, patients with high expression of signature 2 genes had significantly worse prognosis in 9 cancers: bladder (HR=1.46, P=0.039), cervical (HR=1.97, P=0.035), head and neck (HR=1.39, P=0.038), renal clear cell (HR=1.47, P=0.012), papillary renal cell (HR=3.89, P=0.0015), liver (HR=5.10, P<0.0001 and HR=2.26, P<0.001), lung (HR=1.54, P=0.011), pancreatic (HR=2.09, P=0.002, HR=1.46, P=0.018, and HR=1.99, P<0.0001) and stomach (HR=1.78, P=0.004) (Fig.2B).

Cross-platform subgroup and multivariate analyses confirmed the validity of signatures 1 and 2 as independent prognostic factors.

To assess the independence of signatures 1 and 2 over current tumor staging systems, we performed subgroup analyses of their prognostic effect in patients with early (stages 1 and/or 2), intermediate (stages 2 and/or 3) and late (stages 3 and/or 4) cancer stages. Kaplan-Meier analyses revealed that signature 1 successfully identified high-risk (low-expression score) and low-risk (high-expression score) patients with early (bladder, liver, lung, pancreatic and endometrial), intermediate (liver and endometrial) late (papillary renal cell) disease stages (Fig.3A; Fig.S1A). For signature 2, patients with high-expression scores had a significantly poorer survival (Fig.3B). Signature 2 was also independent of disease stage as it successfully predicted survival in early (liver, lung and pancreatic), intermediate (bladder, liver, lung and stomach) and late (bladder, head and neck, papillary renal cell, liver and stomach) stages (Fig.3B; Fig.S1B).

Receiver operating characteristic (ROC) analyses were employed to determine the predictive performance (sensitivity and specificity) of signatures 1 and 2 on 5-year overall survival. Signature 1 performed best, as measured by area under the curve (AUC), in pancreatic cancer (AUC=0.754) followed by papillary renal cell (AUC=0.738), liver (AUC=0.652 and AUC=0.613), bladder (AUC=0.645), endometrial (AUC=0.635) and lung (AUC=0.625) cancers (Fig.3C; Fig.S1C). Signature 2 performance in papillary renal cell cancer (AUC=0.810) was the best, followed by cervical (AUC=0.692), liver (AUC=0.675 and AUC=0.625), pancreatic (AUC=0.668), renal clear cell (AUC=665), head and neck (AUC=0.632), lung (AUC=0.623), stomach (0.618) and bladder (AUC=0.605) cancers (Fig.3D; Fig.S1D). Performance of both signatures 1 and 2 is superior to current tumor staging benchmarks except for the following: signature 1 (liver, lung and endometrial) and signature 2 (bladder, renal clear cell, liver and lung) (Fig.3C,D; Fig.S1C,D). Remarkably, when used in combination with tumor, node and metastasis (TNM) staging parameters, both signatures consistently outperformed each of the individual classifiers (TNM staging or gene signatures when considered alone), reinforcing their incremental prognostic values (Fig.3C,D; Fig.S1C,D). Significantly, while tumor staging could not predict outcome in cervical cancer patients (AUC<0.5), signature 2 sufficiently served as an adverse prognostic factor (Fig.3D).

Univariate Cox regression analyses revealed that tumor stage was associated with survival except for cervical and pancreatic cancers (Table S3). This is expected given the low AUC values for tumor

stage in both cancers: pancreatic (AUC=0.593) and cervical (AUC=0.455), suggesting that current staging systems for these cancers are inadequate (Fig.3C,D). Multivariate Cox analyses after adjusting for tumor staging showed that signatures 1 and 2 remained significantly associated with survival (Table S3). For 2 liver cancer cohorts we considered additional clinicopathological features. The GSE14520 cohort consisted of Chinese patients with hepatitis B-associated hepatocellular carcinoma[17] while LIRI-JP is a Japanese-based cohort of mixed etiology[18]. Tumor size, cirrhosis, TNM staging, Barcelona Clinic Liver Cancer staging and alpha-fetoprotein levels were all significantly associated with survival in GSE14520 (Table S3). Tumor size could also predict survival in LIRI-JP (Table S3). When these significant covariates along with signatures 1 or 2 were included in multivariate Cox models, the signatures remained significant risk factors: signature 1 (LIRI-JP: HR=0.54, P=0.043) and signature 2 (LIRI-JP: HR=4.539, P=1.8e-4 and GSE14520: HR=2.01, P=0.003) (Table S3). These results highlight the potentially superior prognostic ability of our signatures; signatures 1 and 2 identified high- and low-risk patients in 8 (n=2,627) and 12 (n=4,134) independent cohorts respectively to collectively represent 10 diverse cancers (Fig.1A).

Significance of somatic mutations in risk stratified patients

Patients were risk stratified into low- and high-risk groups using signatures 1 and 2. For signature 1, high-risk patients had significantly lower expression levels of good prognosis genes *ALKBH4, ALKBH7, KDM8, KDM6B* and *P4HTM* (Fig.S2A). In contrast, high-risk patients as stratified by signature 2 had significantly higher expression levels of adverse prognosis genes *ASPH, KDM3A, P4HA1, PLOD1* and *PLOD2* (Fig.S2B). To ascertain the relationship between tumor hypoxia and expression of signature genes, hypoxia scores were computed for each patient as mean expression values ($\log_2$) of 52 hypoxia signature genes[19]. Signature 1 expression scores in patients negatively correlated with hypoxia score (Fig.S3A). Since tumor hypoxia is associated with distant metastasis, recurrence and reduced therapeutic response[20], patients with high expression of signature 1 genes (low hypoxia score) correlated with less advance disease states consistent with it being a marker of good prognosis (Fig.S3A).

Conversely, signature 2 scores positively correlated with tumor hypoxia and hence poor survival outcomes (Fig.S4A). We anticipate that patients' individual risks of death, as determined from signatures 1 and 2, will positively correlate tumor hypoxia. Indeed, risk scores for each patient, as

calculated by taking the sum of Cox regression coefficient for each of the individual genes multiplied with its corresponding expression value[21], significantly correlated with hypoxia scores (Fig.S3B,S4B). Hence, high risk patients have more hypoxic tumors suggesting that our gene signatures are efficient and adequate in predicting death.

To ascertain the association between patients' risks, as determined by our gene signatures, and somatic mutations, we retrieved the top five most commonly mutated genes for each cancer. Mutations in *PCDHA1*, a cell adhesion gene from the cadherin superfamily, were associated with poor survival in bladder (HR=1.65, P=0.027) and stomach cancers (HR=1.53, P=0.046) but improved survival in endometrial cancer (HR=0.52, P=0.042) (Table S3). Mutations in another gene from the protocadherin alpha cluster, *PCDHA2*, were also associated with adverse outcomes in stomach cancer (HR=1.60, P=0.025) (Table S3). Mutations in *TTN* (connectin) and the tumor suppressor *TP53* were associated with poor survival in bladder (HR=1.61, P=0.016) and endometrial cancers (HR=1.78, P=0.041) (Table S3). Interestingly, another tumor suppressor *PTEN*, when mutated, is linked to better outcomes in endometrial cancer (HR=0.43, P=0.006) (Table S3). Similar observations were made for a lipid kinase gene *PIK3CA* in endometrial cancer (HR=0.36, P=0.002) (Table S3). Likewise, mucin *MUC4* mutations improved survival in clear cell renal cancer patients (HR=0.57, P=0.012) (Table S3); an observation that is consistent with another study[22].

Multivariate Cox analyses on signatures 1 and 2 while controlling for significant somatic mutation variables revealed that the gene signatures were independent survival predictors: bladder (signature 1: HR=0.69, P=0.047 and signature 2: HR=1.58, P=0.043), renal clear cell (signature 2: HR=1.52, P=0.007), stomach (signature 2: HR=1.80, P=0.006) and endometrial (signature 1: HR=0.52, P=0.024) (Table S3). Patients stratified by signatures 1 or 2 and mutation status indicated that they collectively were significantly associated with survival (Fig.4). In bladder cancer, high risk patients (low signature 1 score) harboring mutant alleles of *PCDHA1* had ~50% increased mortality at 5 years compared to low risk patients (high signature 1 score) with wild-type *PCDHA1* (P=0.016) (Fig.4). Although results were less dramatic for *PCDHA1* and signature 2, we still observed a ~25% elevated mortality at 5 years for these two patient groups with bladder cancer (P=0.04) (Fig.4). In patients with stomach cancer, high-risk patients (high signature 2 score) with mutant *PCDHA1* had the worst outcomes (P=0.0017) (Fig.4). Conversely, *PCDHA1* mutation was associated with good prognosis in endometrial cancer, hence high-risk patients with wild-type *PCDHA1* had the lowest survival rates while survival is improved by ~20% in low-risk patients with mutant *PCDHA1* (P=0.003) (Fig.4).

PIK3CA (P<0.001*)* and *PTEN* (P=0.0013) mutations were associated with good outcomes in endometrial cancer (Fig.4). Mutations in another cadherin gene *PCDHA2* when considered alongside signature 2 were also associated with survival in stomach cancer (P<0.001) (Fig.4). Survival rates were reduced by ~37% in high-risk patients with mutant *PCDHA2* (Fig.4). Joint relation between *TP53* mutation and signature 1 significantly influenced survival in endometrial cancer (P=0.0015) (Fig.4). Since *MUC4* mutations were associated with better outcomes, survival rates were the lowest in high-risk patients (high signature 2 scores) with wild-type *MUC4* (P=0.0025) (Fig.4).

Tumor suppressive roles of *KDM8* through cell cycle regulation and cell adhesion maintenance

Of all the signature genes, *KDM8* was identified as one of the most downregulated genes in tumors compared to non-tumor samples (Fig.S5A). Patients with high *KDM8* levels had significantly lower risk of death in pancreatic (HR=0.57, P=0.016) and liver cancer cohorts (HR=0.52, P=0.0031; HR=0.49, P=0.026 and HR=0.64, P=0.039) (Fig.5). Prognostic significance of *KDM8* was also independent of tumor stage (Fig.5). *KDM8* expression decreased as tumor malignant grade increased in that the least advanced stage 1 tumors had the highest median *KDM8* values (Fig.S5B). Moreover, *KDM8* expression negatively correlated with hypoxia score, indicating that patients with low levels of *KDM8* had more hypoxic tumors and poorer survival outcomes (Fig.S5C). Together, these observations suggest that *KDM8* may function as a tumor suppressor. This hypothesis is corroborated by an independent report on the role of *KDM8* in cell cycle regulation[23]. Indeed, we observed that *KDM8* expression negatively correlated with the expression levels of canonical cell-cycle genes: cyclins (*CCNA2, CCNB1, CCNB2, CCND1, CCNE1* and *CCNE2*) and cyclin-dependent kinases (*CDK1, CDK2, CDK4, CDK6, CDK7* and *CDK8*), which were consistent across all liver and pancreatic cancer cohorts (Fig.S5D). This implied that *KDM8* is required for tight control of the cell-cycle machinery and its reduction may lead to aberrant proliferation commonly seen in cancer cells.

To ascertain the biological consequences of deregulated *KDM8* expression, we conducted different expression analysis on liver cancer patients categorized into low and high groups based on *KDM8* expression. A total of 745 genes were differentially expressed (DEGs) between the two groups (fold change>2 or <-2, P<0.05) (Table S4). Significant enrichments of biological pathways involved in metabolism, immune regulation, *VEGF* production, inflammation and cell adhesion were observed (Fig.S5E;Table S5). Furthermore, DEGs were overrepresented as targets of *HNF4A, HNF4G, FOXA1,*

FOXA2 and *NR2F2* transcription factors (TFs) (Fig.S5F). These TFs play central roles in cell polarity maintenance and epithelial differentiation[24–26], hence downregulation of *KDM8* may drive epithelial-mesenchymal transition (EMT) and tumor progression. *HNF4A* is one of the key TFs responsible for regulating a myriad of hepatic functions including cell junction assembly[26,27]. Pathway analysis of *KDM8* DEGs revealed enrichment of processes related to cell adhesion, suggesting a potential crosstalk between *KDM8* and *HNF4A*. Of the 745 DEGs, analysis on a hepatoma-based HNF4A chromatin immunoprecipitation-sequencing dataset demonstrated that 148 genes were directly bound by HNF4A[28]. To further reinforce the interplay between *HNF4A* and *KDM8*, we observed that 110 of the 745 DEGs were overrepresented in HNF4A-null mice[26] and 45 of these genes were direct HNF4A targets (Fig.S5G). Differential expression analysis between *HNF4A*-deficient and wild-type mice showed that a majority of the 45 genes were downregulated, as expected, suggesting that HNF4A directly activates their gene expression, many of which are involved in a multitude of cell adhesion processes (Fig.S5H).

Discussion

A multi-cohort retrospective study led to the identification of two novel pan-cancer prognostic gene signatures derived from oxygen-sensing genes. Cross-platform validations confirmed prognosis in 10 cancers to collectively include 6,761 patients spanning 20 diverse cohorts (Fig.1A). The gene signatures had opposing prognostic value; signature 1 is a marker of good prognosis while signature 2 is associated with poor outcomes. The key strengths of our signatures as powerful prognostic tools are: 1) pan-cancer utility, 2) involving a mere 5 genes each that provide continuous assessment of death risks and 3) superior to current tumor staging parameters. The prognostic capability of our signatures in diverse cancers is evidence that conserved biological processes; oxygen-sensing, cell-cycle deregulation and loss of cell polarity, governs disease progression and thus clinical outcomes.

Anti-tumorigenic functions have been reported for several genes from signature 1. Loss of *KDM6B* resulted in more aggressive pancreatic ductal adenocarcinoma[29]. In colorectal cancer, high *KDM6B* is a predictor of good prognosis and knock-down of *KDM6B* augmented cell proliferation and inhibited apoptosis[30]. Yet, *KDM6B* function is enigmatic. Others have reported that high *KDM6B* expression is associated with increased metastasis and invasion in renal clear cell carcinoma that correlated with poor prognosis[31]. *KDM6B* also promotes TGF-B induced EMT and invasiveness in breast cancer[32]. While we could neither confirm nor deny the validity of these studies, it is striking that our observation of favorable prognosis associated with high *KDM6B* in pancreatic ductal adenocarcinoma was consistent with the report from Yamamoto and colleagues[29]. We also did not observe any prognostic significance of signature 1, in which *KDM6B* is part of, in either breast or renal clear cell cancers, which indirectly substantiates findings from the two other reports on *KDM6B* not being a marker of good prognosis[31,32].

KDM8 is a gene associated with favorable prognosis. Our results suggest determinative crosstalk between *KDM8* and *HNF4A*, particularly in the context of morphogenesis, cell adhesion, maintenance of cell polarity and epithelial formation. Moreover, survival rates for bladder cancer at 5 years dropped to ~12% in patients that simultaneously harbor low expression of signature 1 genes (high-risk), which included *KDM8*, and *PCDHA1* mutations. Additive effects conferred by mutations in this cell-adhesion protein provides support for the hypothesis that *KDM8* is likely a tumor suppressor and downregulation of this gene may lead to a loss of epithelial phenotype and cell adhesion to promote cancer invasion. Additionally, loss of *KDM8* expression is correlated with

increased expression of cell-cycle genes that may contribute to deranged cell cycle regulation and tumor progression (Fig.S5D). Like *KDM6B*, the function of *KDM8* appears to be cell-type dependent. While *KDM8* expression is downregulated in liver and pancreatic tumors compared to adjacent cells (Fig.S5A), it is overexpressed in breast cancer to induce EMT and invasion[33]. Nonetheless, *KDM8* roles are not limited to cell cycle regulation. *KDM8* exert tumor suppressive functions in hematopoietic cancer by mediating DNA repair[34]. Collectively, imbalance in the JmjC subfamily of lysine demethylases such as *KDM8* and *KDM6B* are likely to result in broad-ranging but cell-type specific biological effects.

Overall, our gene signatures will enhance decision making in clinic and provide individualized treatment plans based on patients' tumor biology to maximize treatment efficacy and prolong lifespan by directing resources to those most in need. Upon diagnosis, the signatures can be used for risk stratification and patients with poor prognosis can be subjected to additional neoadjuvant treatments. These signatures would need to be assessed prospectively and we aim to evaluate their use in clinical decision making of patients presented with these cancers.

Figure legends

Figure 1. Schematic diagram of the study design and development of signatures derived from oxygen-sensing genes. (A) Three pancreatic cancer cohorts were used to define both signatures 1 and 2. Genes found to be prognostic in Cox regression analysis in two out of three pancreatic cancer cohorts were included in signatures 1 and 2. Signature 1 is a marker of good prognosis and consists of 5 genes (*KDM8, KDM6B, P4HTM, ALKBH4* and *ALKBH7*). Signature 2 is a marker of adverse prognosis and consists of 5 genes (*KDM3A, P4HA1, ASPH, PLOD1* and *PLOD2*). Prognosis of both signatures were further confirmed in 10 cancer types using Kaplan-Meier, Cox regression and receiver operating characteristic analyses. **(B)** Forest plots of prognostic genes found to be significant by Cox proportional hazards analysis in pancreatic cancer cohorts abbreviated as PAAD, PACA-AU and PACA-CA. Genes were separated into two groups, good and bad prognostic genes. Hazard ratios (HR) were denoted as red circles and turquoise bars represent 95% confidence interval. Significant Wald test P values were indicated in blue. Y-axes represent gene symbols followed by cohort abbreviations. Signature 1 genes were marked in green. Signature 2 genes were marked in red. Full description of cancers is listed in Table S1.

Figure 2. Kaplan-Meier analyses confirming that gene signatures were associated with patients' overall survival. (A) Validation of signature 1 (green panels) across multiple cancer types. Kaplan-Meier plots of overall survival in cancer patients stratified based on signature 1 mean expression scores. Patients were median-dichotomized into high- and low-score groups. Signature 1 is a marker of good prognosis and hence patients with high signature 1 scores had improved survival rates. **(B)** Validation of signature 2 (red panels) across multiple cancer types. Kaplan-Meier plots of overall survival in cancer patients stratified based on signature 2 mean expression scores. Patients were median-dichotomized into high- and low-score groups. Signature 2 is a marker of adverse prognosis and hence patients with high signature 2 scores had poorer survival rates. P-values were calculated from the log-rank test. Pancreas #1 = PAAD cohort; Pancreas #2 = PACA-AU cohort; Pancreas #3 = PACA-CA cohort; Liver #1 = LIHC cohort; Liver #2 = LIRI-JP cohort and Liver #3 = GSE14520 cohort (Table S1).

Figure 3. Tumor subgroup analyses and evaluation of predictive performance. Gene signatures were prognostic across different malignant grades. Kaplan-Meier plots show independence of **(A)** signature 1 (green panels) and **(B)** signature 2 (red panels) over current staging systems in cancer cohorts. Patients were sub-grouped according to TNM stages and further stratified using either

signature 1 or signature 2 scores. Both signatures successfully identified high-risk patients in different TNM stages. P-values were calculated from the log-rank test. Analysis of specificity and sensitivity of **(C)** signature 1 (green panels) and **(D)** signature 2 (red panels) in cancer cohorts using with Receiver Operating Characteristic (ROC). Plots depict comparison of ROC curves of signatures 1 or 2 and clinical tumor staging parameters. Both signatures demonstrated incremental values over current staging systems. AUC: area under the curve. TNM: tumor, node, metastasis staging. Liver #2 = LIRI-JP cohort and Liver #3 = GSE14520 cohort (Table S1). Representative plots are depicted in this figure. Additional plots are available in Figure S1.

Figure 4. Relationship between patients' risks as determined by gene signatures and common genetic mutations. Patients were median-stratified into low or high-risk groups using **(A)** signature 1 (green panels) and **(B)** signature 2 (red panels). Since signature 1 is a marker of good prognosis, high-risk patients had lower mean expression of signature 1 genes. Signature 2 is a marker of poor prognosis, hence high-risk patients had higher mean expression of signature 2 genes. Kaplan-Meier plots depicting combined relation of somatic mutations with signatures 1 or 2 on overall survival in cancer patients. P-values were calculated from the log-rank test.

Figure 5. *KDM8* expression is associated with good prognosis in liver and pancreatic cancers. Kaplan-Meier analysis of patients stratified by *KDM8* expression. Patients were median-dichotomized into low- and high-expression groups. Patients with low *KDM8* expression had significantly poorer overall survival. This was consistent in patients analyzed as a full cohort or sub-categorized according to tumor TNM stage. Liver #1 = LIHC cohort; Liver #2 = LIRI-JP cohort and Liver #3 = GSE14520 cohort (Table S1).

Supplementary figure legends

Figure S1. Additional tumor subgroup analyses and evaluation of predictive performance. Additional plots from Figure 3. Gene signatures were prognostic across different malignant grades. Kaplan-Meier plots show independence of **(A)** signature 1 (green panels) and **(B)** signature 2 (red panels) over current staging systems in cancer cohorts. Patients were sub-grouped according to TNM stages and further stratified using either signature 1 or signature 2 scores. Both signatures successfully identified high-risk patients in different TNM stages. P-values were calculated from the log-rank test. Analysis of specificity and sensitivity of **(C)** signature 1 (green panels) and **(D)** signature 2 (red panels) in cancer cohorts using with Receiver Operating Characteristic (ROC). Plots depict comparison of ROC curves of signatures 1 or 2 and clinical tumor staging parameters. Both signatures demonstrated incremental values over current staging systems. AUC: area under the curve. TNM: tumor, node, metastasis staging. Liver #1 = LIHC cohort; Liver #2 = LIRI-JP cohort and Liver #3 = GSE14520 cohort (Table S1).

Figure S2. Expression of distribution of signature genes in low- and high-risk patients. (A) signature 1 (green panels) and **(B)** signature 2 (red panels). Patients were median-stratified into low- and high-risk groups based on mean expression of signature genes. Box plots depict expression distribution of each of the 5 genes in both signatures in these two patient groups. **(A)** Since signature 1 is a marker of good prognosis, high-risk patients showed significantly lower expression of individual signature genes. **(B)** In contrast, signature 2 is a marker of poor prognosis, hence high-risk patients showed significantly higher expression of individual signature genes. Nonparametric Mann-Whitney-Wilcoxon tests were used to compare low- and high-risk patients. Asterisks represent significant P values: * < 0.01, ** < 0.001 and *** < 0.0001. LR = low risk. HR = high risk.

Figure S3. Correlation of patients' risk scores derived from signature 1 with tumor hypoxia. (A) Significant negative correlation between signature 1 expression scores and tumor hypoxia. **(B)** Significant positive correlation between signature 1 risk scores and tumor hypoxia. Calculations of expression scores, risk scores and hypoxia scores are explained in the methods. Liver #1 = LIHC cohort and Liver #2 = LIRI-JP cohort.

Figure S4. Correlation of patients' risk scores derived from signature 2 with tumor hypoxia. (A) Significant positive correlation between signature 2 expression scores and tumor hypoxia. **(B)**

Significant positive correlation between signature 2 risk scores and tumor hypoxia. Calculations of expression scores, risk scores and hypoxia scores are explained in the methods. Liver #2 = LIRI-JP cohort and Liver #3 = GSE14520 cohort.

Figure S5. Putative tumor suppressive functions of *KDM8* occur through processes related to cell cycle regulation and cell adhesion. (A) Expression distributions of *KDM8* were significantly lower in tumor (T) compared to non-tumor (NT) samples in liver and pancreatic cancer cohorts. Mann-Whitney-Wilcoxon tests were used to compare T and NT samples. Asterisks represent significant P values: *** < 0.0001. **(B)** Expression levels of *KDM8* decreased according to disease progression and malignant grade in liver and pancreatic cancer cohorts. **(C)** Significant negative correlation between patients' *KDM8* expression and tumor hypoxia (hypoxia score) in liver and pancreatic cancer cohorts. **(D)** Correlation between *KDM8* expression and canonical cell-cycle regulators in patients with liver or pancreatic cancers. A majority of genes involved in cell-cycle regulation are negatively correlated with *KDM8* expression. Liver #1 = LIHC cohort; Liver #2 = LIRI-JP cohort and Liver #3 = GSE14520 cohort (Table S1). **(E)** Patients were median-stratified according to KDM8 expression. Differential expression analysis between *KDM8* high- and low- groups in liver cancer revealed 745 differentially expressed genes (DEGs; fold-change >2 or <-2). Enrichment of biological pathways associated with DEGs, which include processes related to cell adhesion, inflammation, metabolism and signal transduction pathways related to cancer. **(F)** Enrichment of transcription factors (TFs) from the ENCODE database that are potential regulators of *KDM8* DEGs. These TFs were predicted to bind near *KDM8* DEGs. **(G)** Venn diagram depicts the overlap between HNF4A targets (as identified by ENCODE chromatin-immunoprecipitation sequencing dataset) and genes affected by HNF4A loss-of-function (as identified in HNF4A-null mice). Of the 745 DE genes, 148 genes were identified as direct HNF4A targets and 110 genes were affected by HNF4A loss-of-function. In the Venn intersection, 45 genes were both HNF4A targets and altered in HNF4A-null mice. **(H)** Scatter plot depicts expression patterns of 110 genes affected by HNF4A loss-of-function. Gene names of the 45 HNF4A targets were annotated on the plot. A majority of *KDM8*-associated genes were downregulated in the HNF4A-null mice.

Supplementary table legends

Table S1. Cancer cohort descriptions.

Table S2. List of 61 2OG oxygenases.

Table S3. Univariate and multivariate Cox proportional hazards analysis of risk factors associated with overall survival in multiple cancers. Refer to Table S1 for cancer abbreviations.

Table S4. Differentially expressed genes between *KDM8*-high and -low groups in hepatocellular carcinoma (LIHC cohort).

Table S5. Significantly enriched biological pathways of differentially expressed genes.

Acknowledgements

A special thank you to Graham Foster for his critical insights and help on manuscript preparation.

Authors contribution

WHC and AGL designed the study and analyzed the data. All authors interpreted the data. WHC and AGL wrote the initial manuscript draft. All authors revised the manuscript and approved the final version.

Competing interests

None.

Supplementary methods

Study cohorts

Cancer datasets used in this study were obtained from The Cancer Genome Atlas (TCGA)[15], International Cancer Genome Consortium[16] and Gene Expression Omnibus (n=20,752). Cohort descriptions were listed in Table S1. Gene expression profiles for TCGA cancers were downloaded from Broad Institute GDAC Firehose. Gene expression profiles for ICGC cancers were downloaded from the ICGC data portal. TCGA Illumina HiSeq rnaseqv2 Level 3 RSEM normalized gene expression and ICGC normalized gene expression profiles were converted to $\log_2(x + 1)$ scale.

KDM8 differential expression analysis

TCGA liver cancer cohort (LIHC; Table S1) was used. Patients (n=371) were median dichotomized into low and high groups based on *KDM8* expression. To determine differentially expressed genes between both groups, the Bayes method and linear model were employed using the R package limma. P-values were adjusted using the false discovery rate controlling procedure of Benjamini-Hochberg. Genes with $\log_2$ fold change of > 1 or < -1 and adjusted P-values < 0.05 were considered significant.

Gene signature and risk scores

Expression scores for signature 1 and 2 were calculated for each patient by taking the average $\log_2$ expression values of signature genes. As a measure of tumor hypoxia, hypoxia scores were determined by calculating the average $\log_2$ expression values of 52 hypoxia signature genes[19]. Risk scores for each patient were determined by taking the sum of Cox regression coefficient for each of the signature genes multiplied with its corresponding expression value. Nonparametric Spearman's rank correlation was employed to assess the relationship of expression scores and risk scores with tumor hypoxia (hypoxia score).

Survival analyses

Cox proportional hazards regression was employed to investigate the association between patient survival duration and one or more risk factors, e.g. signature 1 or signature 2 scores, tumor stage and other clinical variables. Univariate analyses were performed to determine the influence of individual risk factors on overall survival. Multivariate analyses were performed by including risk factors that were significantly associated with overall survival identified in univariate analyses (P <

0.05) since multiple covariates can influence patient prognosis. Hazard ratios (HR) were determined from Cox models with HR greater than one indicating that a covariate was positively associated with event probability (increased hazard) and negatively associated with survival length. Cox regression analyses were performed using the R survival and survminer packages. Proportional hazards assumption was supported by a non-significant relationship between scaled Schoenfeld residuals and time using the R survival package. In addition, Kaplan-Meier and log-rank tests were used in univariate analyses of the gene signatures in relation to patient survival and were performed using the survival and survminer packages. Patients were median-dichotomized into low and high-score groups based on mean expression scores of signature genes. Difference between high and low-score groups were tested using the log-rank test implemented with the survival package.

Time-dependent receiver operating characteristic (ROC) curve analysis was used to assess the predictive performance of both signatures 1 and 2 in comparison with standard tumor staging parameters. The R survcomp package was employed to compute time-dependent ROC curves[35]. ROC curves depicted true positive rates (sensitivity) versus false positive rates (1-specificity). The area under the curve (AUC) is a measure of how well the gene signatures can predict patient survival where AUC ranges from 0.5 (for an uninformative marker) to 1 (for a perfect predictive marker).

Biological enrichment analysis

Analysis of biological pathway enrichment on the 745 differentially expressed genes between *KDM8*-low and -high groups was conducted using GeneCodis against the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) databases[36]. The Enrichr tool was used to identify transcription factors from the ENCODE database that are potential regulators of these 745 genes[37,38].

***HNF4A* loss of function analysis**

A total of 148 genes were identified as *HNF4A* targets in the HepG2 hepatoma cell line determined using the Enrichr tool[37,38]. Differential expression analysis between *HNF4A* wild-type and null mice livers (GSE3126) performed using the GEO2R tool[26] identified 110 differentially expressed genes from the initial 745-gene set identified previously (Fig.S5G). Of these 110 genes, 45 genes were identified as direct *HNF4A* targets and were downregulated in the *HNF4A*-null mice (Fig.S5H).

Somatic mutation identification.

Level 3 mutation datasets were downloaded from GDAC. Kaplan-Meier analysis and log-rank tests were employed to determine the association of somatic mutations in combination with the signatures 1 or 2, on overall survival.

All graphs were generated using the ggplot2 package in R[39].

References

1. Zeng W, Liu P, Pan W, Singh SR, Wei Y. Hypoxia and hypoxia inducible factors in tumor metabolism. Cancer Lett. Elsevier; 2015;356:263–7.

2. Denko NC. Hypoxia, HIF1 and glucose metabolism in the solid tumour. Nat Rev Cancer. Nature Publishing Group; 2008;8:705.

3. Nagy JA, Chang SH, Dvorak AM, Dvorak HF. Why are tumour blood vessels abnormal and why is it important to know? Br J Cancer. Nature Publishing Group; 2009;100:865.

4. Semenza GL, Wang GL. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. Mol Cell Biol. Am Soc Microbiol; 1992;12:5447–54.

5. Ratcliffe PJ. Oxygen sensing and hypoxia signalling pathways in animals: the implications of physiology for cancer. J Physiol. Wiley Online Library; 2013;591:2027–42.

6. Loenarz C, Coleman ML, Boleininger A, Schierwater B, Holland PWH, Ratcliffe PJ, et al. The hypoxia-inducible transcription factor pathway regulates oxygen sensing in the simplest animal, Trichoplax adhaerens. EMBO Rep. EMBO Press; 2011;12:63–70.

7. Semenza GL. Evaluation of HIF-1 inhibitors as anticancer agents. Drug Discov Today. Elsevier; 2007;12:853–9.

8. Melillo G. Inhibiting hypoxia-inducible factor 1 for cancer therapy. Mol Cancer Res. AACR; 2006;4:601–5.

9. Ploumakis A, Coleman ML. OH, the places you'll go! hydroxylation, gene expression, and cancer. Mol Cell. Elsevier; 2015;58:729–41.

10. Ko M, An J, Pastor WA, Koralov SB, Rajewsky K, Rao A. TET proteins and 5-methylcytosine oxidation in hematological cancers. Immunol Rev. Wiley Online Library; 2015;263:6–21.

11. Huang Y, Rao A. Connections between TET proteins and aberrant DNA modification in cancer. Trends Genet. Elsevier; 2014;30:464–74.

12. Højfeldt JW, Agger K, Helin K. Histone lysine demethylases as targets for anticancer therapy. Nat Rev Drug Discov. Nature Publishing Group; 2013;12:917.

13. Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature. Nature Publishing Group; 2010;463:360.

14. Van Haaften G, Dalgliesh GL, Davies H, Chen L, Bignell G, Greenman C, et al. Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. Nat Genet. Nature Publishing Group; 2009;41:521.

15. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. Nat Genet. Nature Publishing Group; 2013;45:1113.

16. Zhang J, Baran J, Cros A, Guberman JM, Haider S, Hsu J, et al. International Cancer Genome Consortium Data Portal—a one-stop shop for cancer genomics data. Database. Oxford University Press; 2011;2011.

17. Roessler S, Jia HL, Budhu A, Forgues M, Ye QH, Lee JS, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. Cancer Res. 2010;70:10202–12.

18. Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, et al. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. Nat Genet. Nature Publishing Group; 2016;48:500.

19. Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. Br J Cancer. 2010;102:428–35.

20. Vaupel P, Mayer A. Hypoxia in cancer: significance and impact on clinical outcome. Cancer Metastasis Rev. Springer; 2007;26:225–39.

21. Kim SM, Leem SH, Chu IS, Park YY, Kim SC, Kim SB, et al. Sixty-five gene-based risk score classifier predicts overall survival in hepatocellular carcinoma. Hepatology. 2012;55:1443–52.

22. King RJ, Yu F, Singh PK. Genomic alterations in mucins across cancers. Oncotarget. Impact Journals, LLC; 2017;8:67152.

23. Wu B-H, Chen H, Cai C-M, Fang J-Z, Wu C-C, Huang L-Y, et al. Epigenetic silencing of JMJD5 promotes the proliferation of hepatocellular carcinoma cells by down-regulating the transcription of CDKN1A. Oncotarget. Impact Journals, LLC; 2016;7:6847.

24. Zhang C, Han Y, Huang H, Qu L, Shou C. High NR2F2 transcript level is associated with increased survival and its expression inhibits TGF-$β$-dependent epithelial-mesenchymal transition in breast cancer. Breast Cancer Res Treat. Springer; 2014;147:265–81.

25. Song Y, Washington MK, Crawford HC. Loss of FOXA1/2 is essential for the epithelial-to-mesenchymal transition in pancreatic cancer. Cancer Res. 2010;70:2115–25.

26. Battle M a, Konopka G, Parviz F, Gaggl AL, Yang C, Sladek FM, et al. Hepatocyte nuclear factor 4alpha orchestrates expression of cell adhesion proteins during the epithelial transformation of the developing liver. Proc Natl Acad Sci U S A. 2006;103:8419–24.

27. Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, et al. Control of pancreas and liver gene expression by HNF transcription factors. Science (80-). American Association for the Advancement of Science; 2004;303:1378–81.

28. Consortium EP, others. An integrated encyclopedia of DNA elements in the human genome. Nature. Nature Publishing Group; 2012;489:57.

29. Yamamoto K, Tateishi K, Kudo Y, Sato T, Yamamoto S, Miyabayashi K, et al. Loss of histone demethylase KDM6B enhances aggressiveness of pancreatic cancer through downregulation of C/EBP$α$. Carcinogenesis. Oxford University Press UK; 2014;35:2404–14.

30. Tokunaga R, Sakamoto Y, Nakagawa S, Miyake K, Izumi D, Kosumi K, et al. The prognostic significance of histone lysine demethylase JMJD3/KDM6B in colorectal cancer. Ann Surg Oncol. Springer; 2016;23:678–85.

31. Li Q, Hou L, Ding G, Li Y, Wang J, Qian B, et al. KDM6B induces epithelial-mesenchymal transition and enhances clear cell renal cell carcinoma metastasis through the activation of SLUG. Int J Clin Exp Pathol. e-Century Publishing Corporation; 2015;8:6334.

32. Ramadoss S, Chen X, Wang C-Y. Histone demethylase KDM6B promotes epithelial-mesenchymal transition. J Biol Chem. ASBMB; 2012;jbc--M112.

33. Zhao Z, Sun C, Li F, Han J, Li X, Song Z. Overexpression of histone demethylase JMJD5 promotes metastasis and indicates a poor prognosis in breast cancer. Int J Clin Exp Pathol. e-Century Publishing Corporation; 2015;8:10325.

34. Suzuki T, Minehata K, Akagi K, Jenkins NA, Copeland NG. Tumor suppressor gene identification using retroviral insertional mutagenesis in Blm-deficient mice. EMBO J. EMBO Press; 2006;25:3422–31.

35. Schroeder M, Culhane A, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. Bioinformatics. 2011;27(22):3206–3208.

36. Tabas-Madrid D, Nogales-Cadenas R, Pascual-Montano A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res. Oxford University Press; 2012;40:W478--W483.

37. Kuleshov M V, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. Oxford University Press; 2016;44:W90--W97.

38. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. BioMed Central; 2013;14:128.

39. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016.

Figure 1

A

2 OG oxygenases (61 genes)
Cox regression analyses

25 TCGA cancers
n (tumor samples) = 19,781

3 TRAINING COHORTS

TCGA pancreatic cancer (29 / 61 prognostic genes)
PAAD, n = 178

ICGC pancreatic cancer cohorts
PACA-AU, n = 269
PACA-CA, n = 234

Genes prognostic in 2 out of 3 training cohorts

Signature 1 (good prognosis)
KDM8
KDM6B
P4HTM
ALKBH4
ALKBH7

Signature 2 (bad prognosis)
KDM3A
P4HA1
ASPH
PLOD1
PLOD2

PAN-CANCER VALIDATION

Signature 1
8 prognostic cohorts

Bladder	BLCA	n = 408
Papillary cell	KIRP	n = 290
Liver #1	LIHC	n = 371
Liver #2	LIRI-JP	n = 226
Lung	LUAD	n = 515
Pancreas #1	PAAD	n = 178
Pancreas #2	PACA-AU	n = 269
Endometrial	UCEC	n = 370

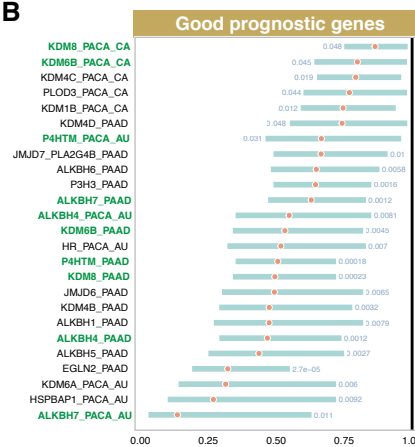
total = 2,627

Signature 2
12 prognostic cohorts

Bladder	BLCA	n = 408
Cervical	CESC	n = 304
Head and neck	HNSC	n = 520
Renal clear cell	KIRC	n = 533
Papillary cell	KIRP	n = 290
Liver #2	LIRI-JP	n = 226
Liver #3	GSE14520	n = 242
Lung	LUAD	n = 515
Pancreas #1	PAAD	n = 178
Pancreas #2	PACA-AU	n = 269
Pancreas #3	PACA-CA	n = 234
Stomach	STAD	n = 415

total = 4,134

B



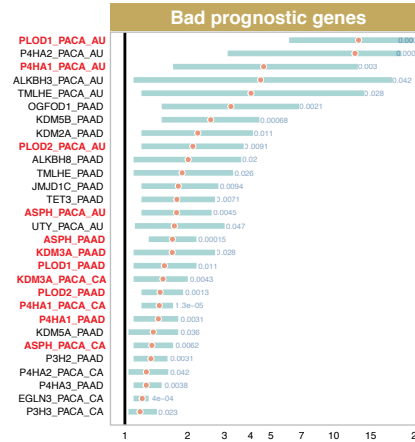


Figure 2

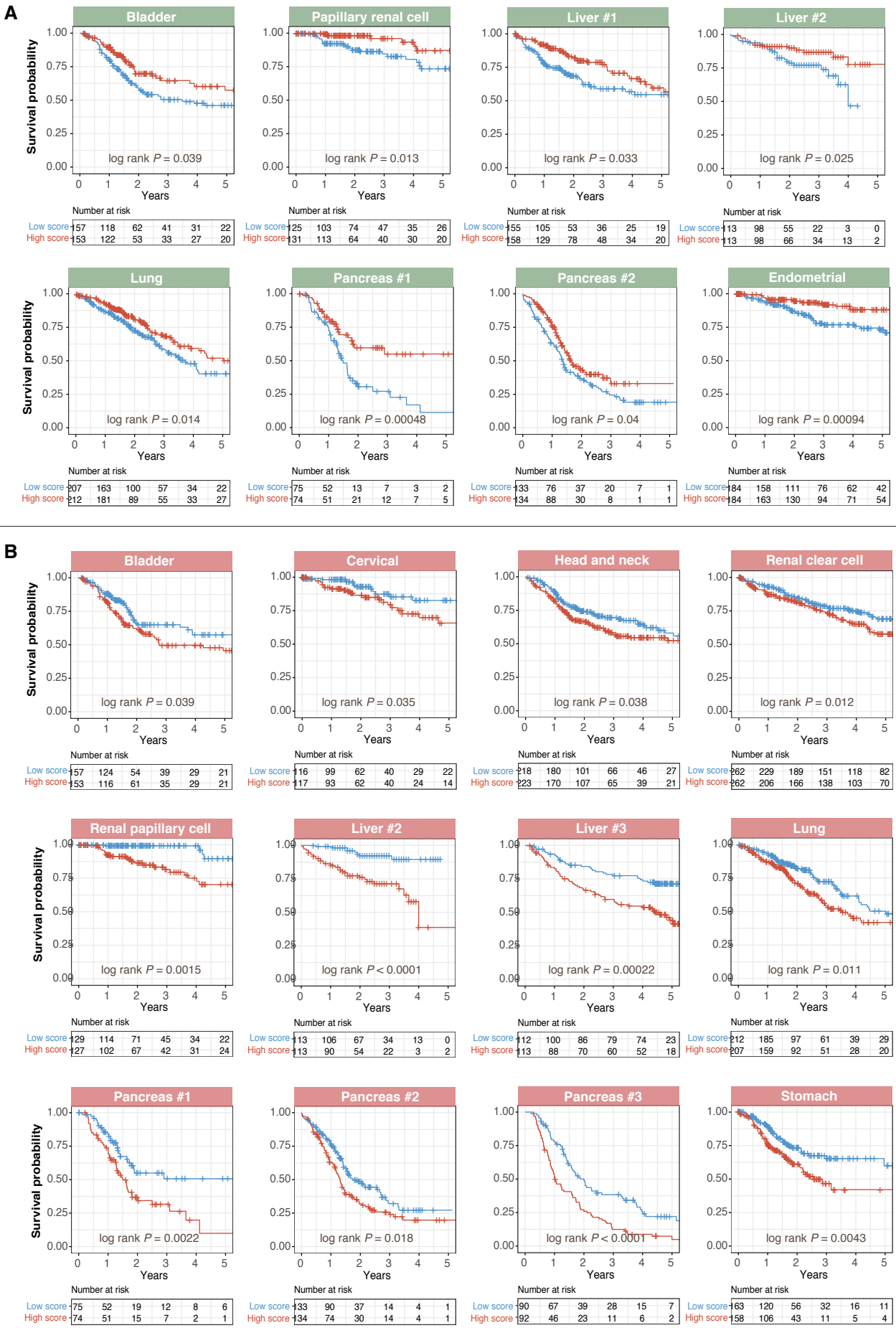

A
Bladder
log rank P = 0.039
Number at risk
Low score 157 118 62 41 31 22
High score 153 122 53 33 27 20
Papillary renal cell
log rank P = 0.013
Number at risk
Low score 125 103 74 47 35 26
High score 131 113 64 40 30 20
Liver #1
log rank P = 0.033
Number at risk
Low score 155 105 53 36 25 19
High score 158 129 78 48 34 20
Liver #2
log rank P = 0.025
Number at risk
Low score 113 98 55 22 3 0
High score 113 98 66 34 13 2
Lung
log rank P = 0.014
Number at risk
Low score 207 163 100 57 34 22
High score 212 181 89 55 33 27
Pancreas #1
log rank P = 0.00048
Number at risk
Low score 75 52 13 7 3 2
High score 74 51 21 12 7 5
Pancreas #2
log rank P = 0.04
Number at risk
Low score 133 76 37 20 7 1
High score 134 88 30 8 1 1
Endometrial
log rank P = 0.00094
Number at risk
Low score 184 158 111 76 62 42
High score 184 163 130 94 71 54
B
Bladder
log rank P = 0.039
Number at risk
Low score 157 124 54 39 29 21
High score 153 116 61 35 29 21
Cervical
log rank P = 0.035
Number at risk
Low score 116 99 62 40 29 22
High score 117 93 62 40 24 14
Head and neck
log rank P = 0.038
Number at risk
Low score 218 180 101 66 46 27
High score 223 170 107 65 39 21
Renal clear cell
log rank P = 0.012
Number at risk
Low score 262 229 189 151 118 82
High score 262 206 166 138 103 70
Renal papillary cell
log rank P = 0.0015
Number at risk
Low score 129 114 71 45 34 22
High score 127 102 67 42 31 24
Liver #2
log rank P < 0.0001
Number at risk
Low score 113 106 67 34 13 0
High score 113 90 54 22 3 2
Liver #3
log rank P = 0.00022
Number at risk
Low score 112 100 86 79 74 23
High score 113 88 70 60 52 18
Lung
log rank P = 0.011
Number at risk
Low score 212 185 97 61 39 29
High score 207 159 92 51 28 20
Pancreas #1
log rank P = 0.0022
Number at risk
Low score 75 52 19 12 8 6
High score 74 51 15 7 2 1
Pancreas #2
log rank P = 0.018
Number at risk
Low score 133 90 37 14 4 1
High score 134 74 30 14 4 1
Pancreas #3
log rank P < 0.0001
Number at risk
Low score 90 67 39 28 15 7
High score 92 46 23 11 6 2
Stomach
log rank P = 0.0043
Number at risk
Low score 163 120 56 32 16 11
High score 158 106 43 11 5 4
Survival probability
Years

Figure 3

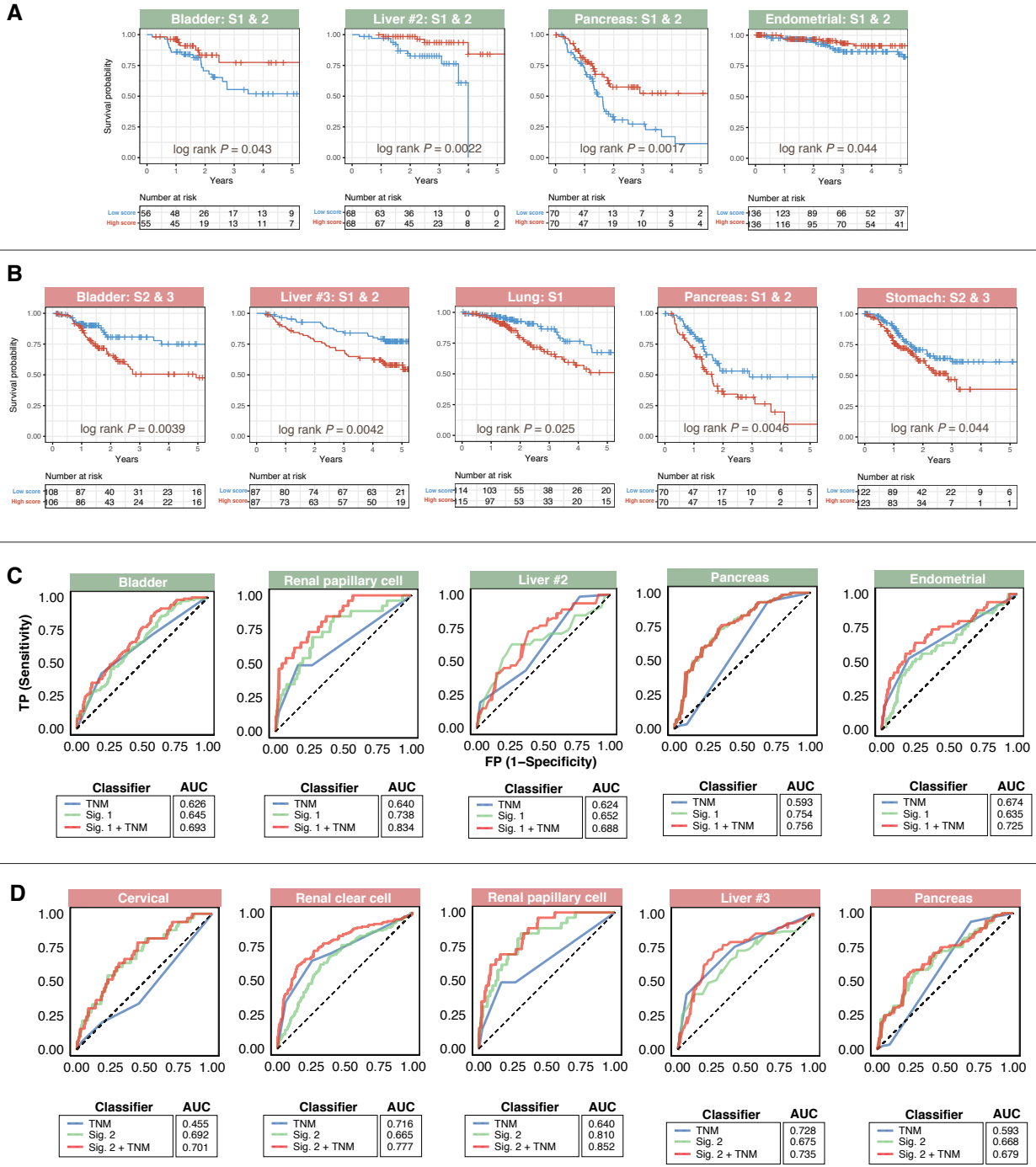
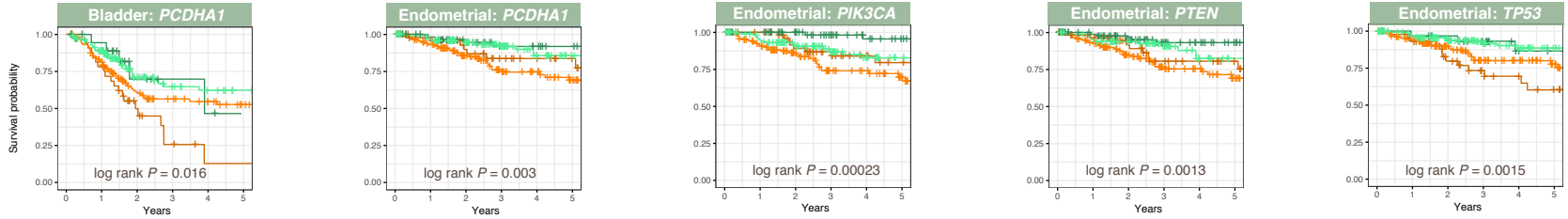

A
Bladder: S1 & 2
Liver #2: S1 & 2
Pancreas: S1 & 2
Endometrial: S1 & 2
Survival probability
Years
log rank P = 0.043
log rank P = 0.0022
log rank P = 0.0017
log rank P = 0.044
Number at risk
Low score 56 48 26 17 13 9
High score 55 45 19 13 11 7
Low score 68 63 36 13 0 0
High score 68 67 45 23 8 2
Low score 70 47 13 7 3 2
High score 70 47 19 10 5 4
Low score 136 123 89 66 52 37
High score 136 116 95 70 54 41
B
Bladder: S2 & 3
Liver #3: S1 & 2
Lung: S1
Pancreas: S1 & 2
Stomach: S2 & 3
log rank P = 0.0039
log rank P = 0.0042
log rank P = 0.025
log rank P = 0.0046
log rank P = 0.044
Low score 108 87 40 31 23 16
High score 106 86 43 24 22 16
Low score 87 80 74 67 63 21
High score 87 73 63 57 50 19
Low score 114 103 55 38 26 20
High score 115 97 53 33 20 15
Low score 70 47 17 10 6 5
High score 70 47 15 7 2 1
Low score 122 89 42 22 9 6
High score 123 83 34 7 1 1
C
Bladder
Renal papillary cell
Liver #2
Pancreas
Endometrial
TP (Sensitivity)
FP (1−Specificity)
Classifier AUC
TNM 0.626
Sig. 1 0.645
Sig. 1 + TNM 0.693
Classifier AUC
TNM 0.640
Sig. 1 0.738
Sig. 1 + TNM 0.834
Classifier AUC
TNM 0.624
Sig. 1 0.652
Sig. 1 + TNM 0.688
Classifier AUC
TNM 0.593
Sig. 1 0.754
Sig. 1 + TNM 0.756
Classifier AUC
TNM 0.674
Sig. 1 0.635
Sig. 1 + TNM 0.725
D
Cervical
Renal clear cell
Renal papillary cell
Liver #3
Pancreas
Classifier AUC
TNM 0.455
Sig. 2 0.692
Sig. 2 + TNM 0.701
Classifier AUC
TNM 0.716
Sig. 2 0.665
Sig. 2 + TNM 0.777
Classifier AUC
TNM 0.640
Sig. 2 0.810
Sig. 2 + TNM 0.852
Classifier AUC
TNM 0.728
Sig. 2 0.675
Sig. 2 + TNM 0.735
Classifier AUC
TNM 0.593
Sig. 2 0.668
Sig. 2 + TNM 0.679

Figure 4

A



Number at risk

	0	1	2	3	4	5
High Risk, *PCDHA1* mutant	34	25	10	4	1	1
High Risk, *PCDHA1* WT	122	92	51	36	30	21
Low Risk, *PCDHA1* mutant	20	17	6	3	2	0
Low Risk, *PCDHA1* WT	132	104	47	30	25	20

Number at risk

	0	1	2	3	4	5
High Risk, *PCDHA1* mutant	47	42	31	22	18	13
High Risk, *PCDHA1* WT	137	116	80	54	44	29
Low Risk, *PCDHA1* mutant	67	57	45	34	31	27
Low Risk, *PCDHA1* WT	117	106	85	60	40	27

Number at risk

	0	1	2	3	4	5
High Risk, *PIK3CA* mutant	56	53	40	29	21	16
High Risk, *PIK3CA* WT	128	105	71	47	41	26
Low Risk, *PIK3CA* mutant	74	69	58	45	36	30
Low Risk, *PIK3CA* WT	110	94	72	49	35	24

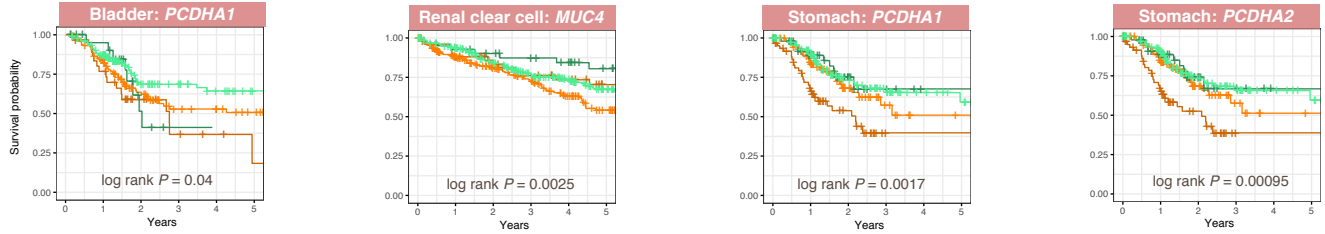
Number at risk

	0	1	2	3	4	5
High Risk, *PTEN* mutant	59	51	36	25	21	15
High Risk, *PTEN* WT	125	107	75	51	41	27
Low Risk, *PTEN* mutant	98	85	67	49	42	34
Low Risk, *PTEN* WT	86	78	63	45	29	20

Number at risk

	0	1	2	3	4	5
High Risk, *TP53* mutant	45	40	28	19	16	11
High Risk, *TP53* WT	139	118	83	57	46	31
Low Risk, *TP53* mutant	32	30	27	21	12	10
Low Risk, *TP53* WT	152	133	103	73	59	44

B



Number at risk

	0	1	2	3	4	5
High Risk, *PCDHA1* mutant	31	23	11	5	3	1
High Risk, *PCDHA1* WT	120	91	49	29	26	20
Low Risk, *PCDHA1* mutant	23	19	5	2	0	0
Low Risk, *PCDHA1* WT	134	105	49	37	29	21

Number at risk

	0	1	2	3	4	5
High Risk, *MUC4* mutant	52	43	36	31	27	19
High Risk, *MUC4* WT	210	163	130	107	76	51
Low Risk, *MUC4* mutant	41	38	34	30	26	20
Low Risk, *MUC4* WT	219	190	155	121	92	62

Number at risk

	0	1	2	3	4	5
High Risk, *PCDHA1* mutant	61	36	16	1	1	1
High Risk, *PCDHA1* WT	90	66	26	10	4	3
Low Risk, *PCDHA1* mutant	51	35	13	5	1	1
Low Risk, *PCDHA1* WT	108	84	43	27	15	10

Number at risk

	0	1	2	3	4	5
High Risk, *PCDHA2* mutant	59	34	16	1	1	1
High Risk, *PCDHA2* WT	92	68	26	10	4	3
Low Risk, *PCDHA2* mutant	49	34	13	5	1	1
Low Risk, *PCDHA2* WT	110	85	43	27	15	10

Figure 5

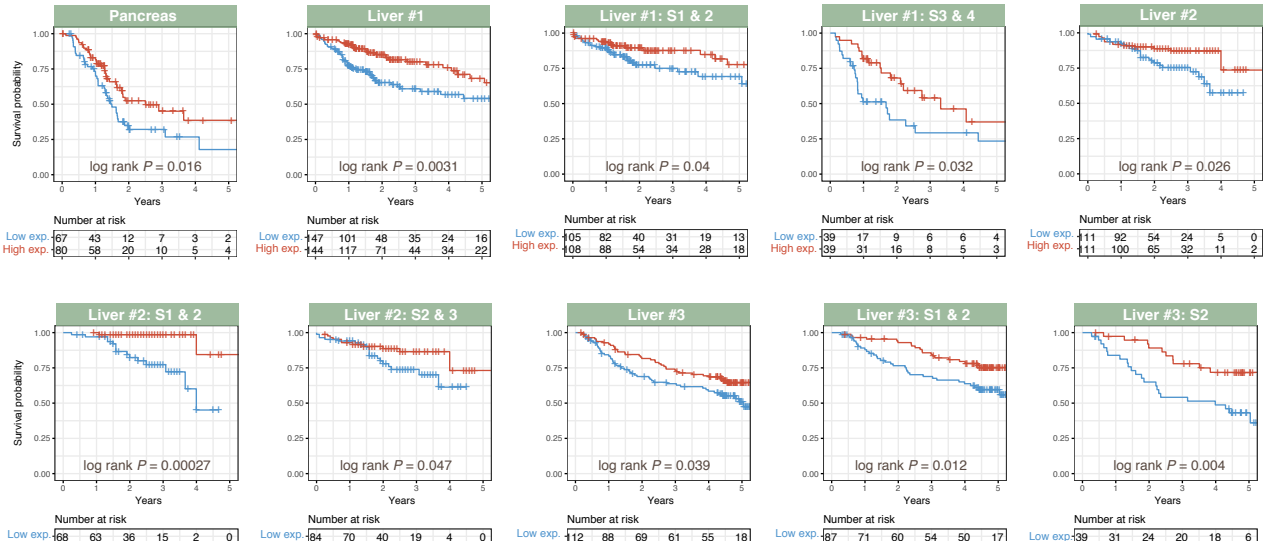
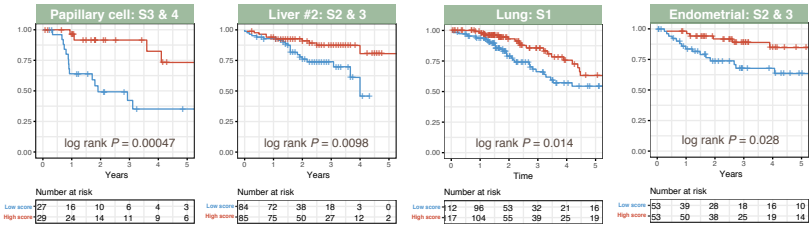
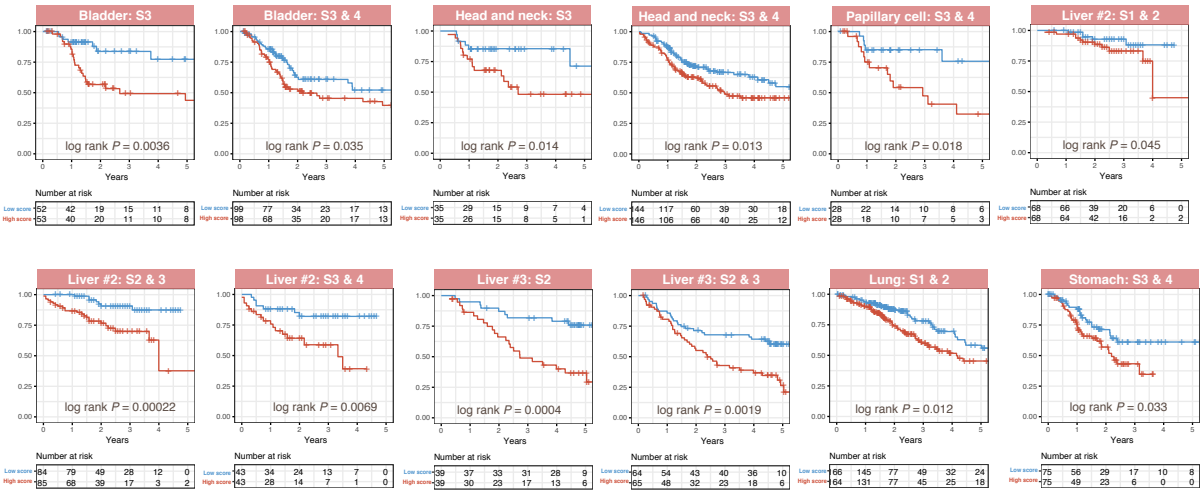

Pancreas
Survival probability
1.00
0.75
0.50
0.25
0.00
log rank P = 0.016
0 1 2 3 4 5
Years
Number at risk
Low exp. 67 43 12 7 3 2
High exp. 80 58 20 10 5 4
Liver #1
log rank P = 0.0031
Years
Number at risk
Low exp. 147 101 48 35 24 16
High exp. 144 117 71 44 34 22
Liver #1: S1 & 2
log rank P = 0.04
Years
Number at risk
Low exp. 105 82 40 31 19 13
High exp. 108 88 54 34 28 18
Liver #1: S3 & 4
log rank P = 0.032
Years
Number at risk
Low exp. 39 17 9 6 6 4
High exp. 39 31 16 8 5 3
Liver #2
log rank P = 0.026
Years
Number at risk
Low exp. 111 92 54 24 5 0
High exp. 111 100 65 32 11 2
Liver #2: S1 & 2
log rank P = 0.00027
Years
Number at risk
Low exp. 68 63 36 15 2 0
High exp. 68 67 45 21 6 2
Liver #2: S2 & 3
log rank P = 0.047
Years
Number at risk
Low exp. 84 70 40 19 4 0
High exp. 85 77 48 26 11 2
Liver #3
log rank P = 0.039
Years
Number at risk
Low exp. 112 88 69 61 55 18
High exp. 113 100 87 78 71 23
Liver #3: S1 & 2
log rank P = 0.012
Years
Number at risk
Low exp. 87 71 60 54 50 17
High exp. 87 82 77 70 63 23
Liver #3: S2
log rank P = 0.004
Years
Number at risk
Low exp. 39 31 24 20 18 6
High exp. 39 36 32 28 23 9

Figure S1

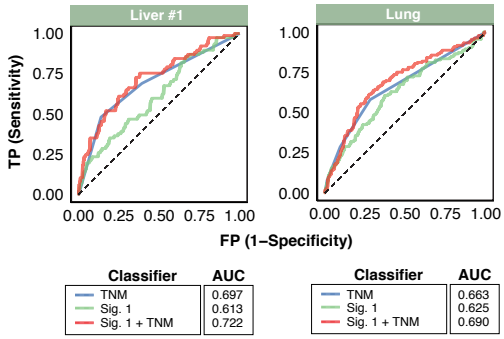
A



B



C



D

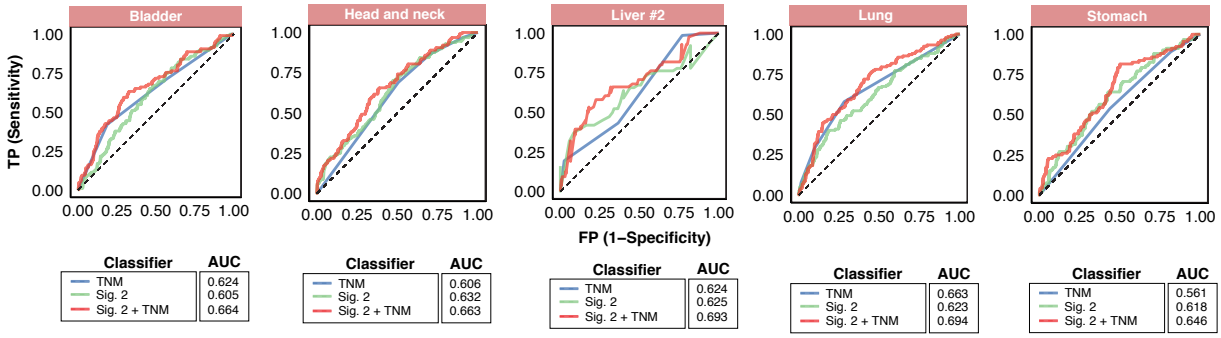


Figure S2

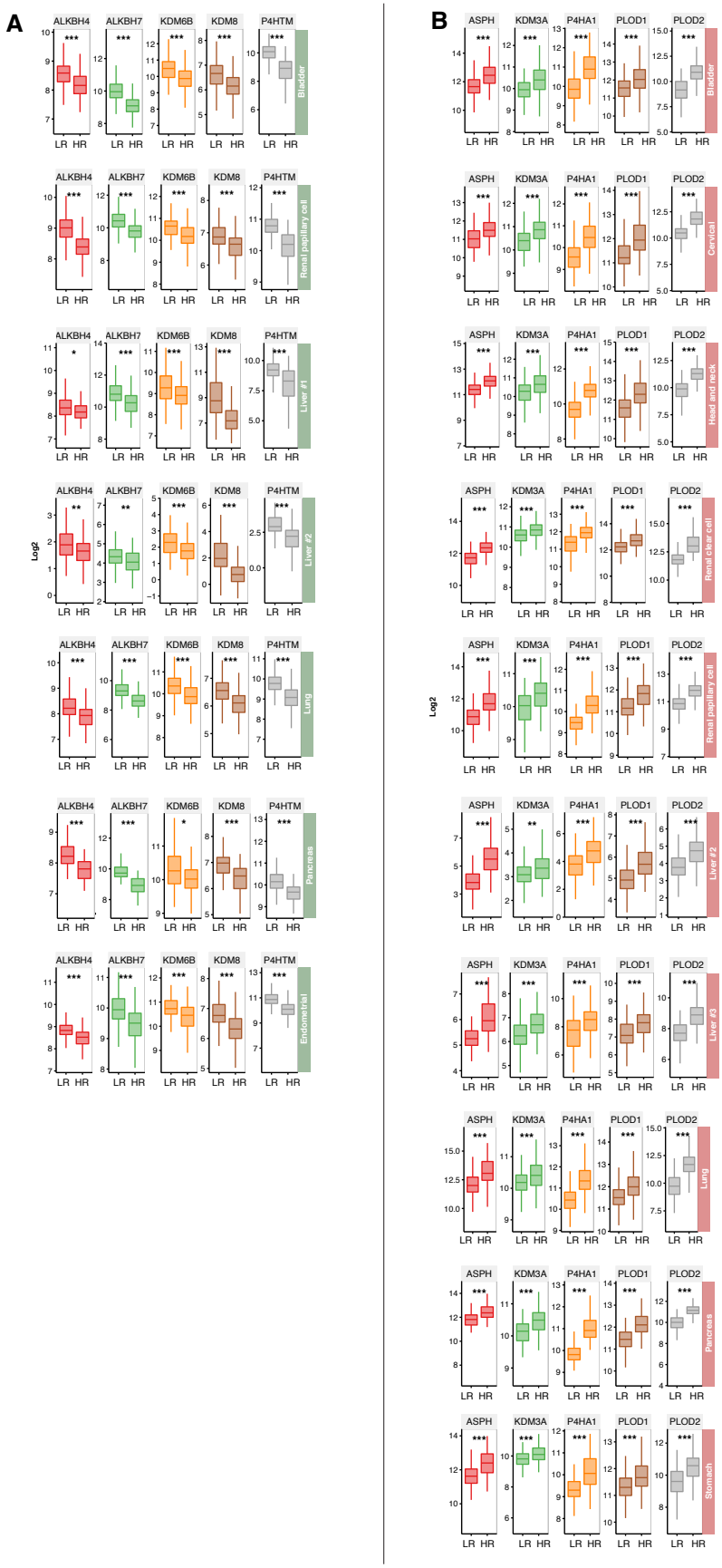

A
B
ALKBH4
ALKBH7
KDM6B
KDM8
P4HTM
ASPH
KDM3A
P4HA1
PLOD1
PLOD2
LR
HR
Log2
Bladder
Renal papillary cell
Liver #1
Liver #2
Lung
Pancreas
Endometrial
Cervical
Head and neck
Renal clear cell
Liver #3
Stomach

Figure S3

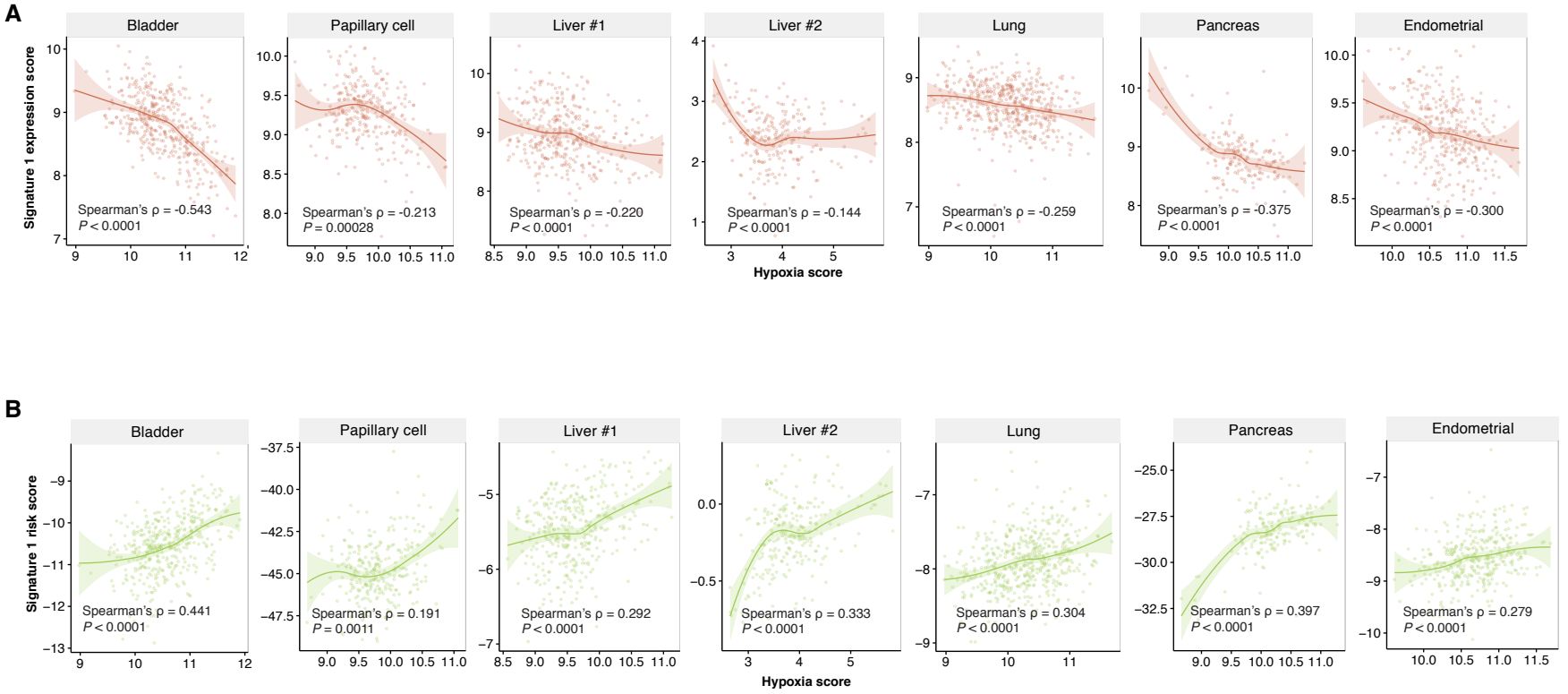

A
Bladder
Papillary cell
Liver #1
Liver #2
Lung
Pancreas
Endometrial
Signature 1 expression score
Spearman's ρ = -0.543
P < 0.0001
Spearman's ρ = -0.213
P = 0.00028
Spearman's ρ = -0.220
P < 0.0001
Spearman's ρ = -0.144
P < 0.0001
Spearman's ρ = -0.259
P < 0.0001
Spearman's ρ = -0.375
P < 0.0001
Spearman's ρ = -0.300
P < 0.0001
Hypoxia score
B
Bladder
Papillary cell
Liver #1
Liver #2
Lung
Pancreas
Endometrial
Signature 1 risk score
Spearman's ρ = 0.441
P < 0.0001
Spearman's ρ = 0.191
P = 0.0011
Spearman's ρ = 0.292
P < 0.0001
Spearman's ρ = 0.333
P < 0.0001
Spearman's ρ = 0.304
P < 0.0001
Spearman's ρ = 0.397
P < 0.0001
Spearman's ρ = 0.279
P < 0.0001
Hypoxia score

Figure S4

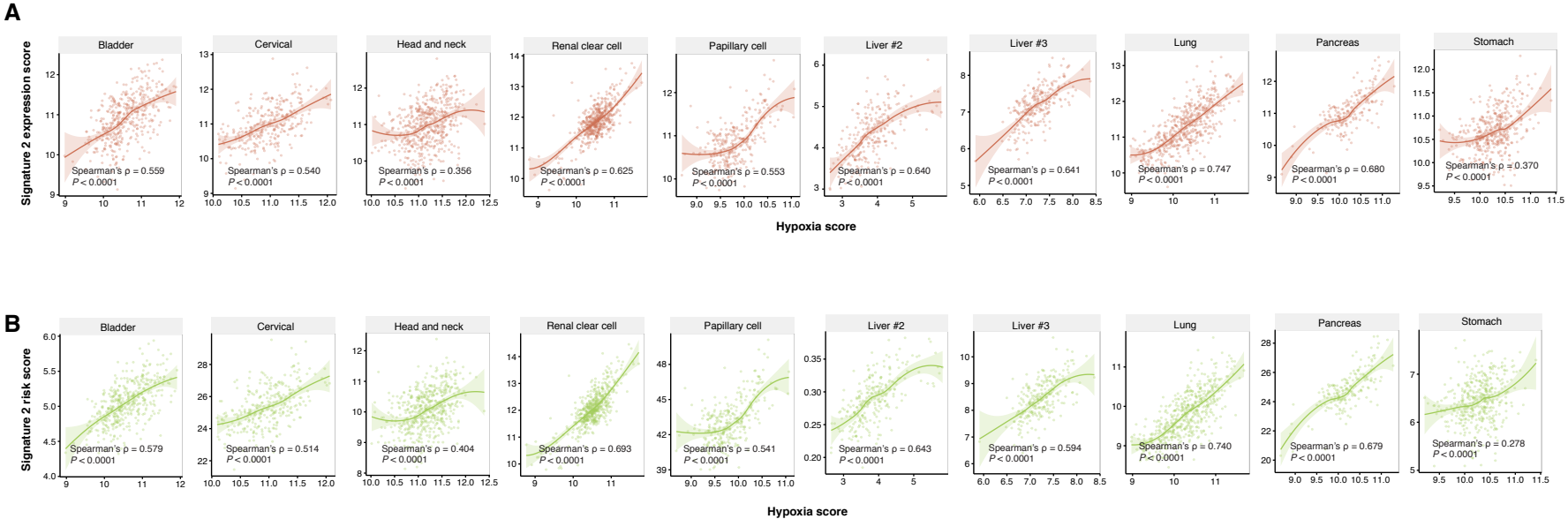

A
Signature 2 expression score
Bladder
Spearman's ρ = 0.559
P < 0.0001
Cervical
Spearman's ρ = 0.540
P < 0.0001
Head and neck
Spearman's ρ = 0.356
P < 0.0001
Renal clear cell
Spearman's ρ = 0.625
P < 0.0001
Papillary cell
Spearman's ρ = 0.553
P < 0.0001
Liver #2
Spearman's ρ = 0.640
P < 0.0001
Liver #3
Spearman's ρ = 0.641
P < 0.0001
Lung
Spearman's ρ = 0.747
P < 0.0001
Pancreas
Spearman's ρ = 0.680
P < 0.0001
Stomach
Spearman's ρ = 0.370
P < 0.0001
Hypoxia score
B
Signature 2 risk score
Bladder
Spearman's ρ = 0.579
P < 0.0001
Cervical
Spearman's ρ = 0.514
P < 0.0001
Head and neck
Spearman's ρ = 0.404
P < 0.0001
Renal clear cell
Spearman's ρ = 0.693
P < 0.0001
Papillary cell
Spearman's ρ = 0.541
P < 0.0001
Liver #2
Spearman's ρ = 0.643
P < 0.0001
Liver #3
Spearman's ρ = 0.594
P < 0.0001
Lung
Spearman's ρ = 0.740
P < 0.0001
Pancreas
Spearman's ρ = 0.679
P < 0.0001
Stomach
Spearman's ρ = 0.278
P < 0.0001
Hypoxia score

Figure S5

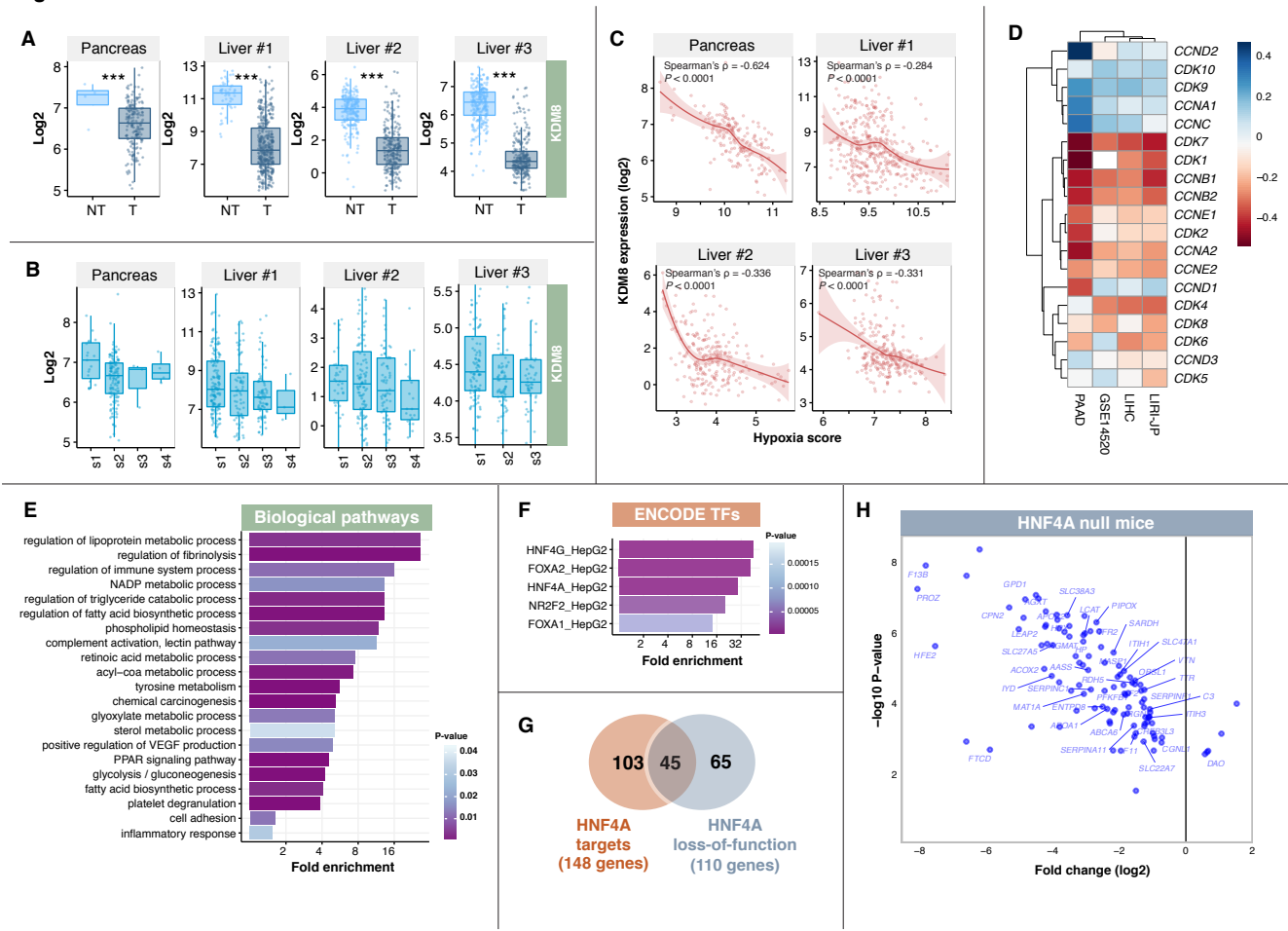

A
Pancreas
Liver #1
Liver #2
Liver #3
Log2
***
NT
T
KDM8
B
s1
s2
s3
s4
C
Spearman's ρ = -0.624
P < 0.0001
Spearman's ρ = -0.284
P < 0.0001
Spearman's ρ = -0.336
P < 0.0001
Spearman's ρ = -0.331
P < 0.0001
KDM8 expression (log2)
Hypoxia score
D
CCND2
CDK10
CDK9
CCNA1
CCNC
CDK7
CDK1
CCNB1
CCNB2
CCNE1
CDK2
CCNA2
CCNE2
CCND1
CDK4
CDK8
CDK6
CCND3
CDK5
PAAD
GSE14520
LIHC
LIRI-JP
E
Biological pathways
regulation of lipoprotein metabolic process
regulation of fibrinolysis
regulation of immune system process
NADP metabolic process
regulation of triglyceride catabolic process
regulation of fatty acid biosynthetic process
phospholipid homeostasis
complement activation, lectin pathway
retinoic acid metabolic process
acyl−coa metabolic process
tyrosine metabolism
chemical carcinogenesis
glyoxylate metabolic process
sterol metabolic process
positive regulation of VEGF production
PPAR signaling pathway
glycolysis / gluconeogenesis
fatty acid biosynthetic process
platelet degranulation
cell adhesion
inflammatory response
P-value
Fold enrichment
F
ENCODE TFs
HNF4G_HepG2
FOXA2_HepG2
HNF4A_HepG2
NR2F2_HepG2
FOXA1_HepG2
P-value
Fold enrichment
G
103
45
65
HNF4A targets (148 genes)
HNF4A loss-of-function (110 genes)
H
HNF4A null mice
−log10 P−value
Fold change (log2)

Table S1.
Cancer abbreviations.

Cohort	Non-tumour #	Tumour #	Description
BLCA	19	408	Bladder Urothelial Carcinoma
BRCA	112	10939	Breast invasive carcinoma
CESC	3	304	Cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	9	36	Cholangiocarcinoma
COAD	41	285	Colon adenocarcinoma
ESCA	11	184	Esophageal carcinoma
GBM	5	153	Glioblastoma multiforme
GBMLGG	5	669	Glioma
HNSC	44	520	Head and Neck squamous cell carcinoma
KICH	25	66	Kidney Chromophobe
KIPAN	129	889	Pan-kidney cohort
KIRC	72	533	Kidney renal clear cell carcinoma
KIRP	32	290	Kidney renal papillary cell carcinoma
LIHC (Liver cohort 1)	50	371	Hepatocellular carcinoma
LUAD	59	515	Lung adenocarcinoma
LUSC	51	501	Lung squamous cell carcinoma
PAAD (Pancreas cohort 1)	4	178	Pancreatic adenocarcinoma
PCPG	3	179	Pheochromocytoma and Paraganglioma
PRAD	52	497	Prostate adenocarcinoma
SARC	2	259	Sarcoma
STAD	35	415	Stomach adenocarcinoma
STES	46	599	Stomach and Esophageal carcinoma
THCA	59	501	Thyroid carcinoma
THYM	2	120	Thymoma
UCEC	11	370	Uterine Corpus Endometrial Carcinoma
LIRI-JP (Liver cohort 2)	NA	226	Hepatocellular carcinoma
GSE14520 (Liver cohort 3)	NA	242	Hepatocellular carcinoma
PACA-AU (Pancreas cohort 2)	NA	269	Pancreatic adenocarcinoma
PACA-CA (Pancreas cohort 3)	NA	234	Pancreatic adenocarcinoma

Table S2. List of 61 2OG oxygenases

Gene Symbol	Description
ALKBH1	alkB homolog 1, histone H2A dioxygenase
ALKBH2	alkB homolog 2, alpha-ketoglutarate dependent dioxygenase
ALKBH3	alkB homolog 3, alpha-ketoglutaratedependent dioxygenase
ALKBH4	alkB homolog 4, lysine demethylase
ALKBH5	alkB homolog 5, RNA demethylase
ALKBH6	alkB homolog 6
ALKBH7	alkB homolog 7
ALKBH8	alkB homolog 8, tRNA methyltransferase
ASPH	aspartate beta-hydroxylase
BBOX1	gamma-butyrobetaine hydroxylase 1
EGLN1	egl-9 family hypoxia inducible factor 1
EGLN2	egl-9 family hypoxia inducible factor 2
EGLN3	egl-9 family hypoxia inducible factor 3
FTO	FTO, alpha-ketoglutarate dependent dioxygenase
HR	HR, lysine demethylase and nuclear receptor corepressor
HSPBAP1	HSPB1 associated protein 1
JARID2	jumonji and AT-rich interaction domain containing 2
JMJD1C	jumonji domain containing 1C
JMJD4	jumonji domain containing 4
JMJD6	arginine demethylase and lysine hydroxylase
JMJD7_PLA2G4B	jumonji Domain Containing 7-Phospholipase A2, Group IVB (Cytosolic) Read-Through
JMJD8	jumonji domain containing 8
KDM1A	lysine demethylase 1A
KDM1B	lysine demethylase 1B
KDM2A	lysine demethylase 2A
KDM2B	lysine demethylase 2B
KDM3A	lysine demethylase 3A
KDM3B	lysine demethylase 3B
KDM4A	lysine demethylase 4A
KDM4B	lysine demethylase 4B
KDM4C	lysine demethylase 4C
KDM4D	lysine demethylase 4D
KDM4E	lysine demethylase 4E
KDM5A	lysine demethylase 5A
KDM5B	lysine demethylase 5B
KDM5C	lysine demethylase 5C
KDM5D	lysine demethylase 5D
KDM6A	lysine demethylase 6A
KDM6B	lysine demethylase 6B
KDM7A	lysine demethylase 7A
KDM8	lysine demethylase 8

OGFOD1	2-oxoglutarate and iron dependent oxygenase domain containing 1
P3H1	prolyl 3-hydroxylase 1
P3H2	prolyl 3-hydroxylase 2
P3H3	prolyl 3-hydroxylase 3
P4HA1	prolyl 4-hydroxylase subunit alpha 1
P4HA2	prolyl 4-hydroxylase subunit alpha 2
P4HA3	prolyl 4-hydroxylase subunit alpha 3
P4HTM	prolyl 4-hydroxylase, transmembrane
PHF2	PHD finger protein 2
PHF8	PHD finger protein 8
PHYH	phytanoyl-CoA 2-hydroxylase
PHYHD1	phytanoyl-CoA dioxygenase domain containing 1
PLOD1	procollagen-lysine,2-oxoglutarate 5-dioxygenase 1
PLOD2	procollagen-lysine,2-oxoglutarate 5-dioxygenase 2
PLOD3	procollagen-lysine,2-oxoglutarate 5-dioxygenase 3
TET1	tet methylcytosine dioxygenase 1
TET2	tet methylcytosine dioxygenase 2
TET3	tet methylcytosine dioxygenase 3
TMLHE	trimethyllysine hydroxylase, epsilon
UTY	ubiquitously transcribed tetratricopeptide repeat containing, Y-linked

Table S3. Univariate and multivariate Cox proportional hazards analysis of risk factors associated with overall survival in multiple cancers.

	Hazard Ratio (95% CI)	*P*-value
Bladder	**Univariate**	
Signature 1 (high vs. low score)	0.662 (0.450 - 0.974)	**0.036**
Signature 2 (high vs. low score)	1.459 (1.096 - 2.137)	**0.042**
TNM staging (II & III vs. I)	1.679 (1.323 - 2.131)	**2.02E-05**
PCDHA1 mutation (Yes vs. No)	1.649 (1.058 - 2.569)	**0.027**
TTN mutation (Yes vs. No)	1.610 (1.091 - 2.376)	**0.016**
	Multivariate (Signature 1)	
Signature 1 (high vs. low score)	0.714 (0.502 - 1.094)	0.132
TNM staging (II & III vs. I)	1.636 (1.286 - 2.081)	**6.03E-05**
	Multivariate (Signature 1)	
Signature 1 (high vs. low score)	0.686 (0.466 - 0.912)	**0.047**
PCDHA1 mutation (Yes vs. No)	1.573 (1.008 - 2.458)	**0.046**
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	1.481 (1.011 - 2.168)	**0.043**
TNM staging (II & III vs. I)	1.685 (1.328 - 2.137)	**1.76E-05**
	Multivariate (Signature 2)	
Signature 2 (low vs. high score)	1.411 (1.062 - 2.070)	**0.048**
PCDHA1 mutation (Yes vs. No)	1.584 (1.015 - 2.472)	**0.043**
Cervical	**Univariate**	
Signature 2 (high vs. low score)	1.972 (1.003 - 3.877)	**0.045**
TNM staging (II & III vs. I)	1.017 (0.695 - 1.487)	0.93
PCDHA1 mutation (Yes vs. No)	1.364 (0.681 - 2.729)	0.38
PIK3CA mutation (Yes vs. No)	1.906 (0.955 - 3.804)	0.067
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	2.01 (1.011 - 3.996)	**0.046**
TNM staging (II & III vs. I)	0.942 (0.639 - 1.390)	0.76
Head and neck	**Univariate**	
Signature 2 (high vs. low score)	1.479 (1.056 - 2.072)	**0.0228**
TNM staging (II & III vs. I)	1.466 (1.188 - 1.808)	**3.53E-04**
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	1.344 (0.956 - 1.890)	0.089076
TNM staging (II & III vs. I)	1.426 (1.155 - 1.761)	**9.74E-04**
Renal clear cell	**Univariate**	
Signature 2 (high vs. low score)	1.483 (1.096 - 2.007)	**0.011**
TNM staging (II & III vs. I)	1.87 (1.641 - 2.132)	**2.00E-16**
MUC4 mutation (Yes vs. No)	0.570 (0.370 - 0.880)	**0.012**
PBRM1 mutation (Yes vs. No)	0.771 (0.508 - 1.172)	0.22
VHL mutation (Yes vs. No)	1.056 (0.780 - 1.430)	0.72
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	1.197 (0.881 - 1.627)	0.25
TNM staging (II & III vs. I)	1.847 (1.618 - 2.109)	**2.00E-16**
	Multivariate (Signature 2)	

Signature 2 (high vs. low score)	1.520 (1.123 - 2.056)	**6.72E-03**
MUC4 mutation (Yes vs. No)	0.554 (0.359 - 0.854)	**7.59E-03**

Renal papillary cell	**Univariate**	
Signature 1 (high vs. low score)	0.370 (0.157 - 0.871)	**0.023**
Signature 2 (high vs. low score)	3.862 (1.565 - 9.526)	**3.36E-03**
TNM staging (II & III vs. I)	2.71 (1.893 - 3.878)	**5.03E-08**
	Multivariate (Signature 1)	
Signature 1 (high vs. low score)	0.461 (0.194 - 0.981)	**0.048**
TNM staging (II & III vs. I)	2.621 (1.821 - 3.772)	**2.13E-07**
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	3.057 (1.232 - 7.584)	**0.016**
TNM staging (II & III vs. I)	2.616 (1.805 - 3.790)	**3.78E-07**

Liver #1 (LIHC)	**Univariate**	
Signature 1 (high vs. low score)	0.656 (0.424 - 0.915)	**0.048**
TNM staging (II & III vs. I)	2.000 (1.533 - 2.608)	**3.14E-07**
	Multivariate (Signature 1)	
Signature 1 (high vs. low score)	0.717 (0.464 - 1.108)	0.13
TNM staging (II & III vs. I)	1.999 (1.570 - 2.546)	**1.99E-08**

Liver #2 (LIRI-JP)	**Univariate**	
Signature 1 (high vs. low score)	0.490 (0.259 - 0.938)	**0.031**
Signature 2 (high vs. low score)	5.271 (2.429 - 11.44)	**2.61E-05**
Tumour size (> 3cm vs. < 3cm)	2.415 (1.263 - 4.615)	**7.60E-03**
TNM staging (II & III vs. I)	2.267 (1.556 - 3.304)	**2.03E-05**
	Multivariate (Signature 1)	
Signature 1 (high vs. low score)	0.541 (0.283 - 0.904)	**0.043**
Tumour size (> 3cm vs. < 3cm)	1.003 (0.997 - 1.009)	0.38
TNM staging (II & III vs. I)	2.065 (1.374 - 3.104)	**0.000483**
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	4.539 (2.055 - 10.029)	**1.84E-04**
Tumour size (> 3cm vs. < 3cm)	1.001 (0.994 - 1.008)	0.78
TNM staging (II & III vs. I)	1.926 (1.296 - 2.862)	**1.19E-03**

Liver #3 (GSE14520)	**Univariate**	
Signature 2 (high vs. low score)	2.285 (1.458 - 3.580)	**3.12E-04**
Tumour size (> 5cm vs. < 5cm)	2.159 (1.406 - 3.316)	**4.40E-04**
Cirrhosis (Yes vs. No)	4.665 (1.147 - 18.97)	**0.031**
TNM staging (II & III vs. I)	2.26 (1.706 - 2.994)	**1.30E-08**
BCLC staging (B & C vs. A & 0)	2.181 (1.722 - 2.762)	**9.86E-11**
AFP (> 300ng/mL vs. < 300ng/mL)	1.606 (1.049 - 2.46)	**0.029**
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	2.012 (1.267 - 3.195)	**3.04E-03**
Tumour size (> 5cm vs. < 5cm)	0.983 (0.589 - 1.642)	0.94
Cirrhosis (Yes vs. No)	3.187 (0.762 - 13.335)	0.11
TNM staging (II & III vs. I)	1.346 (0.914 - 1.980)	0.13
BCLC staging (B & C vs. A & 0)	1.873 (1.332 - 2.634)	**3.08E-04**

Lung	**Univariate**

Signature 1 (high vs. low score)	0.6245 (0.443 - 0.879)	**6.97E-03**
Signature 2 (high vs. low score)	1.562 (1.116 - 2.188)	**9.44E-03**
TNM staging (II & III vs. I)	1.597 (1.364 - 1.870)	**6.11E-09**
	Multivariate (Signature 1)	
Signature 1 (high vs. low score)	0.715 (0.505 - 0.913)	**0.039**
TNM staging (II & III vs. I)	1.553 (1.324 - 1.822)	**6.74E-08**
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	1.496 (1.067 - 2.097)	**0.019**
TNM staging (II & III vs. I)	1.587 (1.353 - 1.861)	**1.36E-08**

Pancreas	**Univariate**	
Signature 1 (high vs. low score)	0.454 (0.278 - 0.741)	**1.59E-03**
Signature 2 (high vs. low score)	1.969 (1.217 - 3.186)	**5.81E-03**
TNM staging (II & III vs. I)	1.339 (0.897 - 1.998)	0.153
	Multivariate (Signature 1)	
Signature 1 (high vs. low score)	0.472 (0.284 - 0.783)	**3.62E-03**
TNM staging (II & III vs. I)	1.150 (0.722 - 1.832)	0.56
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	1.888 (1.157 - 3.082)	**0.011**
TNM staging (II & III vs. I)	1.205 (0.785 - 1.851)	0.394

Stomach	**Univariate**	
Signature 2 (high vs. low score)	1.725 (1.142 - 2.605)	**9.50E-03**
TNM staging (II & III vs. I)	1.372 (1.067 - 1.765)	**0.013**
PCDHA1 mutation (Yes vs. No)	1.525 (1.007 - 2.307)	**0.046**
PCDHA2 mutation (Yes vs. No)	1.604 (1.061 - 2.427)	**0.025**
TP53 mutation (Yes vs. No)	1.296 (0.853 - 1.967)	0.22
TTN mutation (Yes vs. No)	1.115 (0.740 - 1.682)	0.60
	Multivariate (Signature 2)	
Signature 2 (high vs. low score)	1.800 (1.184 - 2.737)	**5.90E-03**
TNM staging (II & III vs. I)	1.464 (1.130 - 1.898)	**3.90E-03**
PCDHA1 mutation (Yes vs. No)	1.501 (0.988 - 2.281)	0.056
PCDHA2 mutation (Yes vs. No)	1.558 (1.026 - 2.365)	**0.037**

Endometrial	**Univariate**	
Signature 1 (high vs. low score)	0.401 (0.229 - 0.702)	**1.39E-03**
TNM staging (II & III vs. I)	1.802 (1.431 - 2.269)	**5.69E-07**
PCDHA1 mutation (Yes vs. No)	0.516 (0.272 - 0.978)	**0.042**
PIK3CA mutation (Yes vs. No)	0.362 (0.190 - 0.689)	**1.98E-03**
PIK3R1 mutation (Yes vs. No)	1.180 (0.645 - 2.158)	0.592
PTEN mutation (Yes vs. No)	0.427 (0.234 - 0.781)	**5.75E-03**
TP53 mutation (Yes vs. No)	1.780 (1.025 - 3.090)	**0.041**
	Multivariate (Signature 1)	
Signature 1 (high vs. low score)	0.519 (0.293 - 0.920)	**0.024**
TNM staging (II & III vs. I)	1.589 (1.245 - 2.028)	**2.02E-04**
PCDHA1 mutation (Yes vs. No)	0.803 (0.376 - 1.712)	0.56
PIK3CA mutation (Yes vs. No)	0.428 (0.216 - 0.848)	**0.015**
PTEN mutation (Yes vs. No)	0.904 (0.427 - 1.910)	0.79
TP53 mutation (Yes vs. No)	1.567 (0.867 - 2.832)	0.13

Table S4. Differentially expressed genes between KDM8-high and -low groups in hepatocellular carcinoma (LIHC-cohort)

Entrez.id	GeneSymbol	Log2 Fold Change	P value
1548	CYP2A6	4.266217426	0.001070647
1409	CRYAA	3.782386136	0.000184121
1576	CYP3A4	3.698152017	0.004124437
1582	CYP8B1	3.644639707	0.000565742
255167	LINC01018	3.620703779	0.001817311
143941	TTC36	3.602776338	0.000293049
51179	HAO2	3.554412019	0.000892977
64388	GREM2	3.515557145	0.001483312
1549	CYP2A7	3.471653994	0.009032202
388503	C3P1	3.446255927	0.000123142
2642	GCGR	3.442241631	0.004731164
6718	AKR1D1	3.309733658	0.000704509
2645	GCK	3.286380385	0.003725235
173	AFM	3.274019175	0.000487471
28234	SLCO1B3	3.200697658	0.013857647
6554	SLC10A1	3.08409662	0.002198269
2674	GFRA1	3.066645866	0.000515341
252995	FNDC5	3.049912992	0.000542378
33	ACADL	3.043110654	0.000960604
10877	CFHR4	3.023986393	0.000576045
84879	MFSD2A	3.003142647	0.000328791
635	BHMT	2.973505474	0.001225789
127	ADH4	2.925047346	0.006456936
1581	CYP7A1	2.918451657	0.007607397
10249	GLYAT	2.89006854	0.005452309
7069	THRSP	2.87825033	0.011128485
10864	SLC22A7	2.8294376	0.002915624
6694	SPP2	2.799569821	0.004013825
266629	SEC14L3	2.785909232	2.99838E-05
6898	TAT	2.764924205	0.002283852
1610	DAO	2.700033063	0.001676755
8647	ABCB11	2.68853326	0.011722721
5244	ABCB4	2.687798988	0.000106696
10	NAT2	2.679791071	0.001921913
83597	RTP3	2.672686193	0.001081366
189	AGXT	2.600530207	0.000591273
4311	MME	2.597010495	0.028115017
171389	NLRP6	2.576924698	1.73365E-05
1600	DAB1	2.558491234	6.20059E-05
58503	OPRPN	2.531857402	0.004723015
284541	CYP4A22	2.526945213	0.002622958
3626	INHBC	2.514424591	0.000155015
116519	APOA5	2.512852835	0.0008519
2158	F9	2.505159911	0.011036784
200931	SLC51A	2.497961255	0.003342533
2998	GYS2	2.494680586	0.011620599
202299	N/A	2.488584103	0.018415021
92292	GLYATL1	2.484104891	3.09117E-05
51302	CYP39A1	2.45250273	0.005432105
29947	DNMT3L	2.447810506	0.00145115
7038	TG	2.429091048	0.001987957
150280	HORMAD2	2.42859805	0.000156247

51733	UPB1	2.42753975	0.002804778
462	SERPINC1	2.423306661	0.007685312
26998	FETUB	2.423205357	0.006804976
345	APOC3	2.414959144	0.001946819
11199	ANXA10	2.41063564	0.000373832
2892	GRIA3	2.37643902	0.006892948
123876	ACSM2A	2.375149513	0.001677471
8529	CYP4F2	2.356139541	0.003590649
1373	CPS1	2.347456038	0.031293508
1564	N/A	2.344177514	0.001167081
6580	SLC22A1	2.327872835	0.027051449
729	C6	2.324475916	2.06597E-06
1558	CYP2C8	2.317696049	0.001559205
5444	PON1	2.316342538	0.001985561
1579	CYP4A11	2.315220723	0.000471545
157310	PEBP4	2.296432128	0.003869673
57733	GBA3	2.291043867	0.003399187
64902	AGXT2	2.279171841	0.006002827
816	CAMK2B	2.274354004	0.014691834
29943	PADI1	2.265481392	0.018050291
23086	EXPH5	2.258753783	0.00035824
7545	ZIC1	2.242916061	0.009744838
8630	HSD17B6	2.234080306	0.004389488
10878	CFHR3	2.229794129	0.017950652
23316	CUX2	2.219746125	0.018957676
348158	ACSM2B	2.216326931	0.001903964
57451	TENM2	2.215244511	0.01927035
9077	DIRAS3	2.202948942	0.000237404
81494	CFHR5	2.197465122	0.040225546
8856	NR1I2	2.179939607	0.016093996
229	ALDOB	2.179066312	0.010830097
1557	CYP2C19	2.179047891	0.000446408
27232	GNMT	2.178235399	0.012158931
22977	AKR7A3	2.175902379	0.009147347
731	C8A	2.169356533	0.001001301
6514	SLC2A2	2.162013	0.013804129
3950	LECT2	2.14055188	0.046697952
63982	ANO3	2.136028037	0.001540142
1553	CYP2A13	2.132585759	0.016606607
3053	SERPIND1	2.113853805	0.008433887
83854	ANGPTL6	2.113590149	0.000128373
121643	FOXN4	2.099987483	0.019830535
3250	HPR	2.099986491	0.014752115
124	ADH1A	2.094269712	0.004377302
366	AQP9	2.092012283	0.009955353
3483	IGFALS	2.083719247	0.014959325
145264	SERPINA12	2.079006118	0.001337341
83729	INHBE	2.07372899	0.001235156
2775	GNAO1	2.073549717	0.001254919
404037	HAPLN4	2.072589055	0.000548267
8608	RDH16	2.072137893	0.01106105
2822	GPLD1	2.071295157	0.013408377
1565	CYP2D6	2.069534836	0.009360876
6360	CCL16	2.068833248	0.017718212
10747	MASP2	2.067838126	0.016925366
5789	PTPRD	2.065981194	0.015065062
116842	LEAP2	2.062777241	0.002171084

653808	ZG16	2.062233472	0.020038666
10599	SLCO1B1	2.052246346	0.008510799
5207	PFKFB1	2.047236592	0.003951254
26291	FGF21	2.043516335	0.033162519
53345	TM6SF2	2.037578069	0.001128643
80168	MOGAT2	2.031577131	0.011018668
2099	ESR1	2.03155751	0.002132853
56134	PCDHAC2	2.030624527	0.007205624
283537	SLC46A3	2.024509926	0.000149799
9970	NR1I3	2.021507726	0.008723175
2161	F12	2.021505846	0.008103316
280	AMY2B	2.020923782	1.96624E-06
148741	ANKRD35	2.018621863	0.007869293
429	ASCL1	2.013026287	0.014483865
346	APOC4	2.012881975	0.030557686
360200	TMPRSS9	2.010677453	0.011300575
134111	UBE2QL1	2.007580313	0.006600179
1828	DSG1	2.003596641	0.025063577
2165	F13B	2.000652263	0.009233682
3827	KNG1	1.997414991	0.002147747
64850	ETNPPL	1.992437952	0.028921785
23541	SEC14L2	1.988841343	0.000171278
387601	SLC22A25	1.988373871	0.002571215
392636	AGMO	1.982612445	0.007127166
1555	CYP2B6	1.974663087	0.004152261
2690	GHR	1.97120996	3.49149E-05
388646	GBP7	1.969293624	0.04529398
5340	PLG	1.968029917	0.006855863
53841	CDHR5	1.965182625	0.012015737
221223	CES5A	1.963985876	0.002671821
6291	SAA4	1.963317102	0.016743223
151126	ZNF385B	1.959984183	0.008367854
126	ADH1C	1.955094511	0.045514905
151295	SLC23A3	1.952474699	0.007351873
2938	GSTA1	1.94869482	0.043904746
114571	SLC22A9	1.939303807	0.00442756
3263	HPX	1.938423562	0.008971805
335	APOA1	1.937796701	0.02032655
4018	LPA	1.935111258	0.004113881
23584	VSIG2	1.930836804	0.030319734
3818	KLKB1	1.930832589	4.20206E-06
159963	SLC5A12	1.926473098	0.006415145
64816	CYP3A43	1.918812267	0.00162869
114770	PGLYRP2	1.918751789	0.024852669
26	AOC1	1.918632919	0.010663639
51268	PIPOX	1.91608333	0.002293964
339398	LINGO4	1.914251522	0.002619155
389643	NUGGC	1.910340137	0.011092759
9388	LIPG	1.909470603	0.000104456
122664	TPPP2	1.908114969	0.001873702
3240	HP	1.906797774	0.022811484
54988	ACSM5	1.903249856	0.015671305
6716	SRD5A2	1.902798702	0.012995147
7399	USH2A	1.901861001	0.006155153
6822	SULT2A1	1.896825407	0.027345031
389434	IYD	1.889670162	0.0012295
2160	F11	1.886585889	0.001108506

10998	SLC27A5	1.886493737	0.004258579
3080	CFHR2	1.872150388	0.016205216
387778	SPDYC	1.870768999	0.001686804
4153	MBL2	1.865182539	0.020969296
6751	SSTR1	1.853287797	0.0106445
256394	SERPINA11	1.853130059	0.027811459
164656	TMPRSS6	1.848365981	0.004707227
55908	ANGPTL8	1.843952506	0.02120754
1	A1BG	1.841239211	0.019645158
590	BCHE	1.838711089	0.047180947
201651	AADACP1	1.829664266	0.002100734
7276	TTR	1.826134518	0.022078856
55224	ETNK2	1.819980809	0.005221762
554235	ASPDH	1.814205909	0.03195445
1036	CDO1	1.812630613	0.005704286
654790	PCP4L1	1.812336409	0.015486964
1109	AKR1C4	1.798881786	0.024654002
7365	UGT2B10	1.79754988	0.006054442
79831	KDM8	1.797348411	5.14958E-06
83875	BCO2	1.79068093	0.002815352
1592	CYP26A1	1.790051825	0.008741061
8858	PROZ	1.789316028	0.012357988
2894	GRID1	1.784642645	0.038421082
6529	SLC6A1	1.777487014	0.002787926
135138	PACRG	1.776624146	0.016016102
85480	TSLP	1.772797812	0.008926007
346606	MOGAT3	1.77260438	0.010890367
6540	SLC6A13	1.772129375	0.006988695
6539	SLC6A12	1.760580601	0.004598944
50486	G0S2	1.759248375	0.002818565
23491	CES3	1.758411582	0.0052068
91947	ARRDC4	1.754857186	0.000627529
10586	MAB21L2	1.75094477	0.027107321
400830	DEFB132	1.746677951	0.019190447
276	AMY1A	1.745153531	0.000151813
57549	IGSF9	1.740577196	0.04783784
246778	IL27	1.738596308	0.002396079
51458	RHCG	1.732390615	0.046747665
552	AVPR1A	1.729890419	0.04451327
84107	ZIC4	1.72971236	0.02836071
145837	N/A	1.728976489	0.047849345
148738	HFE2	1.722431132	0.048661379
7306	TYRP1	1.713379317	0.0121349
2041	EPHA1	1.698195127	0.02815375
1559	CYP2C9	1.692305393	0.017770237
55753	OGDHL	1.687834165	0.014810533
1238	ACKR2	1.68674921	0.001267128
1101	CHAD	1.681941072	0.000812831
1962	EHHADH	1.679240877	0.001395897
93145	OLFM2	1.678413754	0.000480129
22843	PPM1E	1.676015108	0.030930921
1370	CPN2	1.669817786	0.007900849
3948	LDHC	1.669309207	0.017403664
6579	SLCO1A2	1.668373987	0.038648928
63951	DMRTA1	1.667196203	0.016829527
84866	TMEM25	1.663207975	0.00319644
81029	WNT5B	1.661039019	0.010439984

90523	MLIP	1.659916179	0.012463271
732	C8B	1.652146764	0.006573558
57512	GPR158	1.642390106	0.040030886
23632	CA14	1.641714687	0.018466979
284904	SEC14L4	1.640343954	0.041929967
80069	LINC00574	1.637356735	0.00020239
91703	ACY3	1.634570238	0.006099373
10991	SLC38A3	1.632561777	0.016964618
2328	FMO3	1.624156072	0.007172622
283	ANG	1.61484952	0.000565227
123872	DNAAF1	1.610650547	0.02206757
3700	ITIH4	1.610072628	0.003507918
57529	RTL9	1.608977194	0.014122214
367	AR	1.606428173	0.011577899
389668	XKR9	1.605973821	0.022433951
2538	G6PC	1.603549774	0.023203687
2180	ACSL1	1.602940285	0.001107126
125	ADH1B	1.602850062	0.039904227
54435	N/A	1.598461408	0.028224472
64240	ABCG5	1.587709404	0.014126658
1562	CYP2C18	1.584036942	0.018942124
202151	RANBP3L	1.582539278	0.032435377
23460	ABCA6	1.581982582	0.011053131
26084	ARHGEF26	1.577161349	0.008229287
80059	LRRTM4	1.575267801	0.004757661
7036	TFR2	1.5722435	0.01994296
563	AZGP1	1.570740021	0.006373881
10350	ABCA9	1.569171764	0.000594989
5104	SERPINA5	1.56764613	0.013542499
54363	HAO1	1.561769397	0.028920354
84649	DGAT2	1.561613212	0.00683867
648740	N/A	1.560120967	0.000465744
8110	DPF3	1.558182368	0.0016597
344	APOC2	1.5538243	0.045210088
80303	EFHD1	1.551792486	0.004610191
26232	FBXO2	1.545764839	0.014934684
161247	FITM1	1.543109731	0.00344818
8991	SELENBP1	1.540862076	0.001509125
145645	TERB2	1.539483171	0.005321637
9832	JAKMIP2	1.536765947	0.005393749
26577	PCOLCE2	1.533702026	0.048591081
7018	TF	1.531483027	0.021964166
10841	FTCD	1.522763286	0.029823669
64173	N/A	1.521882145	6.26393E-05
64577	ALDH8A1	1.520105598	0.020042599
7433	VIPR1	1.517951622	0.009486735
283848	CES4A	1.505301997	0.020333399
134526	ACOT12	1.501796858	0.04129444
83758	RBP5	1.499639144	0.008899611
9848	MFAP3L	1.497362893	0.003120189
4143	MAT1A	1.496540752	0.021435203
11001	SLC27A2	1.487431618	0.025563174
3931	LCAT	1.486153855	0.009723114
875	CBS	1.484246243	0.009582261
7498	XDH	1.483324178	0.024182197
339965	CCDC158	1.479959669	0.001183926
3479	IGF1	1.475203183	0.013132302

5648	MASP1	1.473335053	0.003850723
1621	DBH	1.468504375	0.025107006
2564	GABRE	1.468117371	0.023425827
213	ALB	1.46507454	0.037002162
6581	SLC22A3	1.463954569	0.016113803
259	AMBP	1.45064681	0.017444397
1842	ECM2	1.448472726	0.000147365
388815	N/A	1.446506599	0.004335874
84699	CREB3L3	1.445202911	0.024012291
5106	PCK2	1.44360587	0.001587637
399959	MIR100HG	1.43676361	0.005262615
3699	ITIH3	1.433683239	0.015408083
1361	CPB2	1.432397218	0.0163544
50509	COL5A3	1.432109449	0.001899974
27284	SULT1B1	1.430225484	0.018618149
8431	NR0B2	1.429688102	0.029235533
27443	CECR2	1.427703983	0.024054631
29119	CTNNA3	1.426648419	0.01500674
5005	ORM2	1.425526543	0.027492056
6401	SELE	1.424234398	0.016895563
6470	SHMT1	1.421016227	0.0009126
4501	MT1X	1.420866767	0.040183179
4051	CYP4F3	1.420775249	0.015894113
26238	LINC01558	1.41682479	0.005682531
5004	ORM1	1.415826815	0.02806463
10060	ABCC9	1.415596396	0.003366254
6335	SCN9A	1.413961978	0.013336958
7274	TTPA	1.413835647	0.028802054
5046	PCSK6	1.409326956	0.003610406
90586	AOC4P	1.407813954	0.009644706
338821	SLCO1B7	1.405237364	0.019328311
5950	RBP4	1.402345482	0.040406923
10349	ABCA10	1.400760353	0.009108983
22809	ATF5	1.394730619	0.036585322
10157	AASS	1.389967645	0.011910273
79054	TRPM8	1.386587936	0.041276147
56624	ASAH2	1.383674374	0.006694146
79974	CPED1	1.382958699	0.009010569
245	ALOX12P2	1.376601275	0.012907206
84630	TTBK1	1.367764765	0.004811105
6097	RORC	1.366878963	0.00145782
8022	LHX3	1.366438412	0.047556874
55107	ANO1	1.360123078	0.03733801
29958	DMGDH	1.351341837	0.020766218
3697	ITIH1	1.351313833	0.029753442
4610	MYCL	1.350329714	0.002795463
84253	GARNL3	1.350145275	0.00658309
89822	KCNK17	1.3485726	0.028163104
27329	ANGPTL3	1.345721441	0.036205928
388182	N/A	1.344102608	0.021757967
84460	ZMAT1	1.340505762	0.000361876
2147	F2	1.335866243	0.025530006
84873	ADGRG7	1.331848934	0.02768844
2277	VEGFD	1.325837861	0.048006646
6098	ROS1	1.320258609	0.031317343
2203	FBP1	1.319634995	0.037390222
51181	DCXR	1.319467687	0.008432687

83715	ESPN	1.313818872	0.008096425
152195	NUDT16P1	1.313575016	0.033129674
199920	FYB2	1.30678752	0.023628838
494470	RNF165	1.306142836	0.017249724
64072	CDH23	1.297260629	0.006524927
10351	ABCA8	1.293599393	0.043942392
57569	ARHGAP20	1.293154723	0.013124018
10144	FAM13A	1.29188075	0.006341895
432	ASGR1	1.291283747	0.019956348
26053	AUTS2	1.288010803	0.008313303
285154	CYP1B1-AS1	1.286684001	0.007298905
79632	FAM184A	1.280735183	0.036457415
2329	FMO4	1.27786324	0.000127037
5176	SERPINF1	1.275011915	0.015883909
1812	DRD1	1.274528749	0.049149818
57678	GPAM	1.271366558	0.015807361
7216	TRO	1.270414852	0.004374857
219670	ENKUR	1.252328995	0.001468013
349149	GJC3	1.247486863	0.007512978
338323	NLRP14	1.245980146	0.007596123
260293	CYP4X1	1.245825942	0.014414288
3698	ITIH2	1.24305569	0.023671309
246181	AKR7L	1.239480534	0.008944274
54923	LIME1	1.239480267	0.007733856
51380	CSAD	1.239471506	0.000374049
2819	GPD1	1.231605305	0.035276503
340811	AKR1C8P	1.229819596	0.037736074
646023	ADORA2A-AS1	1.229310558	0.029797605
5446	PON3	1.227744138	0.028043688
8824	CES2	1.223262489	0.002640713
7448	VTN	1.221142167	0.009252379
84983	FAM222A-AS1	1.22057447	0.00196858
27141	CIDEB	1.220094972	0.00076436
377841	ENTPD8	1.220017899	0.044330609
91614	DEPDC7	1.217544996	0.025594364
23650	TRIM29	1.213453057	0.039024427
6334	SCN8A	1.21340328	0.018486509
348249	N/A	1.209494445	0.016954648
170392	OIT3	1.209213391	0.012045028
55244	SLC47A1	1.209086881	0.024281888
50940	PDE11A	1.204782096	0.030866602
100133205	LINC00240	1.203547604	0.002058614
154	ADRB2	1.202106431	0.025748268
8309	ACOX2	1.200327401	0.035491968
8512	MBL1P	1.199904359	0.0197174
54873	PALMD	1.198587085	0.002881576
51085	MLXIPL	1.195836831	0.000152905
761	CA3	1.194667869	0.015980601
10050	SLC17A4	1.191172477	0.033406065
23562	CLDN14	1.187857923	0.0206028
150356	CHADL	1.187834368	0.034986706
147	ADRA1B	1.186540017	0.007194874
440503	PLIN5	1.185077159	0.002948669
8777	MPDZ	1.183643951	0.001780448
1943	EFNA2	1.177989949	0.004629949
130	ADH6	1.17584133	0.039125191
320	APBA1	1.169781308	0.023853528

154661	RUNDC3B	1.169078724	0.010243899
5024	P2RX3	1.167963719	0.009063695
153769	SH3RF2	1.167762114	0.034071017
7098	TLR3	1.166028552	0.012056776
3795	KHK	1.162429632	0.016468594
2028	ENPEP	1.160645083	0.00165195
127254	ERICH3	1.158219563	0.017475746
51156	SERPINA10	1.156496846	0.035771487
6038	RNASE4	1.150719571	0.000743368
79776	ZFHX4	1.150602446	0.006145916
81543	LRRC3	1.149440558	0.013476296
140699	MROH8	1.149300375	0.000202225
57447	NDRG2	1.149264634	3.18725E-05
285800	PRR18	1.148433442	0.02244634
140828	LINC00261	1.146943493	0.012495279
2780	GNAT2	1.143833603	0.00400733
116844	LRG1	1.143076026	0.025846685
6006	RHCE	1.14172383	0.042967882
89886	SLAMF9	1.141445194	0.042450137
286097	MICU3	1.139195435	0.011572289
441476	STPG3	1.137809695	0.018937208
55277	FGGY	1.135260604	0.038376635
285954	INHBA-AS1	1.135077363	0.011504418
414152	C10ORF105	1.133765997	0.011086218
57026	PDXP	1.133587715	7.47276E-05
18	ABAT	1.132396717	0.016954011
54627	MAP10	1.130319302	0.023493353
23500	DAAM2	1.130304113	0.011901214
5651	TMPRSS15	1.122836763	0.049493534
100188954	N/A	1.119940555	0.030480168
147525	N/A	1.115742152	0.005162458
9104	RGN	1.11441316	0.037408486
54084	TSPEAR	1.11200885	0.04335861
2155	F7	1.109456476	0.008285169
400752	N/A	1.106026795	0.000268752
7704	ZBTB16	1.104977008	0.009865461
653190	N/A	1.101097272	0.007585401
1647	GADD45A	1.100026962	2.00943E-05
5959	RDH5	1.098574216	0.007042391
401288	LINC00242	1.098269741	0.001299677
100129387	N/A	1.096172325	0.000282967
2052	EPHX1	1.093699058	0.015414286
57134	MAN1C1	1.088347318	0.031088089
89927	C16ORF45	1.08559599	0.00544766
54757	FAM20A	1.082322076	0.016216782
254427	PROSER2	1.081743733	0.007964361
7425	VGF	1.08153181	0.008800112
6652	SORD	1.08038233	0.020779186
342931	RFPL4A	1.080173565	0.040723161
377677	CA13	1.079310178	0.009919664
115817	DHRS1	1.074820067	0.001006992
122622	ADSSL1	1.074129809	0.009184543
10402	ST3GAL6	1.072964688	0.00712378
6095	RORA	1.071228748	7.62058E-05
4843	NOS2	1.066715913	0.041703698
3638	INSIG1	1.064815107	0.038862707
32	ACACB	1.059698185	0.001686509

3075	CFH	1.055741774	0.030385611
1593	CYP27A1	1.052799928	0.029272041
84217	ZMYND12	1.051651416	0.03600959
285386	TPRG1	1.048592099	0.031583257
79660	PPP1R3B	1.048426658	0.004569388
391059	FRRS1	1.047308419	0.035947478
978	CDA	1.042511538	0.047707531
718	C3	1.042156504	0.017853286
8542	APOL1	1.039415681	0.002420486
84952	CGNL1	1.03856374	0.020445663
79614	N/A	1.037537142	0.041356081
79814	AGMAT	1.03696062	0.048040584
131	ADH7	1.035696434	0.038098856
64284	RAB17	1.034620697	0.02394295
10826	FAXDC2	1.034324678	0.002391724
23600	AMACR	1.033956168	0.032624265
149465	CFAP57	1.033364636	0.008166289
279	AMY2A	1.033321944	0.020321949
1757	SARDH	1.032628652	0.015510262
3226	HOXC10	1.031904017	0.037868401
5380	N/A	1.028439705	0.000293284
4522	MTHFD1	1.023805714	0.002358802
3577	CXCR1	1.023780695	0.025209753
220004	PPP1R32	1.022166053	0.008785035
285704	RGMB	1.021856316	0.00444471
399671	HEATR4	1.018090387	0.019794184
116	ADCYAP1	1.016673106	0.046225573
116255	MOGAT1	1.014613791	0.009106562
6567	SLC16A2	1.013941147	0.005829829
6342	SCP2	1.012230142	0.000588239
28959	TMEM176B	1.010920124	0.019629853
284723	SLC25A34	1.009042952	0.02716002
368	ABCC6	1.005286833	0.008909763
5919	RARRES2	1.004067478	0.011290875
79729	SH3D21	-1.00112383	0.004751775
991	CDC20	-1.002760237	0.025530959
57468	SLC12A5	-1.0045625	0.008673047
84688	C9ORF24	-1.005609921	0.009267502
770	CA11	-1.006909647	0.026512726
9454	HOMER3	-1.007842144	0.010329995
80763	SPX	-1.00997037	0.047386534
54704	PDP1	-1.010436285	0.007250862
400629	TEX19	-1.011514889	0.045583694
644	BLVRA	-1.014072197	0.004980499
57655	GRAMD1A	-1.020483391	0.000855917
9956	HS3ST2	-1.021026784	0.025808302
1237	CCR8	-1.025273843	0.007499397
636	BICD1	-1.025858787	0.014533123
287	ANK2	-1.028256091	0.029530984
4130	MAP1A	-1.028837764	0.044791204
4619	MYH1	-1.032697225	0.04573196
4828	NMB	-1.033638898	0.009623243
4118	MAL	-1.03400016	0.030564307
9805	SCRN1	-1.034811618	0.034065097
84662	GLIS2	-1.037043451	0.022253202
85479	DNAJC5B	-1.040909474	0.047583649
9966	TNFSF15	-1.043460505	0.032033579

6887	TAL2	-1.043485639	0.020746944
777	CACNA1E	-1.046013803	0.046354211
941	CD80	-1.047868233	0.026317087
9170	LPAR2	-1.048460085	0.028810306
343263	MYBPHL	-1.04887826	0.026680468
84900	RNFT2	-1.049507583	0.011840666
55138	FAM90A1	-1.052241448	0.016488239
27306	HPGDS	-1.053381493	0.018726789
57562	CEP126	-1.058066118	0.032615304
4692	NDN	-1.059132713	0.031835214
26999	CYFIP2	-1.059893426	0.033356281
79801	SHCBP1	-1.063529208	0.030686572
9654	TTLL4	-1.064128284	0.008864083
26996	GPR160	-1.06440735	0.005775609
401884	N/A	-1.064814446	0.022253886
79888	LPCAT1	-1.077109474	0.002447006
92558	BICDL1	-1.084870164	0.030350761
11322	TMC6	-1.085514659	0.005820871
91373	UAP1L1	-1.085598696	0.025285397
26150	RIBC2	-1.093757714	0.040076184
4050	LTB	-1.094388431	0.039634298
342184	FMN1	-1.094884045	0.012581329
6875	TAF4B	-1.102864521	0.015079077
50615	IL21R	-1.103329043	0.043109335
84842	HPDL	-1.10388781	0.039626806
7040	TGFB1	-1.10625981	0.006099409
1235	CCR6	-1.107584342	0.024493721
84941	HSH2D	-1.111837698	0.029570408
9022	CLIC3	-1.118537954	0.045822506
23302	WSCD1	-1.119734191	0.028783814
10518	CIB2	-1.122919185	0.031744582
84935	MEDAG	-1.124468832	0.039571866
23363	OBSL1	-1.127672929	0.032755606
84959	UBASH3B	-1.128549082	0.004436883
9143	SYNGR3	-1.132121288	0.046934128
23612	PHLDA3	-1.13243556	0.033744693
8601	RGS20	-1.134452308	0.00979058
6785	ELOVL4	-1.13682534	0.007128961
26230	TIAM2	-1.136977286	0.018871806
1535	CYBA	-1.139286416	0.014296784
27090	ST6GALNAC4	-1.14607295	0.000500887
11077	HSF2BP	-1.149298465	0.010983236
162979	ZNF296	-1.15163789	0.012408152
23426	GRIP1	-1.164663296	0.041407039
388242	N/A	-1.166220886	0.008689477
27202	C5AR2	-1.16648211	0.036878321
776	CACNA1D	-1.169345714	0.01551653
9545	RAB3D	-1.172134765	0.026886525
256236	NAPSB	-1.175276449	0.031322095
55314	TMEM144	-1.17921781	0.038003221
9518	GDF15	-1.186086909	0.014998885
113157	RPLP0P2	-1.186220842	0.008202717
646851	FAM227A	-1.187099076	0.015739533
414236	C10ORF55	-1.195957796	0.001786196
92521	SPECC1	-1.1974047	0.021947882
2980	GUCA2A	-1.197585716	0.04458967
55789	DEPDC1B	-1.199399506	0.028548115

55282	LRRC36	-1.200214323	0.013313551
924	CD7	-1.201465552	0.014086465
400745	SH2D5	-1.208124794	0.011782893
83850	ESYT3	-1.208825823	0.029734107
23066	CAND2	-1.212122476	0.019659493
10360	NPM3	-1.220103263	0.000654042
23529	CLCF1	-1.22140284	0.00967202
57801	HES4	-1.222554645	0.032596815
84519	ACRBP	-1.223657882	0.008019256
4192	MDK	-1.224639048	0.025373945
576	ADGRB2	-1.229399285	0.047938818
1298	COL9A2	-1.229597063	0.045726575
3684	ITGAM	-1.238551841	0.006721281
9856	KIAA0319	-1.243688085	0.032705162
145200	LINC00239	-1.248696381	0.023904054
747	DAGLA	-1.254390877	0.012672595
25827	FBXL2	-1.256705303	0.020227488
64220	STRA6	-1.259405842	0.027283065
55529	PIP4P2	-1.262381595	0.000696795
6510	SLC1A5	-1.266859375	0.00577065
407977	TNFSF12-TNFSF13	-1.26861428	0.03077803
23025	UNC13A	-1.268929805	0.037057878
1493	CTLA4	-1.271554687	0.02913381
113828	FAM83F	-1.272894852	0.048424928
54503	ZDHHC13	-1.273068318	0.00207166
10863	ADAM28	-1.273624748	0.02419929
7220	TRPC1	-1.280881195	0.012472986
3691	ITGB4	-1.281794737	0.010026231
23682	RAB38	-1.283349609	0.014900612
7941	PLA2G7	-1.28392084	0.004060545
7301	TYRO3	-1.288235317	0.019781783
1400	CRMP1	-1.292107411	0.008018379
57722	IGDCC4	-1.29255191	0.035936175
80003	PCNX2	-1.29901222	0.021772175
3159	HMGA1	-1.30253957	9.75976E-05
132884	EVC2	-1.311132937	0.040614555
57643	ZSWIM5	-1.311689149	0.03798126
9456	HOMER1	-1.313607369	0.026141901
9200	HACD1	-1.315112366	0.044482718
497189	TIFAB	-1.315389465	0.006037284
6277	S100A6	-1.316791843	0.036219682
2539	G6PD	-1.317048445	0.001703053
388931	MFSD2B	-1.317983859	0.00077108
89832	CHRFAM7A	-1.321601004	0.004278256
26011	TENM4	-1.325369943	0.029844517
161582	DNAAF4	-1.330540034	0.005387249
114801	TMEM200A	-1.333459974	0.007894232
80736	SLC44A4	-1.338530263	0.040276028
29108	PYCARD	-1.34676578	0.009911439
55561	CDC42BPG	-1.347265387	0.045563509
284340	CXCL17	-1.348075016	0.008394606
57692	MAGEE1	-1.34926894	0.001380695
80023	NRSN2	-1.34937467	0.005209377
55423	SIRPG	-1.351306906	0.025639179
11074	TRIM31	-1.355253758	0.048876061
57628	DPP10	-1.35578116	0.04172042
348378	FAM159A	-1.36770707	0.01956459

55227	LRRC1	-1.367938683	0.023575252
1117	CHI3L2	-1.369250662	0.049870026
6689	SPIB	-1.380488312	0.027648136
4355	MPP2	-1.381624013	0.008745042
375616	KCP	-1.38341295	0.022425565
4148	MATN3	-1.388748986	0.043829176
6273	S100A2	-1.388918665	0.019282009
933	CD22	-1.390885285	0.033099725
54509	RHOF	-1.392194491	0.034767123
114804	RNF157	-1.39324632	0.013949287
84264	HAGHL	-1.393277939	0.046310798
5450	POU2AF1	-1.40098956	0.039910143
79734	KCTD17	-1.404570401	0.006572265
8877	SPHK1	-1.404672875	0.02451389
1608	DGKG	-1.404689961	0.023411239
90853	SPOCD1	-1.407767405	0.005753487
114787	GPRIN1	-1.410402506	0.007884511
54478	PIMREG	-1.421664708	0.005802751
6349	CCL3L3	-1.42535366	0.006302968
9609	RAB36	-1.430160057	0.03089751
5973	RENBP	-1.439057913	0.009121925
2151	F2RL2	-1.448988807	0.026650913
10626	TRIM16	-1.452216216	0.0126709
51127	TRIM17	-1.456132775	0.016920904
148113	CILP2	-1.463695141	0.037710396
131405	TRIM71	-1.465014572	0.015819164
84033	OBSCN	-1.474618786	0.012984586
83871	RAB34	-1.47493916	0.002970477
55356	SLC22A15	-1.475077726	0.025600542
5270	SERPINE2	-1.484552495	0.024661293
5754	PTK7	-1.485693048	0.008529629
167838	TXLNB	-1.489605993	0.006420802
255231	MCOLN2	-1.496152338	0.010100331
84221	SPATC1L	-1.496327558	0.017820701
93082	NEURL3	-1.497581535	0.040512603
79805	VASH2	-1.506449199	0.01537649
23057	NMNAT2	-1.509176077	0.010704926
6338	SCNN1B	-1.52265235	0.010654044
222962	SLC29A4	-1.525102697	0.048114919
6375	XCL1	-1.530709383	0.001339958
92086	GGTLC1	-1.540560738	0.011919372
150	ADRA2A	-1.542528983	0.016362646
2191	FAP	-1.547925234	0.01500984
22996	TTC39A	-1.552459263	0.038058394
8038	ADAM12	-1.553417812	0.022793618
780	DDR1	-1.555946838	0.012187931
81793	TLR10	-1.557892672	0.007600218
6541	SLC7A1	-1.569101562	0.003991509
147645	VSIG10L	-1.56965073	0.002028702
26027	ACOT11	-1.572772812	0.007549676
619373	MBOAT4	-1.574093059	0.008649112
8745	ADAM23	-1.581051475	0.0441911
90993	CREB3L1	-1.594216574	0.024166104
8712	PAGE1	-1.602764329	0.040192107
83879	CDCA7	-1.626855103	0.018210848
51207	DUSP13	-1.668968611	0.006191255
1410	CRYAB	-1.672400633	0.02255992

3957	LGALS2	-1.682658945	0.005973089
23213	SULF1	-1.683020713	0.008916165
169044	COL22A1	-1.687711872	0.037154477
64208	POPDC3	-1.697012069	0.020110012
26279	PLA2G2D	-1.707557127	0.026349379
57476	GRAMD1B	-1.717122865	0.044382573
10631	POSTN	-1.72077675	0.042113235
55630	SLC39A4	-1.721344593	0.023717347
333	APLP1	-1.724916584	0.029659163
135295	SRSF12	-1.755683062	0.00075067
4320	MMP11	-1.762124576	0.001023686
4318	MMP9	-1.787391604	0.008917117
112609	MRAP2	-1.805823671	0.026185068
5831	PYCR1	-1.808100754	0.028409566
26256	CABYR	-1.810042328	0.040877628
80201	HKDC1	-1.830744644	0.028509176
50804	MYEF2	-1.833725552	0.01546675
6819	SULT1C2	-1.836897601	0.045437191
54997	TESC	-1.856617454	0.00884767
6297	SALL2	-1.860438165	0.001466065
79695	GALNT12	-1.865928632	0.006563507
2679	N/A	-1.870310376	0.001557157
58473	PLEKHB1	-1.872564506	0.006206063
125931	CEACAM20	-1.878921087	0.013203085
51454	GULP1	-1.885698371	0.035325981
4321	MMP12	-1.885756332	0.011100794
771	CA12	-1.893527267	0.03217292
144501	KRT80	-1.894940548	0.040942869
1014	CDH16	-1.89893698	0.019275299
94234	FOXQ1	-1.906900635	0.015768701
9536	PTGES	-1.909019525	0.036247638
284217	LAMA1	-1.910128931	0.017895414
91227	GGTLC2	-1.910946717	0.004041093
80144	FRAS1	-1.911712138	0.025669935
2302	FOXJ1	-1.916580174	0.020127655
26762	HAVCR1	-1.948499845	0.003980754
89780	WNT3A	-1.95071516	0.019441743
8153	RND2	-1.960796744	0.017304413
5225	PGC	-1.995596241	0.043822363
2678	GGT1	-1.998697563	0.00158244
253982	ASPHD1	-2.020689877	0.028169749
51237	MZB1	-2.03119837	0.019798276
23657	SLC7A11	-2.037554309	0.004994662
115584	SLC5A11	-2.064372136	0.021383189
312	ANXA13	-2.066443098	0.038404654
84302	TMEM246	-2.102125265	0.000683001
25817	FAM19A5	-2.111997026	0.02711109
27299	ADAMDEC1	-2.112209899	0.000884125
5603	MAPK13	-2.125895493	0.001565814
2318	FLNC	-2.166737816	0.012697588
26470	SEZ6L2	-2.182053748	0.020941236
4440	MSI1	-2.197551129	0.012551362
126353	MISP	-2.240727761	0.017424813
23532	PRAME	-2.243818832	0.025079322
9245	GCNT3	-2.295442487	0.008008837
54825	CDHR2	-2.357085765	0.043317308
220963	SLC16A9	-2.36514831	0.024378399

2984	GUCY2C	-2.50402611	0.009164089
3352	HTR1D	-2.508409024	0.003310099
65268	WNK2	-2.656948548	0.014960929
22943	DKK1	-2.666057251	0.004906304
4072	EPCAM	-2.682699836	0.03602935
1515	CTSV	-2.699439321	0.000143908
10642	IGF2BP1	-2.713113555	0.006518789
1728	NQO1	-2.736614322	0.004408202
6696	SPP1	-2.740393786	0.016666969

Table S5. Significantly enriched biological pathways of differentially expressed genes.

Category	Ontology	P Value	Fold Enrichment
GO_biological_processes	GO:0006805~xenobiotic metabolic process	3.46E-17	8.87
GO_biological_processes	GO:0055114~oxidation-reduction process	5.66E-15	2.97
KEGG_pathway	hsa01100:Metabolic pathways	1.26E-13	1.96
KEGG_pathway	hsa04610:Complement and coagulation cascades	1.18E-11	6.68
GO_biological_processes	GO:0006699~bile acid biosynthetic process	4.77E-11	15.21
KEGG_pathway	hsa00830:Retinol metabolism	2.51E-09	6.07
GO_biological_processes	GO:0010951~negative regulation of endopeptidase activity	4.15E-09	4.84
GO_biological_processes	GO:0019373~epoxygenase P450 pathway	5.00E-09	14.78
KEGG_pathway	hsa05204:Chemical carcinogenesis	1.15E-08	5.21
KEGG_pathway	hsa00120:Primary bile acid biosynthesis	1.30E-08	12.90
KEGG_pathway	hsa00980:Metabolism of xenobiotics by cytochrome P450	2.12E-08	5.34
GO_biological_processes	GO:0008202~steroid metabolic process	3.70E-08	8.05
KEGG_pathway	hsa00982:Drug metabolism - cytochrome P450	3.87E-08	5.48
KEGG_pathway	hsa04976:Bile secretion	4.83E-08	5.40
GO_biological_processes	GO:0042738~exogenous drug catabolic process	6.80E-08	17.74
GO_biological_processes	GO:0006869~lipid transport	1.18E-07	5.60
GO_biological_processes	GO:0070328~triglyceride homeostasis	2.45E-07	10.24
GO_biological_processes	GO:0017144~drug metabolic process	3.56E-07	9.86
GO_biological_processes	GO:0055085~transmembrane transport	5.19E-07	3.05
GO_biological_processes	GO:0015721~bile acid and bile salt transport	4.43E-06	8.87
KEGG_pathway	hsa03320:PPAR signaling pathway	7.57E-06	4.58
GO_biological_processes	GO:0001523~retinoid metabolic process	1.44E-05	5.24
GO_biological_processes	GO:0006631~fatty acid metabolic process	1.96E-05	5.63
KEGG_pathway	hsa00071:Fatty acid degradation	2.03E-05	5.48
GO_biological_processes	GO:0002576~platelet degranulation	2.99E-05	3.88
GO_biological_processes	GO:0019432~triglyceride biosynthetic process	3.57E-05	8.19
KEGG_pathway	hsa00010:Glycolysis / Gluconeogenesis	3.96E-05	4.26
GO_biological_processes	GO:0006069~ethanol oxidation	4.65E-05	13.31
KEGG_pathway	hsa00260:Glycine, serine and threonine metabolism	4.67E-05	5.62

GO_biological_processes	GO:0042632~cholesterol homeostasis	1.25E-04	4.57
KEGG_pathway	hsa00350:Tyrosine metabolism	1.33E-04	5.64
GO_biological_processes	GO:0051919~positive regulation of fibrinolysis	2.04E-04	26.61
GO_biological_processes	GO:0097267~omega-hydroxylase P450 pathway	2.13E-04	14.78
GO_biological_processes	GO:0009636~response to toxic substance	3.26E-04	3.76
KEGG_pathway	hsa04146:Peroxisome	3.33E-04	3.44
GO_biological_processes	GO:0045723~positive regulation of fatty acid biosynthetic process	3.44E-04	13.31
GO_biological_processes	GO:0043691~reverse cholesterol transport	4.17E-04	8.87
GO_biological_processes	GO:0007597~blood coagulation, intrinsic pathway	4.17E-04	8.87
KEGG_pathway	hsa00590:Arachidonic acid metabolism	4.26E-04	3.89
GO_biological_processes	GO:0042157~lipoprotein metabolic process	4.63E-04	5.60
GO_biological_processes	GO:0051180~vitamin transport	4.97E-04	21.29
GO_biological_processes	GO:0006071~glycerol metabolic process	5.25E-04	12.10
GO_biological_processes	GO:0006953~acute-phase response	5.46E-04	5.46
KEGG_pathway	hsa00500:Starch and sucrose metabolism	5.75E-04	5.32
KEGG_pathway	hsa01130:Biosynthesis of antibiotics	5.96E-04	2.28
GO_biological_processes	GO:0006810~transport	6.94E-04	2.06
KEGG_pathway	hsa02010:ABC transporters	7.03E-04	4.49
GO_biological_processes	GO:0030449~regulation of complement activation	7.48E-04	6.21
GO_biological_processes	GO:0042730~fibrinolysis	9.03E-04	7.60
GO_biological_processes	GO:0016098~monoterpenoid metabolic process	9.66E-04	17.74
GO_biological_processes	GO:0042737~drug catabolic process	9.66E-04	17.74
GO_biological_processes	GO:0006957~complement activation, alternative pathway	0.001070284	10.24
GO_biological_processes	GO:0006637~acyl-CoA metabolic process	0.001132542	7.26
GO_biological_processes	GO:0043252~sodium-independent organic anion transport	0.001402936	6.94
GO_biological_processes	GO:0051289~protein homotetramerization	0.001717738	3.99
GO_biological_processes	GO:0015695~organic cation transport	0.001924609	8.87
GO_biological_processes	GO:0008209~androgen metabolic process	0.00249067	8.32
GO_biological_processes	GO:0010898~positive regulation of triglyceride catabolic process	0.00255538	13.31

GO_biological_processes	GO:0006633~fatty acid biosynthetic process	0.003122356	4.09
GO_biological_processes	GO:0055091~phospholipid homeostasis	0.003726728	11.83
GO_biological_processes	GO:0031638~zymogen activation	0.003726728	11.83
GO_biological_processes	GO:0007596~blood coagulation	0.004023522	2.31
GO_biological_processes	GO:0050746~regulation of lipoprotein metabolic process	0.004111183	26.61
GO_biological_processes	GO:0030573~bile acid catabolic process	0.004111183	26.61
KEGG_pathway	hsa00983:Drug metabolism - other enzymes	0.004345618	3.82
GO_biological_processes	GO:0008152~metabolic process	0.004412615	2.38
GO_biological_processes	GO:0042593~glucose homeostasis	0.004607137	2.90
GO_biological_processes	GO:0042572~retinol metabolic process	0.004786067	5.32
GO_biological_processes	GO:0006730~one-carbon metabolic process	0.004786067	5.32
GO_biological_processes	GO:0006855~drug transmembrane transport	0.004851097	7.00
KEGG_pathway	hsa01200:Carbon metabolism	0.005030137	2.52
KEGG_pathway	hsa04922:Glucagon signaling pathway	0.005057234	2.66
GO_biological_processes	GO:0051156~glucose 6-phosphate metabolic process	0.005176482	10.64
GO_biological_processes	GO:0046470~phosphatidylcholine metabolic process	0.005176482	10.64
GO_biological_processes	GO:0006000~fructose metabolic process	0.005176482	10.64
GO_biological_processes	GO:0031667~response to nutrient levels	0.005533575	5.15
GO_biological_processes	GO:0006094~gluconeogenesis	0.005710899	4.23
KEGG_pathway	hsa00910:Nitrogen metabolism	0.006220473	6.45
GO_biological_processes	GO:0007584~response to nutrient	0.006410732	3.24
GO_biological_processes	GO:0017187~peptidyl-glutamic acid carboxylation	0.00692097	9.68
GO_biological_processes	GO:0010867~positive regulation of triglyceride biosynthetic process	0.00692097	9.68
GO_biological_processes	GO:0032094~response to food	0.007059136	6.34
GO_biological_processes	GO:0001676~long-chain fatty acid metabolic process	0.007059136	6.34
GO_biological_processes	GO:0006651~diacylglycerol biosynthetic process	0.008017984	19.96
GO_biological_processes	GO:0019448~L-cysteine catabolic process	0.008017984	19.96
GO_biological_processes	GO:0036101~leukotriene B4 catabolic process	0.008017984	19.96

GO_biological_processes	GO:0010897~negative regulation of triglyceride catabolic process	0.008017984	19.96
GO_biological_processes	GO:0015711~organic anion transport	0.008973455	8.87
GO_biological_processes	GO:0030212~hyaluronan metabolic process	0.008973455	8.87
GO_biological_processes	GO:0070989~oxidative demethylation	0.008973455	8.87
GO_biological_processes	GO:0007586~digestion	0.009074707	3.38
GO_biological_processes	GO:0042594~response to starvation	0.009351338	4.56
KEGG_pathway	hsa05150:Staphylococcus aureus infection	0.010557477	3.25
GO_biological_processes	GO:0051384~response to glucocorticoid	0.010712715	3.28
GO_biological_processes	GO:0042493~response to drug	0.01097762	1.84
KEGG_pathway	hsa04973:Carbohydrate digestion and absorption	0.011053965	3.66
GO_biological_processes	GO:0019915~lipid storage	0.011469433	5.54
GO_biological_processes	GO:0007188~adenylate cyclase-modulating G-protein coupled receptor signaling pathway	0.011811436	4.32
GO_biological_processes	GO:0007204~positive regulation of cytosolic calcium ion concentration	0.01214725	2.38
GO_biological_processes	GO:0015812~gamma-aminobutyric acid transport	0.013032193	15.97
GO_biological_processes	GO:0019439~aromatic compound catabolic process	0.013032193	15.97
GO_biological_processes	GO:0002682~regulation of immune system process	0.013032193	15.97
GO_biological_processes	GO:0033344~cholesterol efflux	0.013257496	5.32
KEGG_pathway	hsa00051:Fructose and mannose metabolism	0.013760322	4.11
GO_biological_processes	GO:0006508~proteolysis	0.013913326	1.60
GO_biological_processes	GO:0033700~phospholipid efflux	0.014041733	7.60
GO_biological_processes	GO:0042573~retinoic acid metabolic process	0.014041733	7.60
GO_biological_processes	GO:0007155~cell adhesion	0.014451138	1.62
GO_biological_processes	GO:0046487~glyoxylate metabolic process	0.015213524	5.12
GO_biological_processes	GO:0006629~lipid metabolic process	0.015330745	2.20
GO_biological_processes	GO:0006691~leukotriene metabolic process	0.017071014	7.10
GO_biological_processes	GO:0034375~high-density lipoprotein particle remodeling	0.017071014	7.10

GO_biological_processes	GO:0010575~positive regulation of vascular endothelial growth factor production	0.017342128	4.93
GO_biological_processes	GO:0006749~glutathione metabolic process	0.018035879	3.33
KEGG_pathway	hsa00232:Caffeine metabolism	0.01878889	13.16
GO_biological_processes	GO:0042426~choline catabolic process	0.019065477	13.31
GO_biological_processes	GO:0009804~coumarin metabolic process	0.019065477	13.31
GO_biological_processes	GO:0006739~NADP metabolic process	0.019065477	13.31
GO_biological_processes	GO:0046483~heterocycle metabolic process	0.019065477	13.31
GO_biological_processes	GO:0008206~bile acid metabolic process	0.020435709	6.65
GO_biological_processes	GO:0006898~receptor-mediated endocytosis	0.023314405	2.00
KEGG_pathway	hsa04931:Insulin resistance	0.024468004	2.23
GO_biological_processes	GO:0097264~self proteolysis	0.026034557	11.41
GO_biological_processes	GO:0010890~positive regulation of sequestering of triglyceride	0.026034557	11.41
GO_biological_processes	GO:0042908~xenobiotic transport	0.026034557	11.41
GO_biological_processes	GO:0015872~dopamine transport	0.026034557	11.41
GO_biological_processes	GO:0001867~complement activation, lectin pathway	0.026034557	11.41
GO_biological_processes	GO:0044539~long-chain fatty acid import	0.026034557	11.41
KEGG_pathway	hsa00270:Cysteine and methionine metabolism	0.02751594	3.46
GO_biological_processes	GO:0060384~innervation	0.028175749	5.91
GO_biological_processes	GO:0045766~positive regulation of angiogenesis	0.029031787	2.31
GO_biological_processes	GO:0050727~regulation of inflammatory response	0.030367785	2.96
KEGG_pathway	hsa04975:Fat digestion and absorption	0.030429688	3.37
KEGG_pathway	hsa00650:Butanoate metabolism	0.032142023	4.06
GO_biological_processes	GO:0006814~sodium ion transport	0.032219408	2.63
GO_biological_processes	GO:0048260~positive regulation of receptor-mediated endocytosis	0.032549122	5.60
GO_biological_processes	GO:0030198~extracellular matrix organization	0.033745694	1.90
GO_biological_processes	GO:0090181~regulation of cholesterol metabolic process	0.033860949	9.98
KEGG_pathway	hsa00460:Cyanoamino acid metabolism	0.037147063	9.40
GO_biological_processes	GO:0019370~leukotriene biosynthetic process	0.037254292	5.32
GO_biological_processes	GO:0006096~glycolytic process	0.037352177	3.91

GO_biological_processes	GO:0007626~locomotory behavior	0.038188298	2.53
GO_biological_processes	GO:0006935~chemotaxis	0.040165559	2.18
KEGG_pathway	hsa00591:Linoleic acid metabolism	0.040543115	3.78
GO_biological_processes	GO:0006641~triglyceride metabolic process	0.040965189	3.80
GO_biological_processes	GO:0006865~amino acid transport	0.040965189	3.80
GO_biological_processes	GO:0008203~cholesterol metabolic process	0.041920546	2.74
GO_biological_processes	GO:0009749~response to glucose	0.041920546	2.74
GO_biological_processes	GO:0019835~cytolysis	0.042286818	5.07
GO_biological_processes	GO:0016125~sterol metabolic process	0.042286818	5.07
GO_biological_processes	GO:0006544~glycine metabolic process	0.042470726	8.87
GO_biological_processes	GO:0007614~short-term memory	0.042470726	8.87
GO_biological_processes	GO:0072488~ammonium transmembrane transport	0.042470726	8.87
GO_biological_processes	GO:0090331~negative regulation of platelet aggregation	0.042470726	8.87
GO_biological_processes	GO:0070327~thyroid hormone transport	0.042470726	8.87
GO_biological_processes	GO:0045471~response to ethanol	0.043775439	2.28
GO_biological_processes	GO:0035690~cellular response to drug	0.044520986	2.70
GO_biological_processes	GO:0071333~cellular response to glucose stimulus	0.044713088	3.07
GO_biological_processes	GO:0009409~response to cold	0.044769293	3.70
GO_biological_processes	GO:0006956~complement activation	0.044839931	2.45
KEGG_pathway	hsa00140:Steroid hormone biosynthesis	0.046805244	2.65
KEGG_pathway	hsa00561:Glycerolipid metabolism	0.046805244	2.65
GO_biological_processes	GO:0006954~inflammatory response	0.049012626	1.54