Title: Consensus molecular subtype differences linking colon adenocarcinoma and obesity revealed by a cohort transcriptomic analysis

Short Title: Obesity and colon cancer

Michael W. Greene[1]*, Peter T. Abraham[2], Peyton C. Kuhlers[1,3], Elizabeth A. Lipke[2], Martin J. Heslin[4], Stanley T. Wijaya[1] and Ifeoluwa Odeniyi[1]

[1]Department of Nutrition, Dietetics, and Hospitality Management, Auburn University, Auburn, AL 36849, USA
[2]Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA
[3]Department of Biostatistics, University of North Carolina at Chapel Hill, NC 27599, USA
[4]Mitchell Cancer Institute, The University of South Alabama, Mobile, AL 36604, USA

*Corresponding author:
Michael W. Greene
mwgreene@auburn.edu

Abstract

Background

Colorectal cancer (CRC) is the third-leading cause of cancer-related deaths in the United States and worldwide. Obesity - a worldwide public health concern - is a known risk factor for cancer including CRC. However, the mechanisms underlying the link between CRC and obesity have yet to be fully elucidated in part because of the molecular heterogeneity of CRC. We hypothesized that obesity modulates CRC in a consensus molecular subtype (CMS)-dependent manner.

Methods

RNA-seq data and associated tumor and patient characteristics including body weight and height data for 232 patients were obtained from The Cancer Genomic Atlas – Colon Adenocarcinoma (TCGA-COAD) database. Tumor samples were classified into the four CMSs with the CMScaller R package; Body mass index (BMI) was calculated and categorized as normal, overweight, and obese.

Results

We observed a significant difference in CMS categorization between BMI categories. Differentially expressed genes (DEGs) between obese and overweight samples and normal samples differed across the CMSs, and associated prognostic analyses indicated that the DEGs had differing effects on survival. Using Gene Set Enrichment Analysis, we found differences in Hallmark gene set enrichment between obese and overweight samples and normal samples across the CMSs. We constructed Protein-Protein Interaction networks and observed differences in obesity-regulated hub genes for each CMS. Finally, we analyzed and found differences in predicted drug sensitivity between obese and overweight samples and normal samples across the CMSs.

Conclusions

Thus, we conclude that obesity has CMS-specific effects on the CRC tumor transcriptome.

Keywords:

Colon cancer, obesity, consensus molecular subtype, transcriptomics

Background

Improvements in colorectal cancer (CRC) screening, diagnosis, advanced surgical techniques, and preoperative and postoperative treatment have led to reduced CRC incidence and mortality (1). Yet, CRC incidence rates remain high in states with a high prevalence of obesity (2). CRC remains the third most common non skin-related cancer and the third leading cause of cancer-related mortality in the United States (3). In addition to obesity, smoking, an unhealthy diet, high alcohol consumption, and physical inactivity are well-known risk factors that are potentially preventable (4).

Based on bulk transcriptomics, 4 consensus and 1 unclassified CRC consensus molecular subtypes have been proposed (CMS1-4) (5, 6). CMS1 has been termed a microsatellite instability (MSI) immune subtype based on the clustering of MSI tumors and the strong infiltration of immune cells. Worse survival after relapse is associated with patients with CMS1 tumors (6). CMS2 and CMS4 – the two most prevalent molecular subtypes of CRC – represent 37% and 23% of early-stage CRC tumors, respectively (6, 7). The CMS2 subtype has been termed 'canonical' due to the upregulation of classical CRC pathways first proposed by Fearon and Vogelstein (8). The CMS3 subtype has been termed the metabolic subtype due to the metabolic dysregulation at the transcriptome level. Lastly, CMS4 has been termed 'mesenchymal' due to activation of epithelial–mesenchymal transition (EMT) and overexpression of proteins implicated in ECM remodeling and complement signaling thought to be mediated by a stromal-enriched inflamed microenvironment (5). There is a significantly higher risk of distant relapse and death for patients diagnosed with early-stage CMS4 (6).

The epidemiological evidence linking CRC with obesity, and its associated pathophysiological metabolic state is strong (9-12). Even though colon and rectal cancer are grouped together as CRC, abdominal obesity and the metabolic syndrome are more strongly linked to colon cancer than rectal cancer (9, 11). Obesity and the associated pathophysiological conditions of insulin resistance and inflammation afflict approximately a third of the adult population in the United States (13, 14). The pathophysiological state associated with obesity includes visceral adipose tissue and hepatic dysfunction which leads to systemic insulin resistance and inflammation, the dysregulation of adipokines, and dysbiosis (microbial imbalance) (15, 16); all of these pathophysiological alterations have been hypothesized to promote a favorable niche for the pathogenesis of CRC (17-21). Insulin resistant visceral adipose tissue participates in cross-talk with CRC and promotes a favorable niche by secreting metabolites, growth factors, and proinflammatory cytokines (22-24). Elevated pro-inflammatory cytokines are associated with an increased risk of CRC, (9, 25-28) and a circulating inflammatory signature (high miR-21, IL-6, and IL-8) predicts lower progression-free and overall survival of patients with metastatic CRC (29). The obese pathophysiological state of insulin resistance and inflammation have been shown to stimulate CRC tumor growth in animal models (30-37).

Thus, there is compelling epidemiological and experimental evidence linking obesity to CRC (9-12) – although more strongly for colon cancer (9, 11). The obesity-cancer link is thought to be driven by multiple obesity-derived factors that activate pathways mediating cell signaling, proliferation, and tumor progression (38). Yet, there does not exist a framework for activation of these obesity-driven cell pathways. Thus, we questioned whether the effect of obesity on cell signaling, proliferation, and tumor progression pathways in CRC tumors is dependent on the CMS of the tumor. Therefore, we undertook a study to examine the transcriptomic profile of colon adenocarcinoma tumors from obese patients compared to healthy and overweight BMI patients to determine whether obesity modulates cell signaling, proliferation, and tumor progression pathways in a similar manner across the four CMSs. We also examined whether prognostic survival outcomes and predicted drug response in obesity associated differentially expressed genes (DEGs) is similar between the four CMSs. Our secondary objective was to determine whether the transcriptomic profile of tumors from overweight patients compared to healthy BMI patients is similar between the four CMSs.

Methods

Patients:
Our approach is shown in Figure 1. RNA-seq data (HT-Seq counts) and associated tumor and patient characteristics for 454 patients were obtained from The Cancer Genomic Atlas – Colon Adenocarcinoma (TCGA-COAD) database using the R package TCGAbiolinks (39) which extracts and collates data from the Genomic Data Commons (40). The patient characteristics of sex, age, ethnicity, and race were obtained from all 454 patients. Body weight and height were available for only 232 patients. Body mass index (BMI) was calculated, and each patient was categorized as underweight, (BMI < 18.5) normal (BMI 18.5 – 24.9), overweight (BMI 25.0 – 29.9), and obese (BMI >30.0). The following tumor characteristics were obtained for all patients: tumor location, tumor stage (classified as Stage I – IV), number of lymph nodes examined, and number of positive lymph nodes. Lymph node ratio (LNR) was calculated as the relation of tumor-infiltrated to total examined lymph nodes and classified as LNR0 – LNR4 based on the cut-off values 0.17, 0.41, and 0.69 (41). Patient tumor samples from the TCGA-COAD were assigned to a CMS (6) using the R package CMScaller (42), which utilizes a nearest-template prediction algorithm (43). Samples with an FDR greater than 0.05 were not classified. The study was approved by the Auburn University Institutional Review Board (#20-509 EX 2010).

Transcriptomic Analysis:
The R package DESeq2 (44) was used to assess differential gene expression. Raw counts and associated phenotypes were inputted into DESeq2, and the following comparisons were made: obese vs. normal, obese vs. overweight, and overweight vs. normal for each CMS category. Genes with a base mean expression greater than ten and an adjusted p-value less than 0.05 were used for downstream analysis and visualization Volcano plots were generated using the R package EnhancedVolcano [https://github.com/kevinblighe/EnhancedVolcano]. Gene overlap between comparisons were visualized using Euler diagrams and upset plots from the R packages eulerr [https://github.com/jolars/eulerr] and UpsetR [6], respectively.

Gene Set Enrichment Analysis (GSEA) was performed using the desktop GSEA software (version 4.0.3) from the Broad Institute (45, 46). Normalized RNA-seq counts obtained from DESeq2 were used. Permutation type was set to 'gene_set,' and Human Ensembl was selected as the CHIP platform. Hallmark gene sets (47) of well-defined biological states and processes (version 7.2) were assessed in the obese vs. normal, obese vs. overweight, and overweight vs. normal comparisons for each CMS category. Only gene sets with a false discovery rate q-value less than 0.05 were reported. The normalized enrichment score (NES) was reported for the gene set.

To construct a protein-protein interaction (PPI) network from DESeq2-obtained DEGs, the Search Tool for the Retrieval of Interacting Genes (STRING; version 10.0; string-db.org) online database for PPI network construction (48) was used in Cytoscape (v3.7.1, National Resource for Network Biology, https://cytoscape.org/) (49), a bioinformatics platform for visualizing molecular interaction networks. The Cytohubba package in Cytoscape was used to perform the hub gene analysis (50) for the obese vs. normal, obese vs. overweight, and overweight vs. normal comparisons for each CMS category. Hub genes were identified using the maximal clique centrality (MCC) method to obtain the top 10 ranked genes in all modules. The hub genes were queried for mRNA expression using the Gene Expression database of Normal and Tumor tissues 2 (GENT2) (http://gent2.appex.kr.) (51) and the Oncomine database (https://www.oncomine.org/) (52). Colon normal tissue (n = 397) and cancer (n = 3775) microarray data collected from the NCBI GEO public database generated using the Affymetrix U133Plus2 platform was used to assess gene expression using GENT2. Violin plots were used to display the distribution of the data. The violin plots were generated using the ggplot2 library in R Studio. We examined gene expression in colon adenocarcinoma versus normal patient samples from the: 1) Hong Colorectal cohort (normal, n = 12; colon adenocarcinoma, n = 70) from Singapore (53); 2) Skrzypczak Colorectal cohort (normal, n = 24; colon adenocarcinoma, n = 36) from Poland (54), and Kaiser Colorectal cohort (normal, n = 5; colon

adenocarcinoma, n = 41) from the United States (55) using the Oncomine database. Box plots were used to display the distribution of the data.

To examine prognostic patient outcomes the expression of the top 20 significantly upregulated DESeq2-obtained DEGs in the obese vs. normal, obese vs. overweight, and overweight vs. normal comparisons for each CMS category was assessed using the PROGgeneV2 tool was used (56, 57). We selected the CRC cohorts with patients from the United States for which age, gender, and tumor stage were available: GSE17536 (58) and GSE41258 (59) for overall survival and GSE14333 (60) and GSE17536 (58) for relapse-free survival. All cohorts were adjusted for age, stage, and gender covariates and bifurcated based on median expression. The hazard ratios, 95% confidence intervals, and p values were reported.

For assessment of predicted drug sensitivity, the pRRophetic R package (61) was used with normalized RNA-seq counts from normal, overweight, and obese patient tumors to estimate the half maximal inhibitory concentration (IC50) of 130 drugs from the Cancer Genomic Project (CGP; ref. (62)).

Statistical Analysis:
All data analyses were conducted with RStudio and Rx64 3.6.0 software environment. Differences in patient demographic and tumor characteristics between obese, overweight, and normal BMI were analyzed using Fishers exact test. Predicted drug sensitivity was assessed using t-tests to determine differences between obese and overweight BMI patients and normal BMI patients. Gene expression between colon adenocarcinoma and normal colon tissue in the CRC cohorts was assessed within the Oncomine database using t-tests. A significance level of 0.05 was established for all statistical tests.

Results

Patient and tumor characteristics
The TCGA COAD cohort contained both RNA-seq data and weight and height data for 232 patients out of 454 total patients. Weight and height data was used to calculate BMI and classify patients as underweight, normal, overweight, and obese while RNA seq data was used to classify the tumors by the CMS. After exclusion of the one underweight patient, a final cohort of 231 patient's demographic and clinical tumor data was assessed across the normal, overweight, and obese categories (Table 1). Approximately 85% of patient samples in our final cohort were from people located in the US and no significant differences were observed in the geographic site of patient samples across the BMI categories (data not shown). Further, no significant differences in sex, age, or ethnicity were observed across the BMI categories. In contrast, race was significantly different across the BMI categories ($p = 0.001$): a higher proportion of Asian patients were observed in the normal (10%) compared to the obese (0%) BMI category while a greater proportion of Black or African American patients were observed in the obese (30%) compared to the normal (18%) BMI category. No significant differences in the stage and location of the tumor, nor the lymph node ratio, was observed across the BMI categories. In contrast, the CMS classification of tumors was significantly different across the BMI categories ($p = 0.040$): a greater proportion of CMS3 tumors was observed in the obese (22%) compared to the normal (4%) BMI category. Consistent with this observation, the highest average BMI was in the CMS3 group (31.2) followed CMS4 (28.6), CMS1 (27.4), and CMS2 (27.1).

CMS specific differentially expressed genes
Our finding that the proportion of CMS3 tumors differed across BMI categories suggested that there may be CMS specific transcriptomic differences in the obese BMI category compared to the normal and overweight BMI categories. Thus, we examined differentially expressed genes (DEGs) from the RNA seq data between normal, overweight, and obese BMI patients for each CMS. As shown in Figure 2A, Euler diagrams of DEGs demonstrate a unique pattern of overlapping DEGs for each CMS. A significant difference ($p < 0.001$) in the percentage of obesity-regulated DEGs (overlap between the obese vs. normal

and obese vs. overweight comparisons) was observed. CMS1 (70%) and CMS4 (68%) had greater obesity-related DEG overlap than CMS2 (7%) and CMS3 (3%). We next examined whether the obese vs. normal DEGs overlapped between the four CMS categories. The greatest overlap in DEGs was between CMS2 and CMS4 (Fig 2B). A similar result was observed in the overweight vs. normal comparison (Fig 2B). In contrast, we observed in the obese vs. overweight DEGs that the greatest overlap in DEGs was between CMS3 and CMS4 (Fig 2B). Volcano plots of DEGs were generated to examine the pattern of overexpressed vs. underexpressed transcripts between normal, overweight, and obese BMI patients for each CMS (Suppl. Fig 1). We observed that only CMS3 had more overexpressed transcripts than underexpressed transcripts for the obese vs. normal comparison while both CMS3 and CMS4 had more overexpressed transcripts than underexpressed transcripts for the overweight vs. normal comparison. In contrast, all CMS groups had less overexpressed transcripts than underexpressed transcripts for the obese vs. overweight comparison.

To examine the clinical relevance of the CMS specific DEGs, we generated Kaplan-Meier plots (57) to assess overall survival and relapse-free survival in CRC cohorts from the United States. Using the top 20 upregulated DEGs identified in the obese to normal BMI comparison, we observed significantly reduced overall survival, adjusted for age, gender, stage and grade, for high expression of the CMS3 DEGs ($p = 0.016$; HR = 3.73, 95% CI 1.27-10.93) and CMS4 DEGs DEGs ($p = 0.026$; HR = 2.79, 95% CI 1.13-6.88) (Fig 3). Significantly reduced relapse-free survival, adjusted for age, gender, stage and grade, was observed for high expression of CMS2 DEGs ($p = 0.016$; HR = 13.47, 95% CI 1.61-113) and CMS3 DEGs ($p = 0.042$; HR = 2.48, 95% CI 1.03-5.96) (Fig 2B). Only high expression of CMS1 DEGs was associated with significantly reduced relapse-free survival in the obese to overweight BMI comparison ($p = 0.016$; HR = 14.33, 95% CI 1.23-167). Using the top 20 upregulated DEGs identified in the overweight to normal BMI comparison, we also observed significantly reduced overall and relapse-free survival, adjusted for age, gender, stage and grade, for high expression of CMS2, CMS3, and CMS4 DEGs (Suppl. Fig 2). Taken together, these results indicate CMS specific differences in obesity-regulated DEGs.

CMS specific gene set enrichment

To further examine obesity-related transcriptomic differences in the four CMSs, we performed a gene set enrichment analysis comparing RNA seq data from obese to normal BMI patients for each CMS. As shown in Figure 4, immune-related Hallmark gene sets with an FDR ($p < 0.05$) were enriched in CMS1, CMS2, and CMS4. Enrichment in interferon gamma response and interferon alpha response overlapped between CMS1, CMS2, and CMS4. Immune-related gene sets were almost exclusively enriched in both CMS1 (8 out of 10) and CMS2 (2 out of 2). In contrast, Hallmark gene sets enriched in CMS4 contained gene sets regulating DNA repair, MTORC signaling, unfolded protein response, G2M checkpoint, and protein secretion. Interestingly, significant Hallmark gene set enrichment was not observed in CMS3. Analysis of Hallmark gene set enrichment with an FDR ($p < 0.05$) in the obese vs. overweight comparison revealed that CMS1 gene set enrichment was dominated by immune-related gene sets (8 out of 14) (Fig 4). In contrast, immune-related gene sets were not enriched in CMS2, CMS3, and CMS4 except for the coagulation gene set in CMS3. The only significant enrichment in CMS2 was the Hedgehog signaling gene set. The enrichment of the gene sets for Myc targets V1, Myc targets V2, WNT beta catenin signaling, p53 pathway, and DNA repair overlapped between CMS3 and CMS4 while Myc targets V1 also overlapped with CMS1. Unique to CMS4 were 18 enriched gene sets related to proliferation (MTORC signaling), metabolic (cholesterol homeostasis and glycolysis), and stromal signaling (WNT beta catenin signaling, TGF beta signaling, and Notch signaling). Although, the WNT beta catenin signaling gene set overlapped between CMS3 and CMS4. In addition to examining obesity-related transcriptomic differences, we examined transcriptomic differences between overweight vs. normal BMI patients for each CMS (Suppl. Fig 3). Hallmark gene set enrichment with an FDR ($p < 0.05$) revealed immune-related gene set enrichment in CMS2 and CMS4; there was an overlap in interferon gamma response, interferon alpha response, allograft rejection, inflammatory response, and IL-6 JAK STAT3 signaling gene sets. In contrast, gene set enrichment in CMS1 was strongly related to metabolic

processes (heme metabolism, fatty acid metabolism, and bile acid metabolism). The only significant enrichment in CMS3 was a metabolic process gene set (oxidative phosphorylation). These results indicate CMS specific differences in obesity-regulated immune, metabolic, and stromal signaling gene set enrichment.

CMS specific hub genes

To gain insight into obesity-regulated hub genes within each CMS group, we first constructed a Protein-Protein Interaction (PPI) network using the STRING database module in Cytoscape. From the PPI network, hub genes were identified using the MCC algorithm of the CytoHubba module in Cytoscape. The top ten highest scoring genes in the obese to normal BMI patients and the obese to overweight comparisons for each CMS are shown in Figure 5 and Supplemental Table 1. We observed that in CMS1 obese to normal BMI comparison the top three genes in the hub were *Bassoon presynaptic cytomatrix protein* (*BSN*), *Major synaptic vesicle protein p38* (*SYP*), and *Unc-13 homolog A* (*UNC13A*). In contrast, an immune hub containing *interleukin 10* (*IL-10*), *C-C motif chemokine receptor 2* (*CCR2*), and *serum amyloid A1* (*SAA1*) was observed in the obese to overweight BMI comparison for CMS1. Weak interconnectivity was observed in the obese to normal BMI comparison for CMS2, while a hub containing *NK2 homeobox 1* (*NKX2-1*) and *SRY-box transcription factor 2* (*SOX2*) was observed in the obese to overweight BMI comparison for CMS2. A hub containing four Melanoma-associated antigen genes (*MAGEA6*, *MAGEA3*, *MAGEA11*, and *MAGEA12*) was observed in the obese to normal BMI comparison for CMS3, while an interconnected hub network of genes including *somatostatin* (*SST*) and two somatostatin receptor genes (*SSTR2* and *SSTR5*) but also three C-C motif chemokine receptor genes (*CCR2*, *CCR10*, and *CCR13*) was observed in the obese to overweight BMI comparison for CMS3. In the the obese to normal comparison for CMS4 we observed a hub with UDP glucuronosyltransferase 1 family, polypeptide genes (*UGT1A6*, *UGT1A8*, *UGT1A1*, and *UGT1A10*), while a hub network of genes including *neuromedin U receptor 2* (*NMUR2*), *peptide YY* (*PYY*) and *pro-platelet basic protein* (*PPBP*) was observed in the obese to overweight BMI comparison for CMS4. The hub genes for the overweight to normal BMI comparison were also identified (Suppl. Fig 4 and Suppl. Table 1). In CMS1 there was an overlap in *SYP* with the hub observed in the obese comparison to normal BMI while no overlap in hub genes was observed for CMS2. An overlap in the Melanoma-associated antigen genes (*MAGEA6*, *MAGEA3*, *MAGEA11*, and *MAGEA12*) was observed with the hub observed in the CMS3 obese comparison to normal BMI, while no overlaps were observed in CMS4.

To examine the CRC relevance of the hub genes that overlapped in more than one CMS obese BMI comparisons, we used transcriptomic microarray data compiled from NCBI GEO public database using the GENT2 program for an aggregate data analysis and in three independent colon cancer cohorts using the Oncomine database (51, 52) for an analysis of cohort varibility. We observed that *PYY* transcript expression was significantly ($p < 0.001$) downregulated (Log2 fold change = -4.115) in aggregate colon cancer data (Fig 6A) and in all three colon cancer cohorts (Suppl. Fig 6A). In contrast, *PPBP* transcript expression was significantly ($p < 0.001$) upregulated (Log2 fold change = 2.803) in aggregate colon cancer data (Fig 6B) and in two of the cohorts (Suppl. Fig 6B). *INSL5* was significantly ($p < 0.001$) downregulated (Log2 fold change = -3.966) in colon cancer (Fig 6C) and in one of the three colon cancer cohorts (Suppl. Fig 6C). Differential regulation of the *NPW* transcript expression was observed in aggregate colon cancer data (Fig 6D) and in one of the three cohorts (Suppl. Fig 6D). Finally, *CCR2* gene expression was significantly ($p < 0.001$) downregulated (Log2 fold change = -0.706) in colon cancer (Fig 6E).

CMS specific predicted drug sensitivity

We next examined whether obesity modulated the predicted drug sensitivity of 130 drugs from the Genomics Drug Sensitivity in Cancer (GDSC) screen in a similar manner across the four CMSs. We used a phenotype prediction method in which cell line drug response is applied to patient transcriptomic data (63). The front line chemotherapy drugs fluoropyrimidine, irinotecan, and capecitabine (64) were not

available for analysis in our 130 drug set but data for 8 other drugs targeting DNA replication were available and as was data for another 122 drugs in pathways that have been targeted for CRC (apoptosis, cell cycle, chromatin histone acetylation, cytoskeleton, epidermal growth factor receptor (EGFR), extracellular signal regulated kinase (ERK)/mitogen activated protein kinase (MAPK), genome integrity, insulin-like growth factor receptor (IGFR), c-Jun N-terminal kinase (JNK), metabolism, mitosis, other kinases, phosphoinositide 3 kinase (PI3K)/mammalian target of rapamycin (MTOR), protein stability, receptor tyrosine kinase (RTK), and WNT signaling) (65-67). We observed significantly ($p < 0.05$) increased predicted drug sensitivity for 32 and reduced predicted drug sensitivity for 4 of the 130 drugs between obese and normal BMI patient tumors, primarily in CMS1 (22 drugs) but also in CMS4 (10 drugs) and CMS2 (3 drugs) but not CMS3 (1 drug) (Suppl. Table 2). There were no significant differences in predicted drug sensitivity between obese and normal BMI patient tumors in any of the CMS categories for drugs targeting EGFR, ERK/MAPK, IGF1R, and RTK (Suppl. Table 2). In contrast, we observed significantly ($p < 0.05$) increased predicted drug sensitivity for 4 drugs targeting DNA replication (Fig 7A and Suppl. Table 2). Further, significantly ($p < 0.05$) increased predicted drug sensitivity for 4 drugs targeting MTOR (Fig 7B and Suppl. Table 2), including temsirolimus in both CMS1 and CMS4, was observed. Significantly ($p < 0.05$) reduced predicted drug sensitivity for 2 drugs targeting metabolism (peroxisome proliferator-activated receptor (PPAR)) in both CMS2 and CMS4 was also observed (Fig 7C and Suppl. Table 2). Taken together our findings indicate that there are CMS-specific differences in predicted drug sensitivity between obese and normal BMI patient tumors.

We next examined predicted drug sensitivity of 130 drugs from the GDSC screen in a similar manner across the four CMSs for tumors from overweight compared to normal BMI patients. We observed significantly ($p < 0.05$) increased predicted drug sensitivity for 14 and reduced predicted drug sensitivity for 7 of the 130 drugs primarily in CMS2 (17 drugs) but also in CMS3 (3 drugs) and CMS1 (1 drug) but not CMS4 (0 drugs) (Suppl. Table 3). Consistent with the findings in the obese to normal comparison, there was increase predicted drug sensitivity for the drugs targeting DNA replication (including methotrexate) and reduced predicted drug sensitivity for the drug FH535 which targets PPAR (Suppl. Table 3). Further we observed in CMS 2 tumors differential predicted drug sensitivity for two CMS2 canonical pathways: significantly increased predicted drug sensitivity for the drugs targeting EGFR and significantly reduced predicted drug sensitivity for a drug targeting WNT signaling were observed (Suppl. Table 3).

Discussion

In the current study, we examined the transcriptomic profile of tumors from obese patients compared to healthy and overweight BMI patients. We found that obesity differentially affected cancer pathways, hub genes, prognostic patient survival, and predicted drug sensitivity in a molecular subtype specific manner. Our findings are consistent with mechanistic findings indicating that there are multiple mechanisms linking obesity and colon cancer (17, 18, 38). Our findings are significant because they suggest that the CMS of the tumor is an important factor in the link between obesity and colon cancer.

CMS1 tumors are categorized as the immune subtype based on the strong infiltration of immune cells and microsatellite instability (MSI) (5, 6). Our findings that immune-related Hallmark gene sets were enriched in CMS1 tumors in all comparisons with obese patients suggests that obesity enhances the immune phenotype of this subtype. In support of the GSEA findings, we observed that an immune gene hub was identified in CMS1 tumors from obese compared to overweight BMI patients. These findings are consistent with the pathophysiological state of obesity being associated with systemic inflammation (15, 16). Further, we observed immune-related Hallmark gene set enrichment in CMS2 and CMS4 tumors in the obese to normal patient BMI comparison but not in the obese to overweight patient BMI comparison. Consistent with this finding, immune-related Hallmark gene set enrichment in CMS2 and CMS4 tumors was observed in the overweight to normal patient BMI comparison. These findings suggest a differential

role of inflammation in the subtypes of colon cancer tumors based on whether the patient is overweight or obese. The commonality of immune-related Hallmark gene set enrichment in 3 of the 4 subtypes is important because an inflammatory risk score is an independent predictor for stage II colon cancer prognosis (68) and a circulating inflammation signature is a strong prognostic factor of progression-free and overall survival of patients with metastatic CRC (29). Consistent with these findings, a higher dietary inflammatory potential is associated with higher CRC risk (69).

The deregulation of cellular energetics resulting in the reprogramming of energy metabolism plays a role in tumorigenesis (70) and has been identified as an emerging hallmark of cancer (71). Our observation that CMS-specific enrichment in the MTORC signaling gene set and the Myc gene sets Myc targets V1 and Myc targets V2 suggests that key cell signaling and metabolic pathways within tumor cells driving tumor growth and progression are differentially regulated by obesity. It has been hypothesized that obesity-derived factors (e.g. circulating hormones, adipokines, inflammatory cytokines, and dietary factors) converge on these key cell signaling and metabolic pathways (38). The enrichment of metabolism related gene sets was concentrated in CMS4 tumors. Consistent with this finding, we also observed that an obesity-linked network of hub genes related to the *UGT1A* gene locus was downregulated in CMS4 tumors. In agreement with this observation, we found that *UGT1A1*, *UGT1A6*, and *UGT1A8* expression was reduced in colon adenocarcinoma compared to normal colon tissue in 3 separate large CRC cohorts. The UGT1 subfamily of enzymes reduces the biological activity and enhances the solubility of lipophilic substrates through the process of glucuronidation (72). UGT activity has been hypothesized to modulate energy metabolism by altering cellular pools UDP-sugars which are glycolytic intermediates or through interaction with pyruvate kinase (PKM2), a glycolytic enzyme (72, 73).

Enrichment of EMT-related Hallmark gene sets and metabolism-related gene sets was observed in CMS4 tumors in the obese to overweight patient BMI comparison. These findings suggest that obesity enhances the mesenchymal features CMS4 subtype and induces a metabolic phenotype which is typically associated with CMS3 tumors (5, 6). It has been proposed that a metabolic shift in the canonical CMS2 tumors possibly due to KRAS mutations and copy number events results in CMS3 tumors whereas the stromal-enriched inflamed tumor microenvironment is the driver for the development of CMS4 tumors from the CMS2 subtype (5). The obesity-induced enrichment of EMT- and metabolism-related Hallmark gene sets that we observed in CMS4 tumors suggests that obesity may modulate the derivation of CMS3 and CMS4 tumors from canonical CMS2 tumors. Consistent with this proposition, we observed a greater proportion of CMS3 tumors in obese compared to normal BMI patients. This finding is consistent with a report (74) that patients with CMS3 tumors in a Stage II-IV CRC cohort are more likely (OR 3.5, 95% CI 1.1-11.4) to have type 2 diabetes, an obesity-linked disease.

A hallmark feature of the mesenchymal CMS4 subtype is complement activation (5, 6) with a platelet signature (75). Indeed, we observed that complement activation was an obesity-enriched gene set in CMS4 tumors and that the platelet marker PPBP was found to be an obesity-linked hub gene in CMS4 tumors. We also observed complement activation enrichment and *PPBP* as a hub gene in CMS2 tumors in the overweight to normal comparison, suggesting a platelet signature in CMS2 tumors from overweight patients. Platelets help initiate and coordinate the immune response including resolution of inflammation (76, 77). Visceral obesity is associated with persistent platelet activation (78). It has been observed that platelet to lymphocyte ratio (PLR) is higher in CRC patients with the Metabolic syndrome and that PRL is associated with poorer overall survival (79). Further, it has recently been reported that microparticles released from thrombin-activated platelets obtained from obese women induce the expression of EMT and EndMT marker genes when incubated with human colon (HT29) cancer cells (80), suggesting that activated platelets can modulate colon cancer progression.

We found that obese patients with CMS1 tumors had worse prognostic relapse-free survival compared to overweight patients. Our finding is consistent the previously reported poor survival rate after relapse in

patients with CMS1 tumors (6), and the reported stronger association between BMI and MSI-high CRC and microsatellite-stable CRC (18). We also observed that obese patients with CMS4 tumors had worse prognostic overall survival compared to normal BMI patients. Our findings with CMS4 tumors are consistent with the worse patient overall survival has been reported for CRC patients with CMS4 tumors (6) and a report that a fibroblast-like and elevated myeloid signature is correlated with poor patient survival (81). Interestingly, we observed that prognostic overall and relapse-free survival was worse in obese patients with CMS3 tumors while relapse-free survival was also worse in obese patients with CMS2 tumors. Taken together our findings suggest that obesity may contribute to worse survival beyond that previously observed in patients with CMS1 and CMS4 tumors. Our findings on worse prognostic survival in patients with CMS4 tumors was not restricted to obese patients; we observed that overall and relapse-free survival was worse in overweight patients too. There is strong epidemiological evidence that obesity is associated with CRC risk (82). Yet, it has been reported that an obesity paradox exists for CRC where being overweight is associated with improved survival (83). However, a recent meta-analysis has reported that CRC recurrence is increased by 33% in overweight compared to normal BMI ($p < 0.001$) (84). Indeed, it has been suggested that methodological problems in studies using BMI may affect findings and interpretation of those findings (82).

Counter intuitively, we observed that obesity was associated primarily with increased predicted sensitivity with GDSC drugs including those targeting DNA replication and MTOR, but not metabolism pathways which indicates that not all CRC pathways were equally affected. Increased predicted sensitivity was concentrated in the CMS1, the MSI immune subtype. Our observation that inflammation-related gene sets are enriched in CMS1 tumors from obese patients is consistent with the finding that immune infiltration predicts fluoropyrimidine, a DNA replication-based chemotherapy (85). Our findings of differential CMS-dependent differences in predicted drug sensitivity are consistent with the observations that differential response to irinotecan-based compared to oxaliplatin-based chemotherapy (86) and chemotherapy plus bevacizumab compared cetuximab (7, 87) in metastatic CRC clinical trials are CMS-dependent. Even though it has been suggested that the current CMS classification does not provide a rationale for targeted therapy in metastatic CRC (88), our findings suggest that obesity-mediated changes in tumor biological pathways may inform drug discovery and rational combination therapies. In contrast to our finding on predicted drug sensitivity in the obese to normal comparison, differential predicted drug sensitivity in the overweight to normal comparison was concentrated in CMS2, the canonical pathways subtype. Consistent with the observation, we observed significantly different predicted drug sensitivity in EGFR and WNT signaling, two CRC canonical pathways.

A limitation of the current study is that it was only performed in TCGA-COAD cohort which lacked racial diversity. CMS categorization of CRC patients across the BMI categories led to small samples sizes which lowered statistical power for some of the comparisons particularly those with normal BMI patients with CMS3 tumors. Confirmation of our findings will require assembly of a large CRC cohort that contains both transcriptomic and body weight and height data. Another limitation of the current study is that confounding factors that are risk factors for CRC such as smoking and alcohol consumption were not assessed. Finally, it should be noted that BMI is a proxy for adiposity but may not account for the metabolic health of the patients.

Conclusions

We conclude that obesity has CMS-specific effects in colon cancer. This conclusion is based not only on CMS-specific effects of obesity in gene set enrichment but also on findings of obesity-related DEGs and the identification of unique hub genes for each CMS in tumors from obese patients. Prognostic patient survival analysis and predicted drug sensitivity support our findings that obesity has CMS-specific effects in colon cancer. Taken together, our findings are consistent with the hypothesis that the obesity-cancer link is mediated by obesity-derived factors which converge on key cell signaling and metabolic pathways; yet this occurs in a CMS-specific manner in colon cancer.

Acknowledgements
The results shown here are in whole or part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

References

1. Shaukat A, Mongin SJ, Geisser MS, Lederle FA, Bond JH, Mandel JS, et al. Long-term mortality after screening for colorectal cancer. New England Journal of Medicine. 2013;369(12):1106-14.
2. Control CfD, Prevention. Behavioral Risk Factor Surveillance System Survey Data. Overview BRFSS 2015. 2015.
3. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021;71(1):7-33.
4. Marley AR, Nan H. Epidemiology of colorectal cancer. International journal of molecular epidemiology and genetics. 2016;7(3):105.
5. Dienstmann R, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. Nature reviews. 2017;17(2):79-92.
6. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21(11):1350-6.
7. Stintzing S, Wirapati P, Lenz HJ, Neureiter D, Fischer von Weikersthal L, Decker T, et al. Consensus molecular subgroups (CMS) of colorectal cancer (CRC) and first-line efficacy of FOLFIRI plus cetuximab or bevacizumab in the FIRE3 (AIO KRK-0306) trial. Ann Oncol. 2019;30(11):1796-803.
8. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.
9. Aleksandrova K, Boeing H, Jenab M, Bas Bueno-de-Mesquita H, Jansen E, van Duijnhoven FJ, et al. Metabolic syndrome and risks of colon and rectal cancer: the European prospective investigation into cancer and nutrition study. Cancer prevention research. 2011;4(11):1873-83.
10. Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nature reviews. 2004;4(8):579-91.
11. Murphy N, Cross AJ, Abubakar M, Jenab M, Aleksandrova K, Boutron-Ruault M-C, et al. A nested case–control study of metabolically defined body size phenotypes and risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC). PLoS medicine. 2016;13(4).
12. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet. 2008;371(9612):569-78.
13. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and Trends in Diabetes Among Adults in the United States, 1988-2012. JAMA. 2015;314(10):1021-9.
14. Ogden CL, Carroll MD, Fryar CD, Flegal KM. Prevalence of obesity among adults and youth: United States, 2011–2014. NCHS data brief. 2015;219(219):1-8.
15. Johnson AM, Olefsky JM. The origins and drivers of insulin resistance. Cell. 2013;152(4):673-84.
16. Samuel VT, Shulman GI. Mechanisms for insulin resistance: common threads and missing links. Cell. 2012;148(5):852-71.
17. Font-Burgada J, Sun B, Karin M. Obesity and Cancer: The Oil that Feeds the Flame. Cell Metab. 2016;23(1):48-62.
18. Iyengar NM, Gucalp A, Dannenberg AJ, Hudis CA. Obesity and Cancer Mechanisms: Tumor Microenvironment and Inflammation. J Clin Oncol. 2016;34(35):4270-6.

19. Khandekar MJ, Cohen P, Spiegelman BM. Molecular mechanisms of cancer development in obesity. Nature reviews. 2011;11(12):886-95.
20. Schulz MD, Atay C, Heringer J, Romrig FK, Schwitalla S, Aydin B, et al. High-fat-diet-mediated dysbiosis promotes intestinal carcinogenesis independently of obesity. Nature. 2014;514(7523):508-12.
21. Terzic J, Grivennikov S, Karin E, Karin M. Inflammation and colon cancer. Gastroenterology. 2010;138(6):2101-14 e5.
22. Holowatyj AN, Haffa M, Lin T, Scherer D, Gigic B, Ose J, et al. Multi-omics Analysis Reveals Adipose-tumor Crosstalk in Patients with Colorectal Cancer. Cancer prevention research. 2020;13(10):817-28.
23. Zhang Z, Scherer PE. The dysfunctional adipocyte—a cancer cell's best friend. Nature Reviews Endocrinology. 2018;14(3):132-4.
24. Aleman JO, Eusebi LH, Ricciardiello L, Patidar K, Sanyal AJ, Holt PR. Mechanisms of obesity-induced gastrointestinal neoplasia. Gastroenterology. 2014;146(2):357-73.
25. Catalan V, Gomez-Ambrosi J, Rodriguez A, Ramirez B, Ortega VA, Hernandez-Lizoain JL, et al. IL-32alpha-induced inflammation constitutes a link between obesity and colon cancer. Oncoimmunology. 2017;6(7):e1328338.
26. Chung YC, Chang YF. Serum interleukin-6 levels reflect the disease status of colorectal cancer. J Surg Oncol. 2003;83(4):222-6.
27. Izano M, Wei EK, Tai C, Swede H, Gregorich S, Harris TB, et al. Chronic inflammation and risk of colorectal and other obesity-related cancers: The health, aging and body composition study. Int J Cancer. 2016;138(5):1118-28.
28. Kim S, Keku TO, Martin C, Galanko J, Woosley JT, Schroeder JC, et al. Circulating levels of inflammatory cytokines and risk of colorectal adenomas. Cancer Res. 2008;68(1):323-8.
29. Varkaris A, Katsiampoura A, Davis JS, Shah N, Lam M, Frias RL, et al. Circulating inflammation signature predicts overall survival and relapse-free survival in metastatic colorectal cancer. British journal of cancer. 2019;120(3):340-5.
30. Goncalves MD, Hopkins BD, Cantley LC. Dietary fat and sugar in promoting cancer development and progression. Annual Review of Cancer Biology. 2019;3:255-73.
31. Matsui S, Okabayashi K, Tsuruta M, Shigeta K, Seishima R, Ishida T, et al. Interleukin-13 and its signaling pathway is associated with obesity-related colorectal tumorigenesis. Cancer science. 2019;110(7):2156.
32. O'Neill AM, Burrington CM, Gillaspie EA, Lynch DT, Horsman MJ, Greene MW. High-fat Western diet-induced obesity contributes to increased tumor growth in mouse models of human colon cancer. Nutrition research. 2016;36(12):1325-34.
33. O'Neill AM, Gillaspie EA, Burrington CM, Lynch DT, Dauchy RT, Blask DE, et al. Development and Characterization of a Novel Congenic Rat Strain for Obesity and Cancer Research. Nutrition and cancer. 2018;70(2):278-87.
34. Olivo-Marston SE, Hursting SD, Perkins SN, Schetter A, Khan M, Croce C, et al. Effects of calorie restriction and diet-induced obesity on murine colon carcinogenesis, growth and inflammatory factors, and microRNA expression. PLoS One. 2014;9(4):e94765.
35. Rabin-Court A, Rodrigues MR, Zhang X-M, Perry RJ. Obesity-associated, but not obesity-independent, tumors respond to insulin by increasing mitochondrial glucose oxidation. PloS one. 2019;14(6).
36. Wang Y, Nasiri AR, Damsky WE, Perry CJ, Zhang XM, Rabin-Court A, et al. Uncoupling Hepatic Oxidative Phosphorylation Reduces Tumor Growth in Two Murine Models of Colon Cancer. Cell Rep. 2018;24(1):47-55.
37. Wunderlich CM, Ackermann PJ, Ostermann AL, Adams-Quack P, Vogt MC, Tran ML, et al. Obesity exacerbates colitis-associated cancer via IL-6-regulated macrophage polarisation and CCL-20/CCR-6-mediated lymphocyte recruitment. Nat Commun. 2018;9(1):1646.

38. Doerstling SS, O'Flanagan CH, Hursting SD. Obesity and Cancer Metabolism: A Perspective on Interacting Tumor–Intrinsic and Extrinsic Factors. Frontiers in Oncology. 2017;7:216.
39. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic acids research. 2016;44(8):e71-e.
40. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, et al. Toward a shared vision for cancer genomic data. New England Journal of Medicine. 2016;375(12):1109-12.
41. Rosenberg R, Engel J, Bruns C, Heitland W, Hermes N, Jauch K-W, et al. The prognostic value of lymph node ratio in a population-based collective of colorectal cancer patients. Annals of surgery. 2010;251(6):1070-8.
42. Eide PW, Bruun J, Lothe RA, Sveen A. CMScaller: an R package for consensus molecular subtyping of colorectal cancer pre-clinical models. Sci Rep. 2017;7(1):16618.
43. Hoshida Y. Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. PloS one. 2010;5(11):e15543.
44. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
45. Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nature Genetics. 2003;34(3):267-73.
46. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.
47. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417-25.
48. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. Nucleic acids research. 2012;41(D1):D808-D15.
49. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics. 2011;27(3):431-2.
50. Chin C-H, Chen S-H, Wu H-H, Ho C-W, Ko M-T, Lin C-Y. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Systems Biology. 2014;8(4):S11.
51. Park S-J, Yoon B-H, Kim S-K, Kim S-Y. GENT2: an updated gene expression database for normal and tumor tissues. BMC medical genomics. 2019;12(5):1-8.
52. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, et al. ONCOMINE: A Cancer Microarray Database and Integrated Data-Mining Platform. Neoplasia. 2004;6(1):1-6.
53. Hong Y, Downey T, Eu KW, Koh PK, Cheah PY. A 'metastasis-prone'signature for early-stage mismatch-repair proficient sporadic colorectal cancer patients and its implications for possible therapeutics. Clinical & experimental metastasis. 2010;27(2):83-90.
54. Skrzypczak M, Goryca K, Rubel T, Paziewska A, Mikula M, Jarosz D, et al. Modeling oncogenic signaling in colon tumors by multidirectional analyses of microarray data directed for maximization of analytical reliability. PloS one. 2010;5(10):e13091.
55. Kaiser S, Park Y-K, Franklin JL, Halberg RB, Yu M, Jessen WJ, et al. Transcriptional recapitulation and subversion of embryonic colon development by mouse colon tumor models and human colon cancer. Genome biology. 2007;8(7):1-26.
56. Goswami CP, Nakshatri H. PROGgene: gene expression based survival analysis web application for multiple cancers. Journal of clinical bioinformatics. 2013;3(1):1-9.
57. Goswami CP, Nakshatri H. PROGgeneV2: enhancements on the existing database. BMC cancer. 2014;14(1):1-6.
58. Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, et al. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. Gastroenterology. 2010;138(3):958-68.

59. Sheffer M, Bacolod MD, Zuk O, Giardina SF, Pincas H, Barany F, et al. Association of survival and disease progression with chromosomal instability: a genomic exploration of colorectal cancer. Proceedings of the National Academy of Sciences. 2009;106(17):7131-6.
60. Jorissen RN, Gibbs P, Christie M, Prakash S, Lipton L, Desai J, et al. Metastasis-associated gene expression changes predict poor outcomes in patients with Dukes stage B and C colorectal cancer. Clinical Cancer Research. 2009;15(24):7642-51.
61. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PloS one. 2014;9(9):e107468.
62. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic acids research. 2012;41(D1):D955-D61.
63. Geeleher P, Cox NJ, Huang RS. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. Genome biology. 2014;15(3):1-12.
64. Adam R, Haller DG, Poston G, Raoul JL, Spano JP, Tabernero J, et al. Toward optimized front-line therapeutic strategies in patients with metastatic colorectal cancer—an expert review from the International Congress on Anti-Cancer Treatment (ICACT) 2009. Annals of Oncology. 2010;21(8):1579-84.
65. Xie Y-H, Chen Y-X, Fang J-Y. Comprehensive review of targeted therapy for colorectal cancer. Signal Transduction and Targeted Therapy. 2020;5(1):22.
66. Novellasdemunt L, Antas P, Li VS. Targeting Wnt signaling in colorectal cancer. A review in the theme: cell signaling: proteins, pathways and mechanisms. American Journal of Physiology-Cell Physiology. 2015;309(8):C511-C21.
67. Saif MW, Chu E. Biology of colorectal cancer. The Cancer Journal. 2010;16(3):196-201.
68. Schetter AJ, Nguyen GH, Bowman ED, Mathé EA, Yuen ST, Hawkes JE, et al. Association of inflammation-related and microRNA gene expression with cancer-specific mortality of colon adenocarcinoma. Clinical Cancer Research. 2009;15(18):5878-87.
69. Tabung FK, Liu L, Wang W, Fung TT, Wu K, Smith-Warner SA, et al. Association of Dietary Inflammatory Potential With Colorectal Cancer Risk in Men and Women. JAMA Oncology. 2018;4(3):366-73.
70. Pavlova Natalya N, Thompson Craig B. The Emerging Hallmarks of Cancer Metabolism. Cell Metabolism. 2016;23(1):27-47.
71. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. cell. 2011;144(5):646-74.
72. Allain EP, Rouleau M, Lévesque E, Guillemette C. Emerging roles for UDP-glucuronosyltransferases in drug resistance and cancer progression. British Journal of Cancer. 2020;122(9):1277-87.
73. Audet-Delage Y, Rouleau M, Rouleau M, Roberge J, Miard S, Picard F, et al. Cross-talk between alternatively spliced UGT1A isoforms and colon cancer cell metabolism. Molecular pharmacology. 2017;91(3):167-77.
74. Davis JS, Yu R, Banerjee M, Jiang Z-Q, Menter DG, Guinney J, et al. Distinct Patient and Tumor Characteristics of the Consensus Molecular Subtypes of Colorectal Cancer. Gastroenterology. 2017;152(5):S880.
75. Lam M, Roszik J, Kanikarla-Marie P, Davis JS, Morris J, Kopetz S, et al. The potential role of platelets in the consensus molecular subtypes of colorectal cancer. Cancer and Metastasis Reviews. 2017;36(2):273-88.
76. Semple JW, Italiano JE, Freedman J. Platelets and the immune continuum. Nature Reviews Immunology. 2011;11(4):264-74.
77. Margraf A, Zarbock A. Platelets in Inflammation and Resolution. The Journal of Immunology. 2019;203(9):2357-67.
78. Davì G, Guagnano MT, Ciabattoni G, Basili S, Falco A, Marinopiccoli M, et al. Platelet Activation in Obese WomenRole of Inflammation and Oxidant Stress. JAMA. 2002;288(16):2008-14.

79. You J, Zhang H, Shen Y, Chen C, Liu W, Zheng M, et al. Impact of platelet to lymphocyte ratio and metabolic syndrome on the prognosis of colorectal cancer patients. Onco Targets Ther. 2017;10:2199-208.
80. Grande R, Dovizio M, Marcone S, Szklanna PB, Bruno A, Ebhardt HA, et al. Platelet-derived microparticles from obese individuals: characterization of number, size, proteomics, and crosstalk with cancer and endothelial cells. Frontiers in pharmacology. 2019;10:7.
81. Li H, Courtois ET, Sengupta D, Tan Y, Chen KH, Goh JJL, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. Nature genetics. 2017;49(5):708-18.
82. Slawinski C, Barriuso J, Guo H, Renehan AG. Obesity and cancer treatment outcomes: interpreting the complex evidence. Clinical Oncology. 2020;32(9):591-608.
83. Schlesinger S, Siegert S, Koch M, Walter J, Heits N, Hinz S, et al. Postdiagnosis body mass index and risk of mortality in colorectal cancer survivors: a prospective study and meta-analysis. Cancer causes & control. 2014;25(10):1407-18.
84. Jaspan V, Lin K, Popov V. The impact of anthropometric parameters on colorectal cancer prognosis: A systematic review and meta-analysis. Critical Reviews in Oncology/Hematology. 2021;159:103232.
85. Mo X, Huang X, Feng Y, Wei C, Liu H, Ru H, et al. Immune infiltration and immune gene signature predict the response to fluoropyrimidine-based chemotherapy in colorectal cancer patients. Oncoimmunology. 2020;9(1):1832347.
86. Okita A, Takahashi S, Ouchi K, Inoue M, Watanabe M, Endo M, et al. Consensus molecular subtypes classification of colorectal cancer as a predictive factor for chemotherapeutic efficacy against metastatic colorectal cancer. Oncotarget. 2018;9(27):18698.
87. Lenz H-J, Ou F-S, Venook AP, Hochster HS, Niedzwiecki D, Goldberg RM, et al. Impact of consensus molecular subtype on survival in patients with metastatic colorectal cancer: results from CALGB/SWOG 80405 (Alliance). Journal of Clinical Oncology. 2019;37(22):1876.
88. Sveen A, Cremolini C, Dienstmann R. Predictive modeling in colorectal cancer: time to move beyond consensus molecular subtypes. Annals of Oncology. 2019;30(11):1682-5.

Table 1. Patient Demographics and Tumor Characteristics

	Total (n= 231)		Normal† (n= 77)		Overweight† (n= 80)		Obese† (n=74)		
	n	%	n	%	n	%	n	%	P-value
*Sex**									0.157
Female	108	47	37	48	31	39	40	54	
Male	123	53	40	52	49	61	34	46	
*Age**									0.104
30-39	9	4	2	3	6	8	1	1	
40-49	25	11	11	14	8	10	6	8	
50-59	45	19	10	13	15	19	20	27	
60-69	62	27	17	22	19	24	26	35	
70-79	60	26	23	30	23	29	14	19	
80-89	28	12	12	16	9	11	7	9	
90>	2	1	2	3	0	0	0	0	
*Ethnicity**									0.205
Hispanic or Latino	3	1	1	1	1	1	1	0	
Not Hispanic or Latino	221	96	71	92	77	96	73	99	
Not Reported	7	3	5	6	2	2	0	0	
*Race**									0.001
American Indian or Alaska Native	1	0	0	0	0	0	1	1	
Asian	8	3	8	10	0	0	0	0	
Black or African American	51	22	14	18	15	19	22	30	
White	171	74	55	71	65	81	51	69	
*Tumor Location**									0.476
Ascending Colon	41	18	14	18	16	20	11	15	
Cecum	57	25	18	23	19	24	20	27	
Descending Colon	12	5	4	5	2	2	6	8	
Hepatic Flexure	14	6	7	9	4	5	3	4	
Rectosigmoid Junction	1	0	0	0	0	0	1	1	
Sigmoid Colon	62	27	18	23	21	26	23	31	
Splenic Flexure	5	2	0	0	4	5	1	1	
Transverse Colon	23	10	11	14	6	8	6	8	
Not Reported	16	7	5	6	8	10	3	4	
*Tumor Stage**									0.547
Stage I	33	14	11	14	10	12	12	16	
Stage II	94	41	38	49	29	36	27	36	
Stage III	79	34	20	26	32	40	27	36	
Stage IV	25	11	8	19	9	11	8	11	
*Lymph Node Ratio**									0.075
LNR0	123	53	41	53	43	54	39	53	
LNR1	57	25	14	18	22	28	21	18	
LNR2	19	8	5	6	8	10	6	8	
LNR3	10	4	3	4	5	6	2	3	
LNR4	11	5	5	6	2	2	4	5	
Not Reported	11	5	9	12	0	0	2	3	
*Consensus Molecular Subtype**									0.040
CMS1	37	16	17	22	11	14	9	12	
CMS2	54	23	20	26	22	28	12	16	
CMS3	31	13	3	4	12	15	16	22	
CMS4	86	37	30	39	26	32	30	32	
Unassigned	23	10	7	9	9	11	7	11	

* Significance across score categories by Fishers Exact test

† BMI 19-24.9 (Normal), BMI 25-29.9 (Overweight), BMI ≥30 (Obese)

Figure Legends

Figure 1. Transcriptomic Analysis Flow Chart. Colon adenocarcinoma RNA-seq data (HT-Seq counts) and associated tumor and patient characteristics were obtained from The Cancer Genomic Atlas – Colon Adenocarcinoma (TCGA-COAD) database using the R package TCGAbiolinks. Body mass index (BMI) was calculated, and each patient was categorized as normal, overweight, and obese. Patient tumor samples from the TCGA-COAD were assigned to a consensus molecular subtype (CMS) using the R package CMScaller. Raw counts and associated phenotypes were inputted into DESeq2, and the following comparisons were made: obese vs. normal, obese vs. overweight, and overweight vs. normal for each CMS category. Volcano plots were generated using the R package EnhancedVolcano. Gene overlap between comparisons were visualized using Euler diagrams and upset plots from the R packages eulerr and UpsetR, respectively. To examine prognostic patient outcomes (Survival Analysis) DESeq2-obtained DEGs in the BMI comparisons for each CMS category was assessed using the PROGgeneV2 tool and Kaplan-Meier plots were generated to assess overall survival and relapse-free survival. Normalized RNA-seq counts obtained from DESeq2 were used for Gene Set Enrichment Analysis (GSEA). DESeq2-obtained DEGs were used to construct a protein-protein interaction (PPI) network from the Search Tool for the Retrieval of Interacting Genes (STRING) using Cytoscape. The Cytohubba package in Cytoscape was used to perform the hub gene analysis. The hub genes were queried for mRNA expression using the aggregate microarray data using GENT2 and in colon cancer cohorts in the Oncomine database. The pRRophetic R package was used with normalized RNA-seq counts from normal, overweight, and obese patient tumors to estimate the half maximal inhibitory concentration (IC50) of drugs from the Cancer Genomic Project.

Figure 2. Differential expressed gene analysis reveals CMS-specific differences between BMI groups. (A) Euler diagrams were used to visualize the weighted overlap of DESeq2-obtained DEGs (MeanBase > 10, FDR p value < 0.05) between Obese vs. Normal, Obese vs. Overweight, and Overweight vs Normal DEGs for each CMS category. The R package eulerr was used to construct the Euler diagrams. (B) Upset plots were used to visualize the intersection of DEGs between the four CMS categories for each BMI comparison. The R package UpsetR was used to construct the Upset plots.

Figure 3. Prognostic patient outcomes reveal CMS-specific differences between obese BMI groups. The expression of the top 20 significantly upregulated DESeq2-obtained DEGs (MeanBase > 10, FDR p value < 0.05) in the obese vs. normal and obese vs. overweight comparisons for each CMS category were assessed in GSE17536 and GSE41258 for overall survival (A) and GSE14333 and GSE17536 for relapse-free survival (B) using the PROGgeneV2 tool. Survival analyses were adjusted for age, stage, and gender covariates and bifurcated based on median expression. The hazard ratios, 95% confidence intervals, and p values were reported for the Kaplan-Meier plots.

Figure 4. Gene set enrichment of Hallmark gene sets reveal CMS-specific differences between obese BMI groups. Normalized RNA-seq counts obtained from DESeq2 were used for Gene Set Enrichment Analysis (GSEA). Hallmark gene sets were assessed in the obese vs. normal (A) and obese vs. overweight (B) comparisons for each CMS category. Bubbles for gene sets with

a false discovery rate q-value less than 0.05 were reported. The size of bubbles represents the normalized enrichment score (NES). The color of bubbles represents false discovery rate q-value (FDR).

Figure 5. Hub gene analysis reveal CMS-specific differences between obese BMI groups. DESeq2-obtained DEGs (MeanBase > 10, FDR p value < 0.05) were used to construct a protein-protein interaction (PPI) network from the STRING database in Cytoscape. The Cytohubba package in Cytoscape was used to perform the hub gene analysis from the PPI network for the obese vs. normal (left panels) and obese vs. overweight (right panels) comparisons for each CMS category. Hub genes were identified using the maximal clique centrality (MCC) method to obtain the top 10 ranked genes in all modules. The intensity and color (high, red; orange, medium; yellow, low) of the hub genes is shown.

Figure 6. Hub gene expression is relevant to colon cancer. Hub genes were queried for mRNA expression using Affymetrix U133Plus2 platform microarray data from normal colon tissue (Normal) and colon cancer (Cancer) in the GEO database and the GENT2 program. Violin plots were generated using ggplot2 in R studio. The median is represented as a horizontal bar in the violin sample plot. The p value from t-tests and the Log2 fold change (Log2FC) are shown. (A) *PYY*, (B) *PPBP*, (C) *INSL5*, (D) *NPW*, and (E) *CCR2*.

Figure 7. Predicted drug sensitivity analysis reveal CMS-specific differences between obese BMI groups. Estimated half maximal inhibitory concentration (IC50) of drugs targeting DNA Replication (A), MTOR (B), and Metabolism (C) for tumors from normal (blue) and obese (red)

BMI patients across the CMS categories is shown. To assess predicted drug sensitivity, the pRRophetic R package was used with normalized RNA-seq counts. Box plots with p values from t-tests are shown.

Figure 1

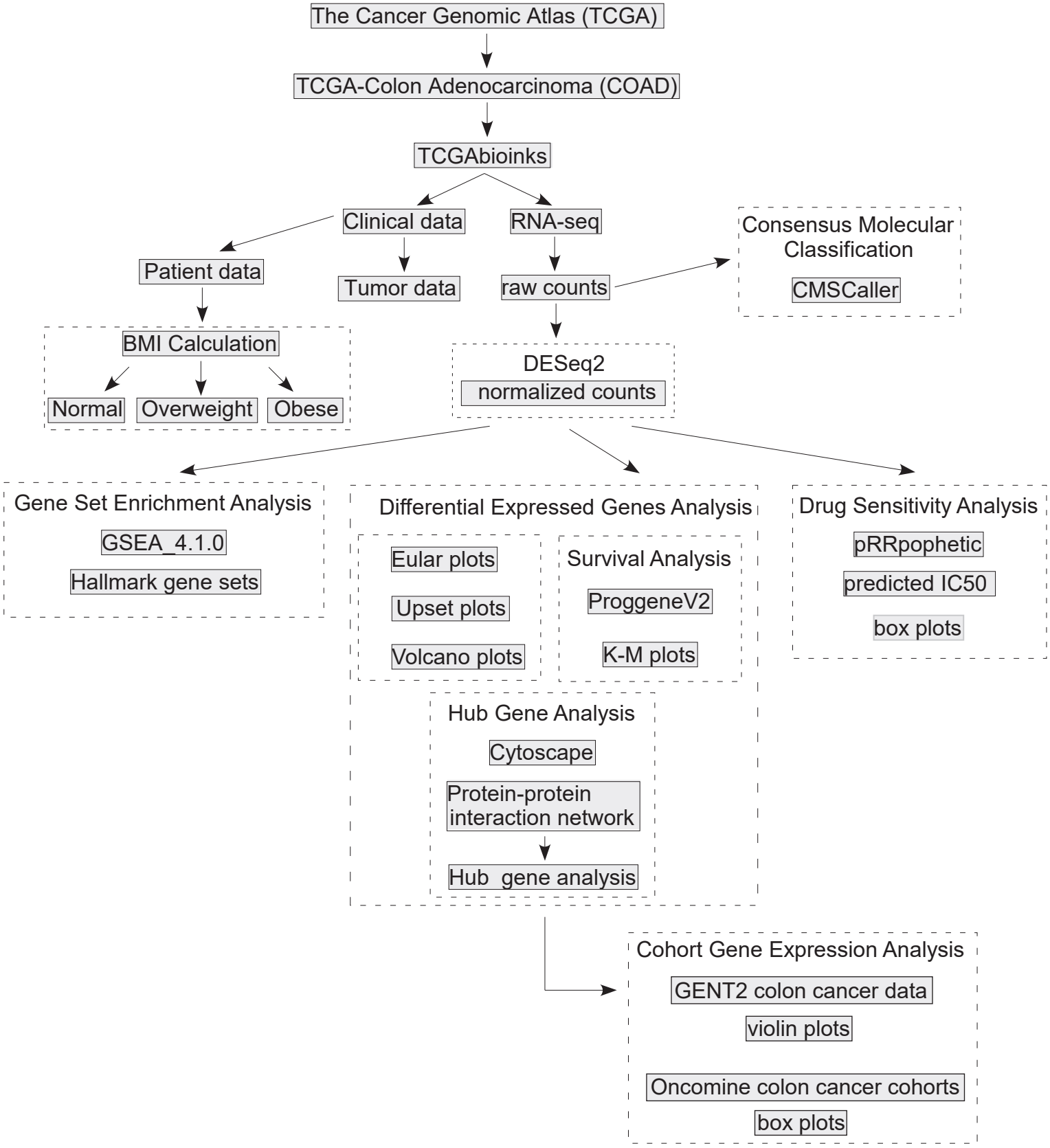


Figure 2

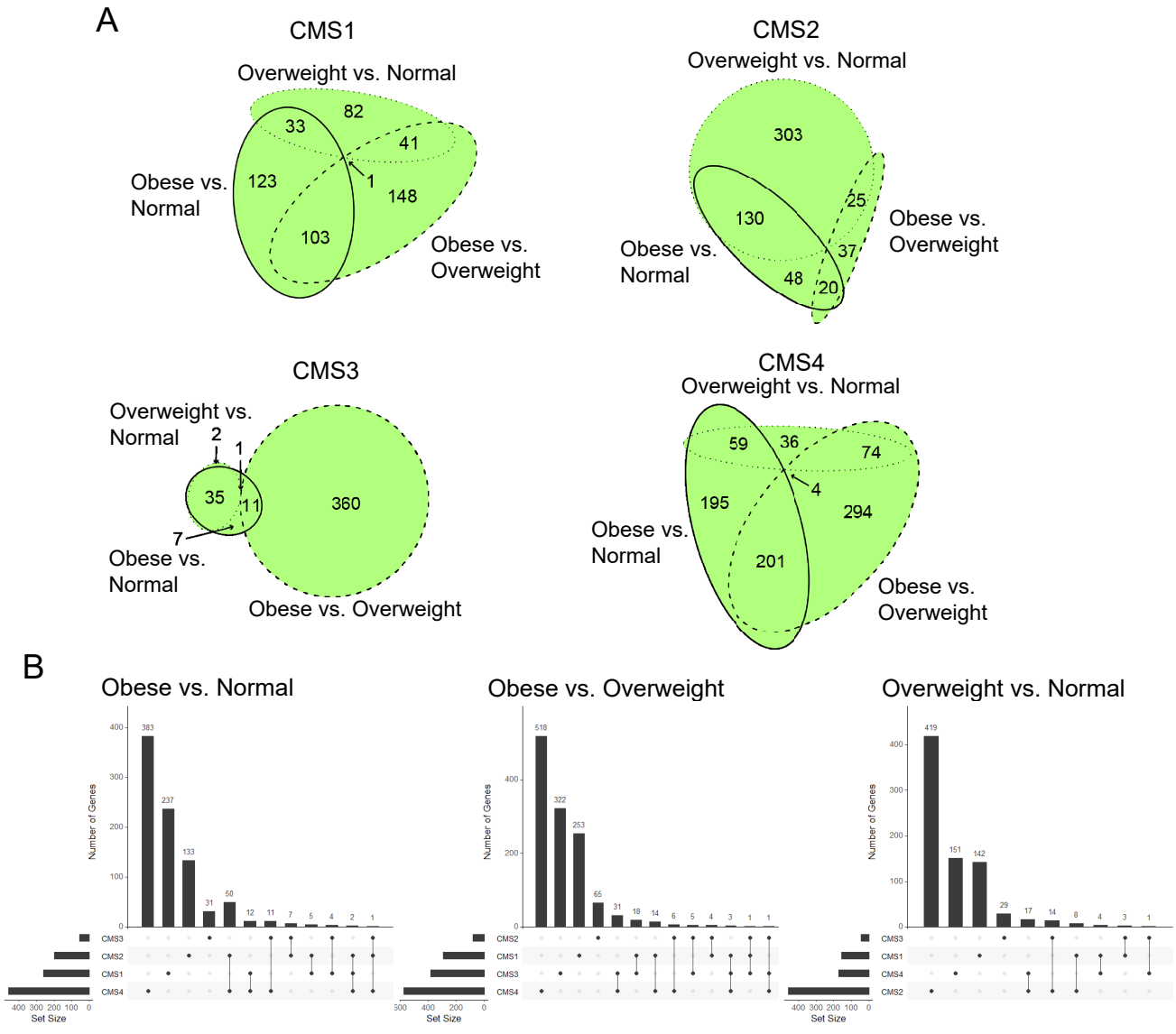


Figure 3

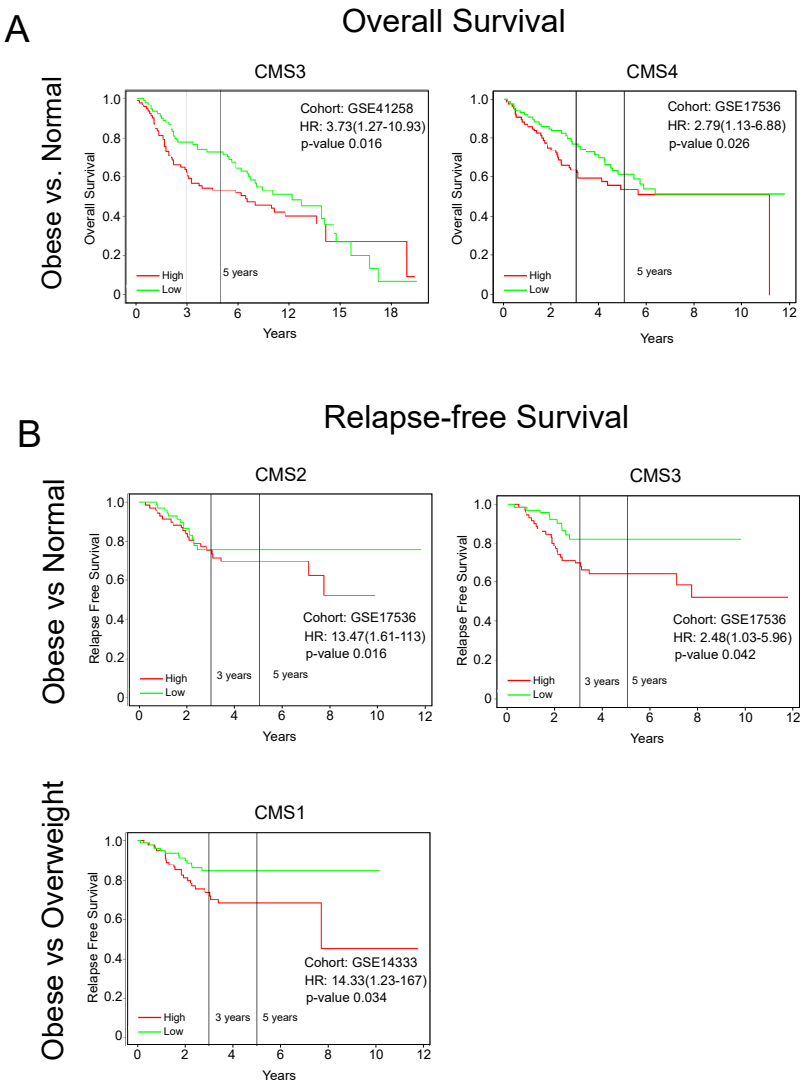
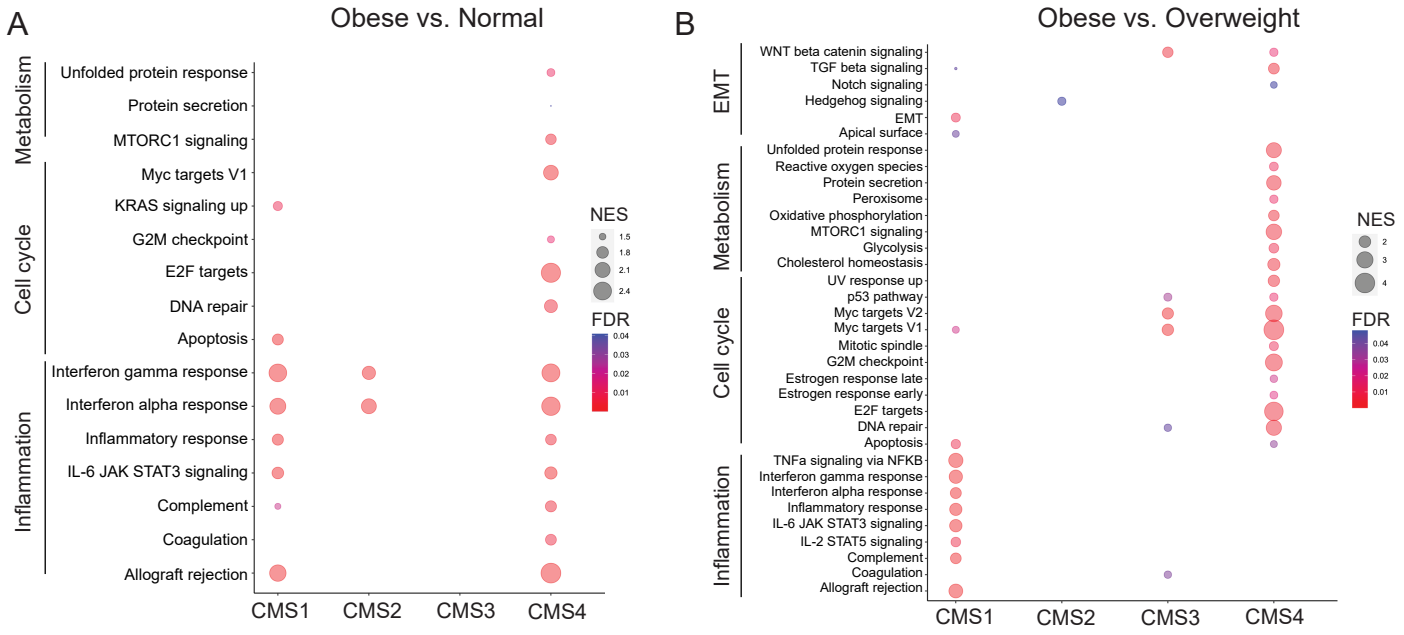


Figure 4



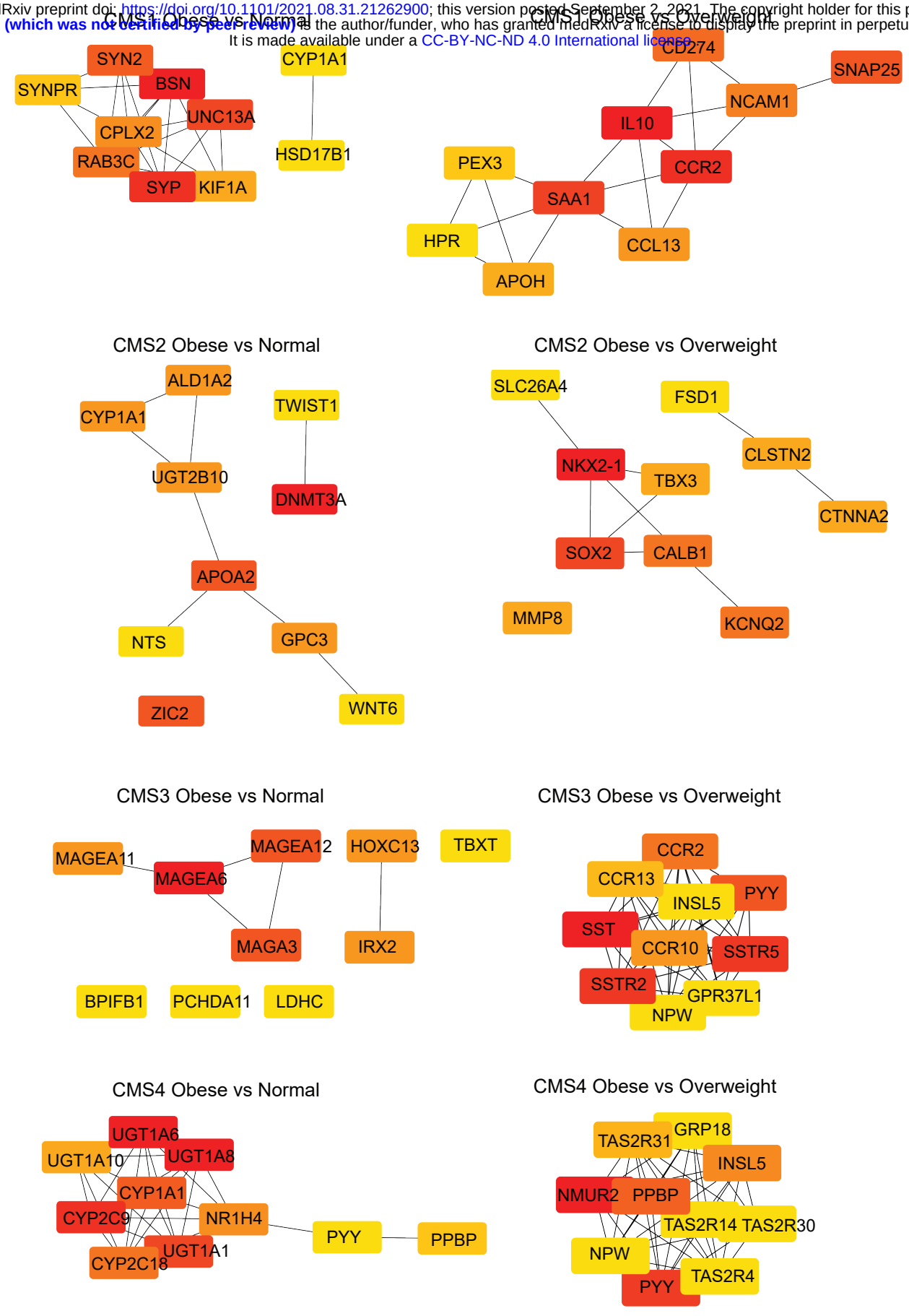


Figure 6

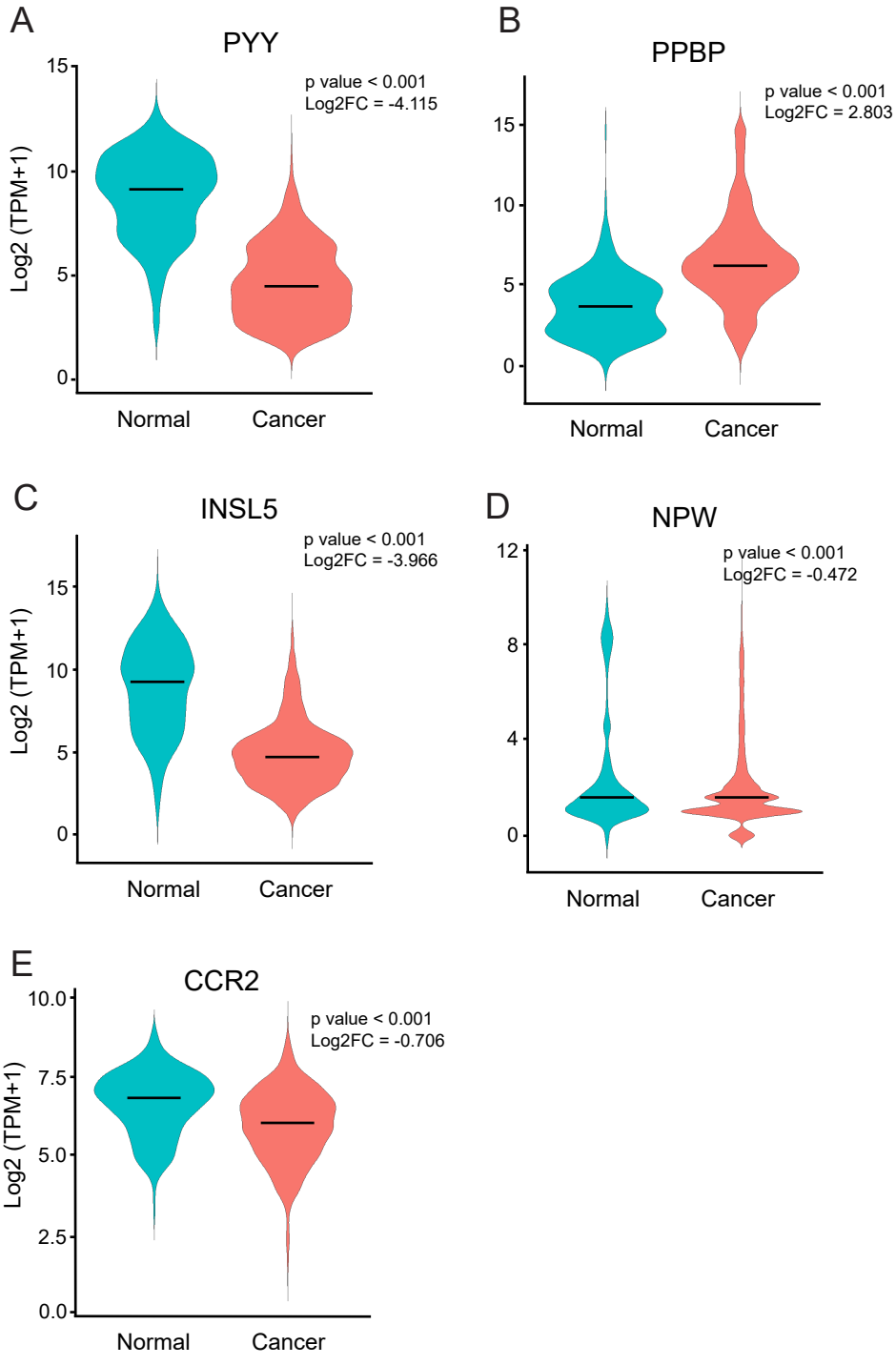
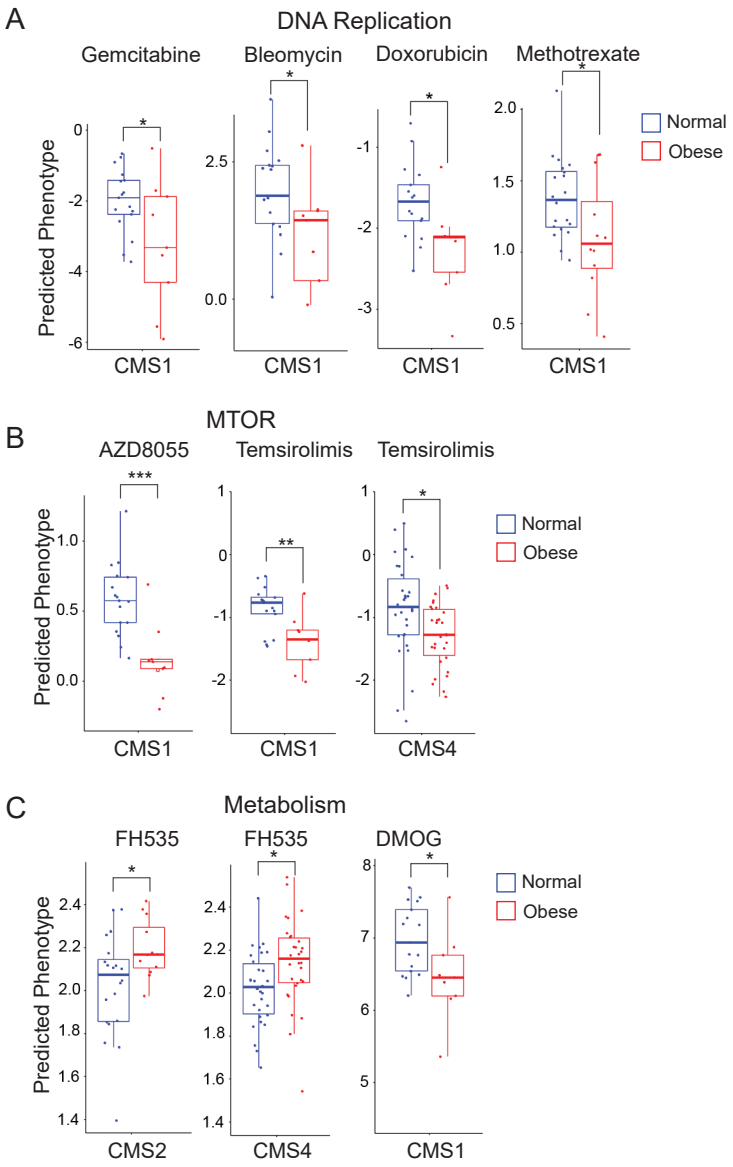


Figure 7



Supplemental Table 1. Hub gene analysis for the obese vs. normal, obese vs. overweight, and overweight vs. normal comparisons for each CMS category.

	hgnc_symbol	log2FC*	Padj$^{\nabla}$	description
Obese vs. Normal†				
CMS1	BSN	2.19	0.007	Bassoon presynaptic cytomatrix protein
	SYP	1.72	0.022	Major synaptic vesicle protein p38
	UNC13A	2.71	0.000	Unc-13 homolog A
	RAB3C	2.99	0.003	RAB3C, member RAS oncogene family
	SYN2	-2.42	0.025	Synapsin II
	CPLX2	4.70	0.000	Complexin II
	KIF1A	4.63	0.000	kinesin family member 1A
	SYNPR	-4.44	0.005	Synaptoporin
	CYP1A1	-5.24	0.035	Cytochrome P450, family 1, subfamily A, polypeptide 1
	HSD17B1	1.02	0.020	Short chain dehydrogenase/reductase family 28C member 1
CMS2	DNMT3A	-5.63	0.000	DNA (cytosine-5-)-methyltransferase 3 alpha
	APOA2	4.78	0.003	Apolipoprotein A-II
	ZIC2	3.24	0.046	Zinc finger protein of the cerebellum 2
	CYP1A1	5.15	0.017	Cytochrome P450, family 1, subfamily A, polypeptide 1
	ALDH1A2	-3.95	0.001	Retinaldehyde-specific dehydrogenase type 2
	UGT2B10	-4.19	0.001	UDP glucuronosyltransferase 2 family, polypeptide B10
	GPC3	-2.09	0.039	Intestinal protein OCI-5
	WNT6	-3.05	0.004	Wingless-type MMTV integration site family, member 6
	NTS	-4.28	0.003	Neurotensin/neuromedin N
	TWIST1	1.46	0.046	Class A basic helix-loop-helix protein 38
CMS3	MAGEA6	23.82	0.000	Melanoma-associated antigen 6
	MAGEA3	24.08	0.000	Melanoma-associated antigen 3
	MAGEA12	20.72	0.000	Melanoma-associated antigen 12
	MAGEA11	22.69	0.000	Melanoma-associated antigen 11
	HOXC13	16.61	0.000	Homeobox protein Hox-C13
	IRX2	7.35	0.012	Iroquois-class homeodomain protein IRX-2
	TBXT	22.27	0.000	T-box transcription factor T
	BPIFB1	18.71	0.000	BPI fold containing family B member 1
	PCDHA11	21.63	0.000	Protocadherin alpha 11
	LDHC	18.07	0.000	L-lactate dehydrogenase C chain
CMS4	UGT1A6	-1.84	0.002	UDP glucuronosyltransferase 1 family, polypeptide A6
	UGT1A8	-2.21	0.003	UDP glucuronosyltransferase 1 family, polypeptide A8
	CYP2C9	-2.14	0.011	Cytochrome P450, family 2, subfamily C, polypeptide 9
	UGT1A1	-3.26	0.000	UDP glucuronosyltransferase 1 family, polypeptide A1
	CYP1A1	-3.71	0.006	Cytochrome P450, family 1, subfamily A, polypeptide 1
	NR1H4	-2.79	0.003	Nuclear receptor subfamily 1, group H, member 4; Isoform 4
	UGT1A10	-1.93	0.002	UDP glucuronosyltransferase 1 family, polypeptide A10

	CYP2C18	-1.58	0.022	Cytochrome P450, family 2, subfamily C, polypeptide 18
	PPBP	5.93	0.000	Pro-platelet basic protein (chemokine (C-X-C motif) ligand 7)
	PYY	-5.12	0.000	Peptide tyrosine tyrosine
Overweight vs. Normal†				
CMS1	TH	-3.00	0.003	tyrosine hydroxylase
	RBFOX1	-5.47	0.000	RNA binding fox-1 homolog 1
	SYN1	-2.08	0.003	synapsin I
	UCHL1	-2.11	0.010	ubiquitin C-terminal hydrolase L1
	CALCA	-4.61	0.001	calcitonin related polypeptide alpha
	SCN2A	-2.58	0.033	sodium voltage-gated channel alpha subunit 2
	NCAM1	2.17	0.009	neural cell adhesion molecule 1
	SYP	2.08	0.001	synaptophysin
	RIMBP2	2.82	0.004	RIMS binding protein 2
	SNAP25	2.75	0.004	synaptosome associated protein 25
CMS2	FBXO2	1.85	0.031	F-box protein 2
	DRD4	-1.00	0.049	dopamine receptor D4
	TRIM71	-3.45	0.014	tripartite motif containing 71
	CXCL11	1.80	0.020	C-X-C motif chemokine ligand 11
	CXCL9	2.06	0.003	C-X-C motif chemokine ligand 9
	FBXW5	-0.81	0.015	F-box and WD repeat domain containing 5
	PPBP	4.53	0.000	pro-platelet basic protein
	SPSB4	-1.51	0.035	splA/ryanodine receptor domain and SOCS box containing 4
	BDKRB1	1.18	0.015	bradykinin receptor B1
	ASB12	-1.45	0.041	ankyrin repeat and SOCS box containing 12
CMS3	HP	-5.35	0.034	haptoglobin
	VTN	-3.74	0.022	vitronectin
	BPIFB1	16.20	0.000	BPI fold containing family B member 1
	FGA	-8.19	0.000	fibrinogen alpha chain
	FGB	-7.64	0.003	fibrinogen beta chain
	PAX7	-11.17	0.023	paired box 7
	MAGEA11	20.54	0.000	MAGE family member A11
	MAGEA12	20.33	0.000	MAGE family member A12
	MAGEA3	23.84	0.000	MAGE family member A3
	MAGEA6	23.68	0.000	MAGE family member A6
CMS4	ANPEP	-2.07	0.025	alanyl aminopeptidase, membrane
	ALPI	-3.21	0.000	alkaline phosphatase, intestinal
	TDRD9	1.55	0.012	tudor domain containing 9
	ABCG5	-3.33	0.000	ATP binding cassette subfamily G member 5
	TEX101	2.73	0.005	testis expressed 101
	DSG1	2.87	0.000	desmoglein 1
	SYCP2	-1.85	0.009	synaptonemal complex protein 2

	TDRD1	2.34	0.000	tudor domain containing 1
	ABCG8	-2.61	0.007	ATP binding cassette subfamily G member 8
	MAEL	2.30	0.047	maelstrom spermatogenic transposon silencer
Obese vs. Overweight†				
CMS1	APOH	-3.27	0.009	apolipoprotein H
	SAA1	4.11	0.002	serum amyloid A1
	HPR	-6.39	0.003	haptoglobin-related protein
	IL10	2.82	0.000	interleukin 10
	CCR2	2.29	0.025	C-C motif chemokine receptor 2
	NCAM1	-2.36	0.017	neural cell adhesion molecule 1
	CCL13	2.71	0.033	C-C motif chemokine ligand 13
	SNAP25	-2.88	0.013	synaptosome associated protein 25
	PEX3	0.85	0.019	peroxisomal biogenesis factor 3
	CD274	2.22	0.004	CD274 molecule
CMS2	CALB1	-4.22	0.032	calbindin 1
	SOX2	4.27	0.017	SRY-box transcription factor 2
	FSD1	-1.59	0.021	fibronectin type III and SPRY domain containing 1
	CLSTN2	2.86	0.001	calsyntenin 2
	SLC26A4	-2.43	0.001	solute carrier family 26 member 4
	NKX2-1	-8.82	0.000	NK2 homeobox 1
	KCNQ2	2.21	0.020	kinase non-catalytic C-lobe domain containing 1
	CTNNA2	3.84	0.006	catenin alpha 2
	TBX3	2.07	0.000	T-box transcription factor 3
	MMP8	7.18	0.000	matrix metallopeptidase 8
CMS3	GPR37L1	-1.41	0.010	G protein-coupled receptor 37 like 1
	PYY	-6.71	0.000	peptide YY
	INSL5	-5.47	0.008	insulin like 5
	SST	-3.85	0.017	somatostatin
	SSTR2	-2.58	0.000	somatostatin receptor 2
	SSTR5	3.67	0.005	somatostatin receptor 5
	CCR2	-2.09	0.011	C-C motif chemokine receptor 2
	CCL13	-2.50	0.013	C-C motif chemokine ligand 13
	NPW	2.19	0.049	neuropeptide W
	CCR10	-1.62	0.022	C-C motif chemokine receptor 10
CMS4	PYY	-4.25	0.000	peptide YY
	GPR18	-2.19	0.000	G protein-coupled receptor 18
	INSL5	-3.40	0.020	insulin like 5
	NPW	1.72	0.010	neuropeptide W
	PPBP	2.32	0.006	pro-platelet basic protein
	NMUR2	3.55	0.004	neuromedin U receptor 2
	TAS2R4	-1.37	0.007	taste 2 receptor member 4

TAS2R14	-1.16	0.003	taste 2 receptor member 14
TAS2R30	-4.47	0.001	taste 2 receptor member 30
TAS2R31	-2.89	0.000	taste 2 receptor member 31

* Log2 fold change (Log2FC) from the DESeq2-obtained DEGs

∇ False discovery rate P value (Padj) from the DESeq2-obtained DEGs

† BMI 19-24.9 (Normal), BMI 25-29.9 (Overweight), BMI ≥30 (Obese)

Supplementary Table 2. Predicted drug sensitivity for normal compared to obese BMI categories.

Name	Targets	Target pathway	log(IC50) Normal - log(IC50) Obese	p-value	CMS
GNF-2	BCR-ABL	ABL signaling	0.058	0.026	CMS4
Obatoclax Mesylate	BCL2, BCL-XL, BCL-W, MCL1	Apoptosis regulation	0.698	0.013	CMS1
Palbociclib	CDK4, CDK6	Cell cycle	0.356	0.017	CMS1
CGP-082996	CDK4	Cell cycle	0.176	0.021	CMS2
RO-3306	CDK1	Cell cycle	0.132	0.036	CMS4
Parthenolide	HDAC1	Chromatin histone acetylation	0.118	0.039	CMS4
IPA-3	PAK1	Cytoskeleton	0.331	0.023	CMS1
PF-562271	FAK, FAK2	Cytoskeleton	0.244	0.022	CMS1
Gemcitabine	Pyrimidine antimetabolite	DNA replication	1.232	0.028	CMS1
Bleomycin	dsDNA break induction	DNA replication	1.092	0.024	CMS1
Doxorubicin	Anthracycline	DNA replication	0.575	0.010	CMS1
Methotrexate	Antimetabolite	DNA replication	0.281	0.030	CMS2
Lapatinib	EGFR, ERBB2	EGFR signaling	0.161	0.024	CMS4
PLX-4720	BRAF	ERK MAPK signaling	0.311	0.041	CMS1
KU-55933	ATM	Genome integrity	-0.032	0.008	CMS1
JNK Inhibitor VIII	JNK	JNK and p38 signaling	0.053	0.031	CMS4
DMOG	HIF-PH	Metabolism	0.503	0.027	CMS1
FH535	PPARgamma, PPARdelta	Metabolism	-0.168	0.034	CMS2
FH535	PPARgamma, PPARdelta	Metabolism	-0.123	0.014	CMS4
ZM447439	AURKA, AURKB	Mitosis	0.324	0.034	CMS1
Vinorelbine	Microtubule destabiliser	Mitosis	0.086	0.039	CMS1
Thapsigargin	SERCA	Other	0.909	0.031	CMS1
Tipifarnib	Farnesyl-transferase (FNTA)	Other	0.427	0.027	CMS1
Midostaurin	PKC, PPK, FLT1, c-FGR, others	Other	0.396	0.021	CMS1
Shikonin	not defined	Other	0.050	0.014	CMS1
Bexarotene	Retinioic X receptor (RXR) agonist	Other	-0.156	0.011	CMS3
Ponatinib	ABL, PDGFRA, VEGFR2, FGFR1, SRC, TIE2, FLT3	Other, kinases	0.355	0.049	CMS1
Dasatinib	ABL, SRC, Ephrins, PDGFR, KIT	Other, kinases	0.463	0.046	CMS4
BMS-509744	ITK	Other, kinases	0.146	0.015	CMS4

Bosutinib	SRC, ABL, TEC	Other, kinases	0.057	0.040	CMS4
Serdemetan	MDM2	p53 pathway	0.349	0.014	CMS1
Temsirolimus	MTOR	PI3K/MTOR signaling	0.519	0.003	CMS1
AZD8055	MTORC1, MTORC2	PI3K/MTOR signaling	0.430	0.000	CMS1
Dactolisib	PI3K (class 1), MTORC1, MTORC2	PI3K/MTOR signaling	0.390	0.042	CMS1
Temsirolimus	MTOR	PI3K/MTOR signaling	0.427	0.013	CMS4
Elesclomol	HSP90	Protein stability and degradation	0.615	0.021	CMS1

Supplementary Table 3. Predicted drug sensitivity for normal compared to overweight BMI categories.

Name	Targets	Target pathway	log(IC50) Normal - log(IC50) Obese	p-value	CMS
RO-3306	CDK1	Cell cycle	0.2360	0.002	CMS2
CGP-082996	CDK4	Cell cycle	0.1739	0.020	CMS2
Camptothecin	TOP1	DNA replication	0.5147	0.041	CMS2
Methotrexate	Antimetabolite	DNA replication	0.2221	0.017	CMS2
Afatinib	ERBB2, EGFR	EGFR signaling	0.2826	0.012	CMS2
Gefitinib	EGFR	EGFR signaling	0.0796	0.000	CMS2
KU-55933	ATM	Genome integrity	0.2150	0.001	CMS2
Bicalutamide	AR	Hormone-related	-0.0598	0.001	CMS2
JNK Inhibitor VIII	JNK	JNK and p38 signaling	0.0788	0.004	CMS2
Doramapimod	p38, JNK2	JNK and p38 signaling	0.0640	0.014	CMS2
AS601245	JNK1, JNK2, JNK2	JNK and p38 signaling	-0.4234	0.023	CMS3
FH535	PPARgamma, PPARdelta	Metabolism	-0.2210	0.009	CMS2
Tretinoin	Retinoic acid	Other	0.1998	0.004	CMS2
FTI-277	Farnesyl-transferase (FNTA)	Other	-0.0904	0.039	CMS2
Bosutinib	SRC, ABL, TEC	Other, kinases	-0.3230	0.029	CMS1
SL0101	RSK, AURKB, PIM1, PIM3	Other, kinases	0.1593	0.031	CMS2
Bosutinib	SRC, ABL, TEC	Other, kinases	0.0804	0.008	CMS2
Sunitinib	PDGFR, KIT, VEGFR, FLT3, RET, CSF1R	RTK signaling	0.0503	0.031	CMS2
GW441756	NTRK1	RTK signaling	0.0752	0.020	CMS3
NVP-TAE684	ALK	RTK signaling	-0.0744	0.037	CMS3
CHIR-99021	GSK3A, GSK3B	WNT signaling	-0.1717	0.032	CMS2

Supplemental Figure Legends

Supplemental Figure 1. Differential expressed gene analysis reveals CMS-specific differences between BMI groups. (A) Volcano plots were used to visualize DESeq2-obtained DEGs (MeanBase > 10, FDR p value < 0.05) between Obese vs. Normal (A), Overweight vs Normal (B), and Obese vs. Overweight (C) comparisons for each CMS category. The R package EnhancedVolcano was used to construct the plots. The ratio of overexpressed to underexpressed DEGs is shown for each volcano plot. The DEGs with a false discovery rate less than 0.05 are shown as red dots while nonsignificant DEGs are represented as green dots. Select highly significant and differentially expressed genes are identified in the plots.

Supplemental Figure 2. Prognostic patient outcomes reveal CMS-specific differences between overweight and normal BMI groups. The expression of the top 20 significantly upregulated DESeq2-obtained DEGs (MeanBase > 10, FDR p value < 0.05) in the overweight vs. normal comparisons for each CMS category were assessed in GSE17536 and GSE41258 for overall survival (A) and GSE14333 and GSE17536 for relapse-free survival (B) using the PROGgeneV2 tool. Survival analyses were adjusted for age, stage, and gender covariates and bifurcated based on median expression. The hazard ratios, 95% confidence intervals, and p values were reported for the Kaplan-Meier plots.

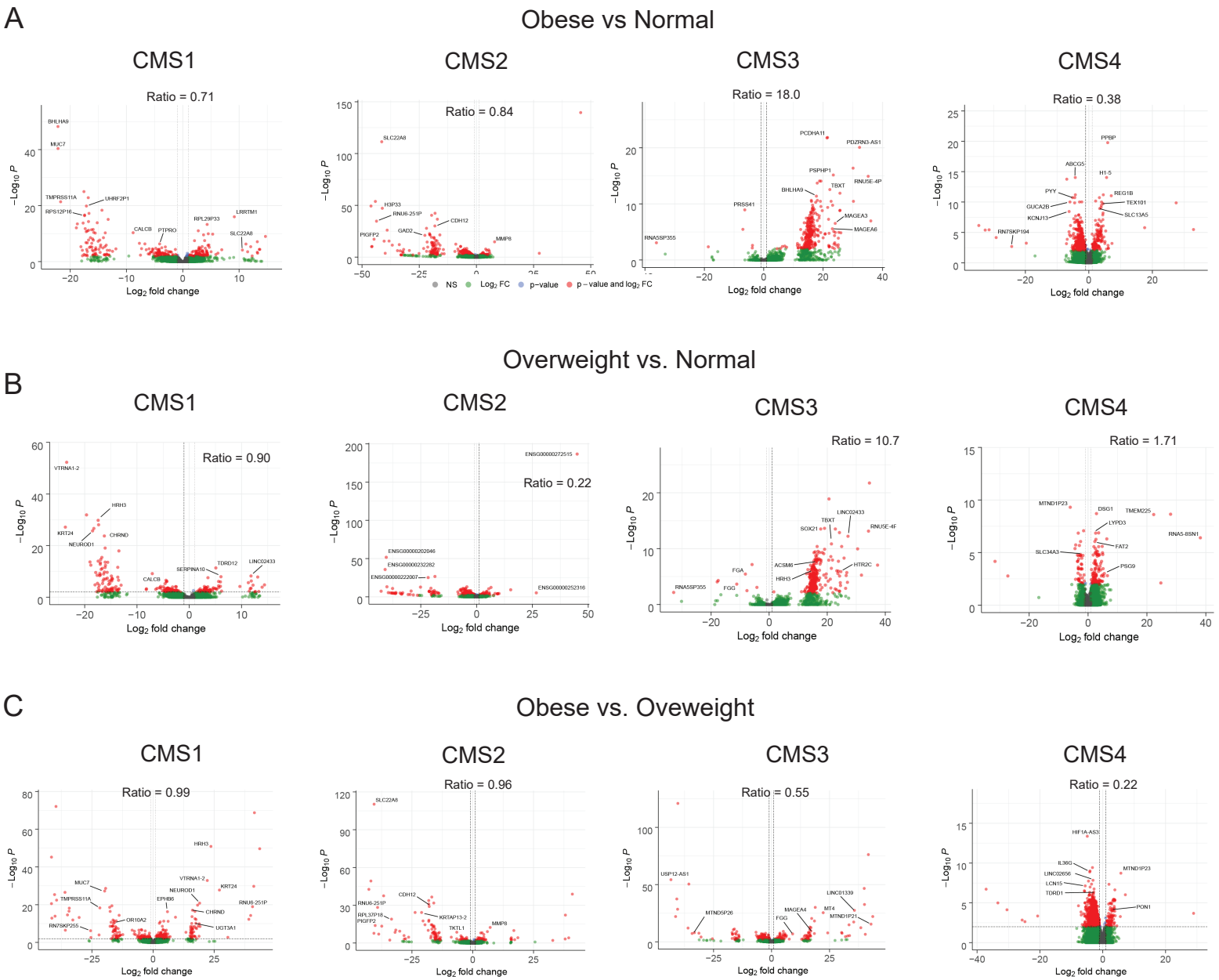
Supplemental Figure 3. Gene set enrichment of Hallmark gene sets reveal CMS-specific differences between overweight and normal BMI groups. Normalized RNA-seq counts obtained from DESeq2 were used for Gene Set Enrichment Analysis (GSEA). Hallmark gene sets

were assessed in the overweight vs. normal comparison for each CMS category. Bubbles for gene sets with a false discovery rate q-value less than 0.05 were reported. The size of bubbles represents the normalized enrichment score (NES). The color of bubbles represents false discovery rate q-value (FDR).

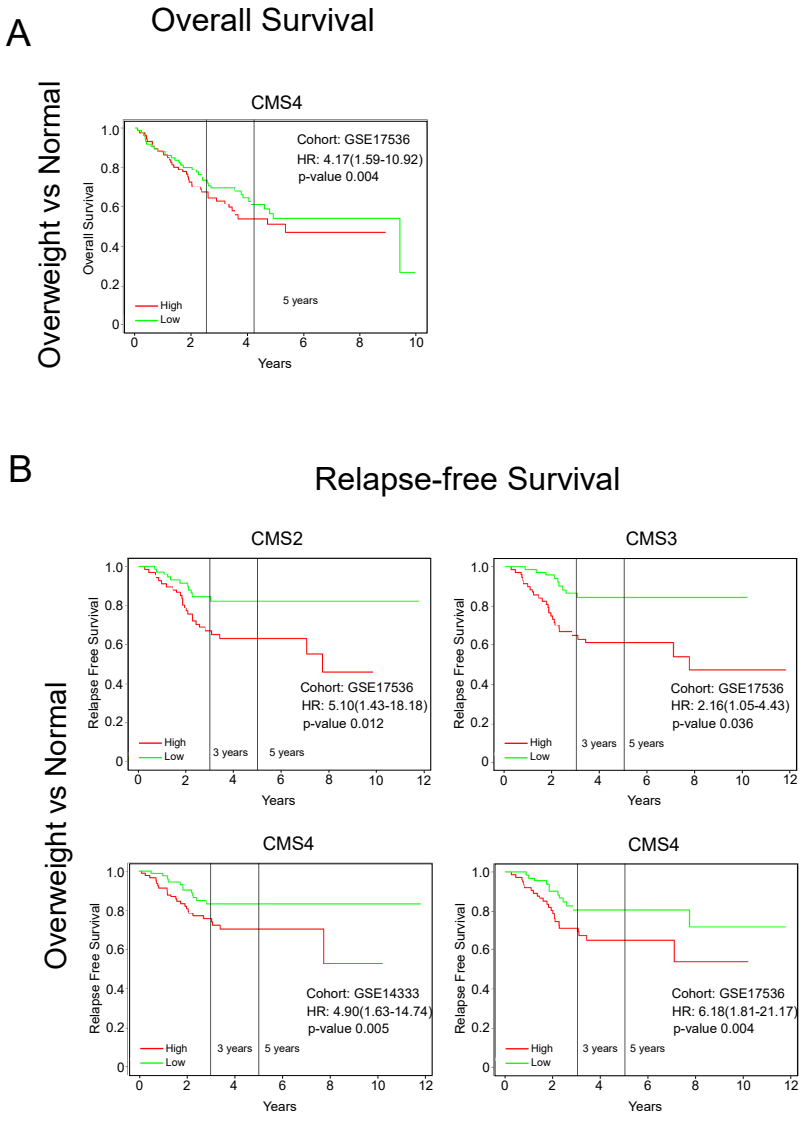
Supplemental Figure 4. Hub gene analysis reveal CMS-specific differences between overweight and normal BMI groups. DESeq2-obtained DEGs (MeanBase > 10, FDR p value < 0.05) were used to construct a protein-protein interaction (PPI) network from the STRING database in Cytoscape. The Cytohubba package in Cytoscape was used to perform the hub gene analysis from the PPI network for the overweight vs. normal comparison for each CMS category. Hub genes were identified using the maximal clique centrality (MCC) method to obtain the top 10 ranked genes in all modules. The intensity and color (high, red; orange, medium, yellow, low) of the hub genes is shown.

Supplemental Figure 5. Hub gene expression is relevant to colon adenocarcinoma in independent cancer patient cohorts. Hub genes were queried for mRNA expression using the Oncomine database. Hub gene expression in colon adenocarcinoma (Carcinoma) versus normal patient samples from the Hong Colorectal (Normal, n = 12; Carcinoma, n = 70), Skrzypczak Colorectal (Normal, n = 24; Carcinoma, n = 36), and Kaiser Colorectal (Normal, n = 5; Carcinoma, n = 41) cohorts. The p value from t-tests are shown. (A) *PYY*, (B) *PPBP*, (C) *INSL5*, and (D) *NPW*.

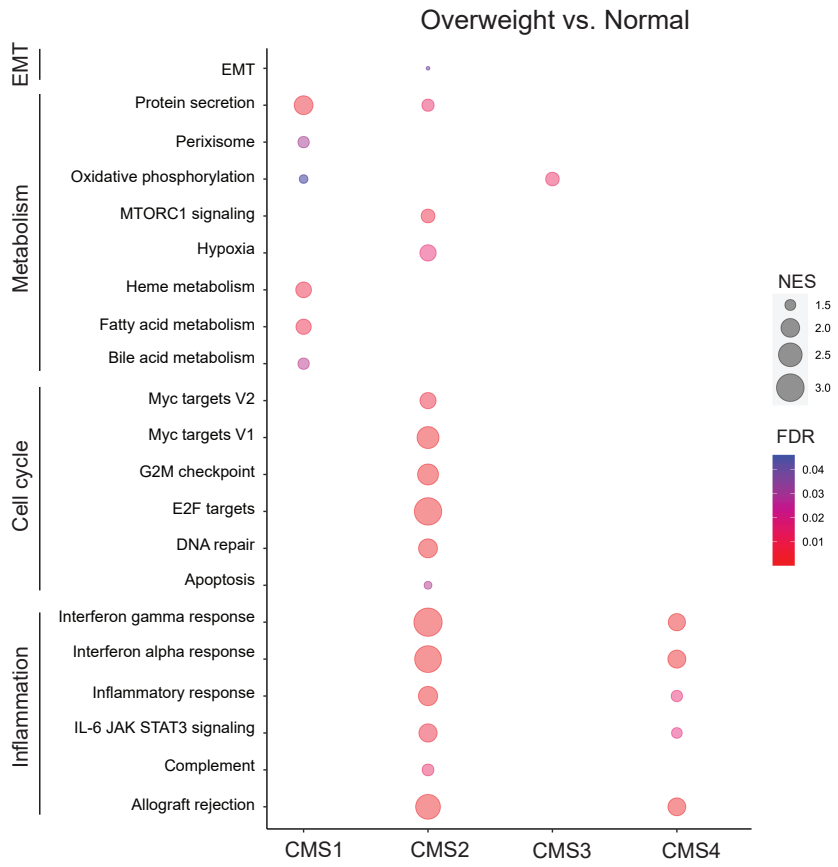
Supplemental Figure 1

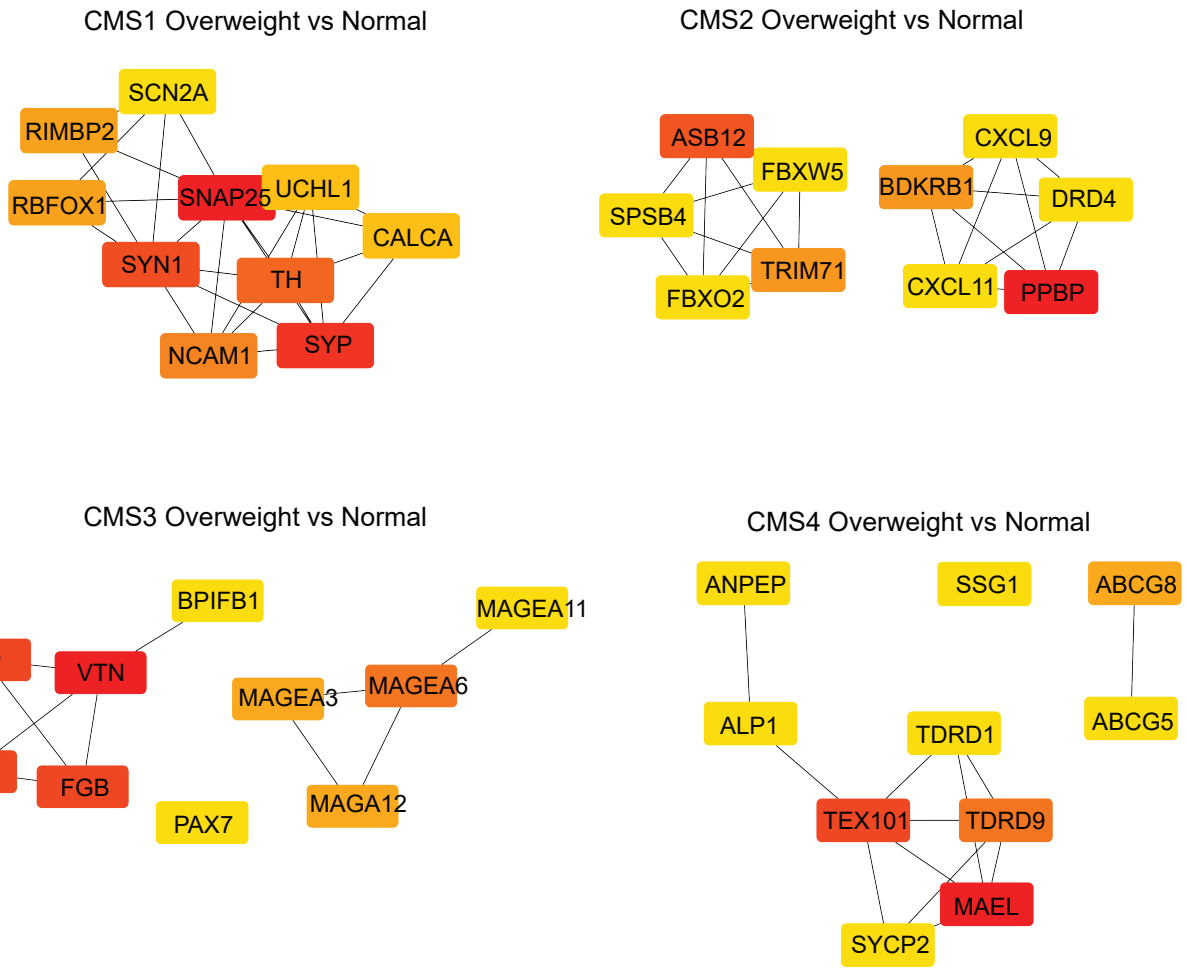


Supplemental Figure 2



Supplemental Figure 3





Supplemental Figure 5