McDonald Merry-Lynn Noelle (Orcid ID: 0000-0002-2647-9422)

# Genetic Variation in Genes Regulating Skeletal Muscle Regeneration and Tissue Remodeling Associated with Weight Loss in COPD

Lakshman Kumar P[1], Wilson A[1,2], Rocco A[1,2], Cho MH[3], Wan E[3,4], Hobbs BD[3,5], Washko GR[5], Ortega VE[6], Christenson SA[7], Li X[8], Wells J.M.[1], Bhatt SP[1], DeMeo DL[3], Lutz S[9], Rossiter H[10], Casaburi R[10], Rennard SI[11], Lomas DA[12], Labaki WW[13], Tal-Singer R[14], Bowler RP[15], Hersh CP[3], Tiwari HK[16], Dransfield M[1], Thalacker-Mercer A[17], Meyers DA[8], Silverman EK[3] and McDonald MN[1], on behalf of the COPDGene, ECLIPSE and SPIROMICS investigators

[1]Division of Pulmonary, Allergy and Critical Care Medicine, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL;

[2]Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL;

[3]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA;

[4]Veterans Affairs Boston Health Care System, Jamaica Plain, MA, USA;

[5] Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA;

[6] Department of Internal Medicine, Section on Pulmonary, Critical Care, Allergy and Immunologic Diseases, Wake Forest School of Medicine, Winston-Salem, NC, USA;

[7] Division of Pulmonary, Critical Care, Allergy, & Sleep Medicine, Department of Medicine, University of California San Francisco, San Francisco, CA, USA;

[8]Department of Medicine, University of Arizona College of Medicine, Tucson, AZ, USA;

[9] Department of Population Medicine, Harvard Medical School, Boston, MA, USA;

[10]Rehabilitation Clinical Trials Center, Los Angeles Biomedical Research Institute at Harbor Harbor-UCLA Medical Center, Torrance, CA, USA;

[11]Department of Medicine, Nebraska Medical Center, Omaha, Nebraska, USA

[12]UCL Respiratory, University College London, London, United Kingdom;

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jcsm.12782

[13]Division of Pulmonary and Critical Care Medicine, University of Michigan, Ann Arbor, MI, USA;

[14] COPD Foundation, Miami, FL, USA;

[15]Department of Medicine, Division of Pulmonary, Critical Care & Sleep Medicine National Jewish Health, Denver, CO, USA;

[16] Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA;

[17]Department of Cell Development and Integrative Biology, University of Alabama at Birmingham, Birmingham, AL, USA;

Corresponding author: Dr. Merry-Lynn McDonald, MSc, PhD, Director of Integrative 'Omics and Assistant Professor, Division of Pulmonary, Allergy and Critical Care Medicine, Department of Medicine, University of Alabama at Birmingham, 701 19th Street S, LHRB 440, Birmingham, AL 35233, T: +1-205-934-0714, F: +1-205-934-1721, Email: mmcdonald@uab.edu

**Abstract:**

***Background***: Chronic Obstructive Pulmonary Disease (COPD) is the third leading cause of death globally. COPD patients with cachexia or weight loss have increased risk of death independent of body mass index (BMI) and lung function. We tested the hypothesis genetic variation is associated with weight loss in COPD using a genome-wide association study (GWAS) approach.

***Methods***: Participants with COPD (N=4,308) from three studies (COPDGene, ECLIPSE, and SPIROMICS) were analyzed. Discovery analyses were performed in COPDGene with replication in SPIROMICS and ECLIPSE. In COPDGene, weight loss was defined as self-reported unintentional weight loss > 5% in the past year or low BMI (BMI < 20 kg/m$^2$). In ECLIPSE and SPIROMICS, weight loss was calculated using available longitudinal visits. Stratified analyses were performed among African American (AA) and Non-Hispanic White (NHW) participants with

COPD. Single variant and gene-based analyses were performed adjusting for confounders. Fine mapping was performed using a Bayesian approach integrating genetic association results with linkage disequilibrium and functional annotation. Significant gene networks were identified by integrating genetic regions associated with weight loss with skeletal muscle protein-protein interaction (PPI) data.

***Results***: At the single variant level, **only the rs35368512 variant**, intergenic to *GRXCR1* and *LINC02383,* was associated with weight loss (OR:3.6, 95%CI:2.3-5.6, P=$3.2\times10^{-8}$) among AA COPD participants in COPDGene. At the gene-level in COPDGene, *EFNA2* and *BAIAP2* were significantly associated with weight loss in AA and NHW COPD participants, respectively. The *EFNA2* association replicated among AA from SPIROMICS (P=0.0014) whereas the *BAIAP2* association replicated in NHW from ECLIPSE (P=0.025). The *EFNA2* gene encodes the membrane-bound protein ephrin-A2 involved in the regulation of developmental processes and adult tissue homeostasis such as skeletal muscle. The *BAIAP2* gene encodes the insulin-responsive protein of mass 53 kD (IRSp53), a negative regulator of myogenic differentiation. Integration of the gene-based findings participants with PPI data revealed networks of genes involved in pathways such as Rho and synapse signaling.

***Conclusions***: The *EFNA2* and *BAIAP2* genes were significantly associated with weight loss in COPD participants. Collectively, the integrative network analyses indicated genetic variation associated with weight loss in COPD may influence skeletal muscle regeneration and tissue remodeling.

**Keywords:** GWAS, cachexia, weight loss, COPD, genetics, biomarkers, skeletal muscle regeneration, tissue remodeling

**Introduction:**

Chronic Obstructive Pulmonary Disease (COPD) is the third leading cause of death internationally with associated mortality continuing to rise[1,2]. Although COPD is primarily diagnosed using lung function, traits not directly related to lung function such as cachexia greatly reduce quality of life and increase risk of mortality[3,4] . Cachexia is a debilitating co-morbidity increasing risk of death[3] and healthcare expenditure[5]. Most often thought of with respect to cancer, it has been estimated that there are 1.4 times as many patients with COPD cachexia than cancer cachexia by population prevalence[6].

Cachexia is defined as weight loss, primarily caused by loss of muscle with or without loss of fat, in individuals suffering from a chronic illness[7]. The consensus definition for cachexia diagnosis includes weight loss > 5% in the last 12 months or low BMI (BMI < 20 kg/mg$^2$) in addition to three out of five of: decreased muscle strength; fatigue; anorexia; low fat-free mass index (FFMI); and any indication of increased inflammatory markers (CRP, IL6, etc.), anemia or low serum albumin[7]. We recently demonstrated participants with COPD with cachexia and/or weight loss greater than 5% in the past year had a greater than three-fold increased mortality independent of BMI and lung function[8]. Monitoring cachexia using weight loss criteria is relatively simple and predictive of mortality among individuals with COPD.

Loss of muscle mass underlying weight loss in cachexia can be influenced by dysregulation of a number of mechanisms involved in the balance between protein synthesis and degradation[9]. Further, impaired ability to regenerate skeletal muscle tissue can contribute to muscle loss in COPD cachexia[9]. COPD patients with advanced disease are more likely to exhibit weight loss[8] in addition to skeletal muscle remodeling from a slow twitch (Type I) to a fast twitch (Type II) myofiber shift[10]. Skeletal muscle remodeling occurs in response to external stimuli leading to activation of intracellular signaling pathways and muscle fiber transition[11].

Although the major risk factor for COPD is smoking, COPD is a heritable disease with multiple genetic loci reproducibly associated[12-15]. The prevalence of cachexia in COPD is correlated with increasing disease severity[8]. Genetic variation may also contribute to the development of cachexia and weight loss in COPD. By performing genome-wide association study (GWAS) analyses, we previously identified genetic variants associated with longitudinal BMI in a small number of participants with COPD (N=237) in the Framingham Heart Study (FHS)[16]. As analyses were performed in a small size, investigation in a larger sample of participants with COPD with more precise phenotyping is merited. The genetics of cachexia have been more thoroughly investigated in cancer with several genes reproducibly associated using primarily candidate gene association approaches[17,18]. Cancer cachexia genes identified in association studies are known to be involved with inflammatory response regulation, pathways directing muscle and fat metabolism and appetite regulation among others[17,19].

We hypothesized genetic variation may be associated with weight loss contributing to cachexia in COPD. To test this hypothesis, we performed GWAS testing in 4,308 participants with COPD from 3 cohorts: COPD Genetic Epidemiology (COPDGene), Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) and Subpopulations and Intermediate Outcome Measures in COPD Study (SPIROMICS). Replication was assessed in remaining cohorts. GWAS findings were further explored by integrating gene-based findings with publicly available transcriptomics and protein-protein interaction (PPI) databases to gain additional insight to underlying biological mechanisms which may be influencing cachexia.

## Materials and Methods:

### *Ethics Statement*

**Institutional Review Board** approval for all analyses was obtained from the University of Alabama at Birmingham and performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

### *Study Participants*

The current analyses utilized COPD participants recruited as part of three studies: COPDGene[20], ECLIPSE[21] and SPIROMICS[22] (**Figure S1**). In all studies, COPD was classified using post-bronchodilator lung function testing ($FEV_1$: forced expiratory volume in one second and $FEV_1$/FVC: FEV1 expressed as a fraction of forced vital capacity) at baseline enrollment. All participants with COPD had moderate to severe disease defined by a GOLD[23] stage of 2 (FEV1/FVC < 0.7 and 50% < $FEV_1$ < 80% predicted) or higher (GOLD 3 and 4). In the current analyses, COPDGene COPD participants had at least a 10 pack-year smoking history, were aged 45 to 80 years at baseline and followed longitudinally with two visits 5 years apart. In ECLIPSE, COPD participants had at least a 10 pack-year smoking history and were aged 45 to 75 years at baseline and followed longitudinally. Visits in ECLIPSE occurred at baseline, 3 months, 6 months, and then every 6 months for three years. **In SPIROMICS, COPD participants had at least a 20 pack-year smoking history**, were aged 41 to 80 years at baseline and were followed with annual visits for three years.

*Weight Loss in COPD*

As diagramed in Supplementary Figure 1, COPDGene weight loss was defined as either self-reported, unintentional weight loss greater than 5% in the past year or as had low BMI (< 20 kg/m$^2$). We performed additional cleaning of the self-reported unintentional weight loss variable in COPDGene by confirming weight loss based on the weight measurements collected at Visits 1 and 2. This led to the exclusion of 2 participants (**Figure S1**). In ECLIPSE and SPIROMICS, weight loss greater than 5% was defined if present at any of the longitudinal visits.

*Genome-wide association study (GWAS) analyses*

Genotyping in COPDGene was performed using the Illumina Human Omni 1-Quad (Illumina, San Diego, CA), in ECLIPSE using the Illumina HumanHap 550v3 chips and in SPIROMICS using the Illumina HumanOmniExpressExome BeadChip (Illumina, Inc., San Diego, CA). Standard quality control steps were performed on DNA samples and single-nucleotide polymorphism (SNP) data as previously described [24-26]. For all three studies, genotypes were imputed using the Haplotype Reference Consortium (HRC) reference panel [27]. Only SNPs with an imputation quality score of 0.5 or greater were included in the analysis. **SNP positions were reported based on the human genome (hg) 19 build. SNPs were annotated to genes or closest genes using the NCBI hg19 RefGene database (version 2017-06-01) as implemented using ANNOVAR[28].** In COPDGene, a total of 5,405,435 and 7,629,332 variants were imputed and passed quality control in the **Non-Hispanic White (NHW) and African American (AA)** COPD participants, respectively. In ECLIPSE, a total of 5,370,356 variants were imputed and passed quality control. In SPIROMICS, a total of 5,421,262 variants were imputed and passed quality control in NHW and AA participants. Statistical analyses were performed in PLINK **v1.90b3.45[29]** and R

**vfoss/2016b**[30]. Discovery analyses were performed using data from AA COPD and NHW COPD participants from COPDGene and results were assessed in the remaining cohorts for replication. Association with weight loss for each SNP with a minor allele frequency (MAF) of 5% or greater was tested assuming an additive model adjusting for age, sex and principal components of genetic ancestry controlling for genetic ancestry. The level of genome-wide significance (GWS) for single nucleotide polymorphisms (SNPs) association tests was defined as $P<5.0x10^{-8}$**. This level of GWS is a value traditionally used in GWAS to account for the large number of variants in linkage disequilibrium**[31]**. It approximates a Bonferroni corrected P=0.05 for 1 million independent tests in the genome.** Regional association plots were generated using LocusZoom[32] and linkage disequilibrium information from the 1,000 Genomes African Ancestry and 1,000 Genomes European Ancestry reference panels[33] were used. Meta-analyses were performed using METAL[34] software. Gene-based analyses were performed using MAGMA[35] software which integrates both single SNP tests with linkage disequilibrium patterns within gene regions. MAGMA assigns SNPs to genes based on physical position (± 50 kb) of known genes in the NCBI site. Statistical significance (GWS) for gene-based tests was defined as $P<2.5x10^{-6}$, which corresponds to Bonferroni corrected **P-value threshold for P=0.05** for the approximate 20,000 genes in the genome.

*Fine Mapping Analysis*

To prioritize biological causality of the genotyped variants, fine mapping was performed using PAINTOR[36]. PAINTOR implements a Bayesian approach incorporating genetic association results, linkage disequilibrium and functional annotation to generate the posterior probability of causality for each variant. Fine-mapping regions were prioritized based on examination of the regional association plots as well as including variants ± 25 kb from the lead variant. Single SNP test statistic (Z score) information from the gene regions for the variants in the partitioned gene regions was used as input and were functionally annotated to the regions in Skeletal Muscle, Lung, Brain, Adipose and Liver. Top 5 regions functionally annotated to the variants based on highest likelihood ratio were included for the analysis.

*Defining modules by integrating GWAS findings with PPI data*

R library dmGWAS version 2.4 was used to integrate GWAS findings with PPI data [37]. The dmGWAS algorithm is applied to integrate GWAS results with PPI data by using the PPI data as a search space to examine gene networks also termed modules. Nodes in the network represent gene-based results with edges representing PPI between proteins encoded by two genes. Every gene in the GWAS results is considered as an initial seed gene by the dmGWAS algorithm with a starting test statistic corresponding to the gene-based result. Using the search space defined by the PPI network additional genes are considered for inclusion in the module. A gene is included in the module if it increases the module test statistic by a factor of r. PPI data was downloaded from PINA[38], which includes collected and annotated data from 6 public databases: MINT, IntAct, DIP, BioGRID, HPRD and MIPS/IMPact, on Dec 16, 2019. We further subset the PPI network data based on proteins in the Compiled Skeletal Muscle Proteome[39]. In the current analyses, gene-based meta-analysis in the AA and NHW COPD participants was used as an input to dmGWAS. A distance constraint of d=2 and r=0.1 were used. A normalized module score accounting for the number of genes in the modules was generated. dmGWAS function simpleChoose was used to choose the top 10 ranked modules based on the normalized module score. Subnetworks created for each ethnicity-based PPI network was visualized using Cytoscape 3.6.1[40].

*Pathway and Tissue Enrichment Analyses*

Gene set enrichment analysis (GSEA) [41] and tissue enrichment analyses was performed using the Functional Mapping and Annotation (FUMA)[42] software to examine known biology of the network modules that were generated by dmGWAS. Tissue enrichment was assessed by testing whether collections of genes exhibit tissue specific expression patterns based on the Genotype-Tissue Expression (GTEx) project version 8 data[43] implemented in FUMA.

**Results:**

*COPD Population Characteristics*

We contrasted descriptive characteristics of COPD participants in the discovery cohort, COPDGene, with the replication cohorts, SPIROMICS and ECLIPSE (**Table 1**). When comparing SPIROMICS and COPDGene, SPIROMICS AA COPD participants, on average, were more likely to be older, have lower BMI, have better lung function, and to have unintentional weight loss (**Table 1**). On average, among NHW COPD participants, those in ECLIPSE were more likely to be male, have lower BMI, have less smoking exposure, have worse lung function, and have higher unintentional weight loss than those in COPDGene and SPIROMICS **(Table 1).** Whereas NHW SPIROMICS subjects tended to be older, on average, compared to COPDGene and ECLIPSE **(Table 1).** The prevalence of weight loss was 17% and 14.6% in COPDGene AA and NHW participants with COPD, respectively. Whereas the prevalence of the weight loss trait ranged from 30.7% to 38.6% in SPIROMICS and ECLIPSE COPD participants.

*Examining association of single SNPs with weight loss in participants with COPD*

Among AA COPD participants, **the rs35368512 variant** was significantly associated with weight loss in the discovery analysis (OR:3.6, 95%CI:2.3-5.6, $P=3.2x10^{-8}$, **Supplementary Table 1**) but did not replicate in the AA COPD participants from SPIROMICS. The **rs35368512** variant is intergenic with the closest gene, *GRXCR1*, located within 200 kb (**Figure 1A**). When the meta-analysis results of weight loss in AA COPD participants were examined, no additional variant was associated at a level of GWS (**Supplementary Table 2**). The top variant associated with weight loss in AA COPD participants in the meta-analysis was intronic to the *TBX15* gene (**Supplementary Table 2**). Among the NHW COPD participants in COPDGene, no single variant was significantly associated with weight loss (**Supplementary Table 3**) nor reached GWS in the meta-analysis of all the populations (**Supplementary Table 4**). The top single variant (rs62015138, OR=2.1, 95% CI:1.6-2.8, P=$6.4 \times 10^{-7}$) associated with weight loss among NHW COPD participants in COPDGene was within the *RBFOX1* gene (**Supplementary Table 3**). The top single variant associated with weight loss in the meta-analysis of NHW COPD participants (**rs35017521**, OR= 1.4, 95% CI:1.2-1.6, P=$7.7 \times 10^{-7}$, **Supplementary Table 4**) is intergenic between a microRNA gene (*MIR6072*) and a long intergenic non-protein coding RNA gene (*LINC00701*).

*EFNA2 and BAIAP2 gene-based regions associated with* weight loss *in participants with COPD*

At the gene level, *EFNA2* was associated at level of GWS with weight loss among AA COPD participants (**Table 2**). This finding replicated (P<0.05) among AA COPD participants from SPIROMICS contributing to a meta-analysis P=$1.4 \times 10^{-8}$ (**Table 2**). The lead *EFNA2* variant, chr19:1304013, was associated with an increased risk of weight loss (OR= 3.6, 95% CI= 1.9-6.7, P=$6.2 \times 10^{-5}$) in discovery analyses, which was not GWS. In the meta-analysis, an additional two genes, *C19orf24* and *CIRBP*, were associated with weight loss at a level of GWS among AA COPD participants (**Table 2**). The *EFNA2* and *CIRBP* genes are located in the same region on chromosome 19 (**Figure 1B**). Fine-mapping of 109 variants indicated 41% were needed to obtain a credible set with modest posterior probability. Fine mapping analyses indicated several variants within *EFNA2* had modest posterior probability (PP) of being causal (PP>0.15, **Figure 2**) with the combined region accounting for 99% of the PP.

Among NHW COPD participants, the *BAIAP2* gene was significantly associated with weight loss (**Table 3**) with the finding replicating in ECLIPSE. Among NHW COPD participants from SPIROMICS, the *BAIAP2* gene was also associated with weight loss at a level near nominal significance (P=0.055) contributing to the significant meta-analysis result (P=$5.20 \times 10^{-7}$). In the discovery analyses, NHW COPD participants in COPDGene the top *BAIAP2* variant, chr17:79084367, was associated with decreased risk of weight loss in COPD (OR= 0.60, 95%

CI=0.48–0.76, P=1.5x10$^{-5}$) but not a level of GWS. The *BAIAP2* gene, on chromosome 17, is near (5 kb) to the *AATK* gene (**Figure 3**). The *AATK* gene is physically close but transcribed in the opposite direction with the two genes having overlapping 3' non-coding regions[44]. The *AATK* gene was significantly associated with weight loss in COPD participants from COPDGene but did not replicate in ECLIPSE or SPIROMICS (**Table 3**). Fine-mapping of 477 variants indicated 41% were needed to obtain a credible set with modest posterior probability. Fine-mapping analysis indicated only one variant within *BAIAP2* had a modest PP of being causal with the remaining variants having low likelihood of being causal (PP<5%, **Figure 3**). However, the collective set which included variants within *AATK* increased the combined likelihood to 99%.

*Integrating GWAS findings with PPI data to identify networks of COPD* weight loss *genes*

COPD weight loss consensus networks were generated using unsupervised integration of gene-based meta-analysis results with PPI data of proteins expressed in skeletal muscle. Integration of the meta-analysis weight loss gene-based results from AA COPD participants identified 12,135 modules. A consensus module was generated from the top 10 most significant modules included 29 genes (**Figure 4**, **Supplementary Table 5**). Several of these genes (*EFNA2, CIRBP*, *WDR88* and *KCNK1*) in the consensus module were among the top 10 most associated with weight loss in AA COPD participants (**Table 2**). GSEA indicated the consensus module genes associated with weight loss in the AA COPD participants were enriched in pathways involved in *NRF1* signaling, adipogenesis, synapse and RNA metabolism among others (**Supplementary Table 6**). Integration of the meta-analysis weight loss gene-based results from NHW COPD participants with PPI network data identified 12,168 modules. A consensus network was created from the top 10 most significant modules and was comprised of 36 genes (**Figure 4, Supplementary Table 7**). Of which, several genes (*BAIAP2, AATK, ZZEF1 and RHOB*) were among the top 10 most significantly associated with weight loss in NHW COPD participants (**Table 3**). GSEA results indicated the consensus module genes were enriched in pathways involved in adipogenesis, synapse signaling, Rho GTPase signaling as well as genes in other known other pathways

(**Supplementary Table 8**). Despite having only one gene, *UBC*, common to both the AA and NHW weight loss consensus networks, there were 6 known gene-sets in common. These were comprised of genes: 1) involved in synapse signaling; 2) involved in formation of the incision complex; 3) with sites recognized by miR-520D; 4) involved in protein tagging for modification, sequestration, transport or degradation; 5) involved in adipocyte differentiation; and 6) involved in subdivision of chromosomal regions.

The *EFNA2* gene is a member of the set of 16 synapse signaling genes represented in consensus weight loss networks in both AA and NHW COPD participants (combined GO SYNAPSE gene list in **Supplementary Tables 6 and 8**). The 16 genes include *STXBP3, DISC1, KCNK1, KPNA2, ITSN1* and *PPP2CA* as well as *PPP1CA, SNAP23, NSF, ELAVL1, BIN1, PACSIN2, ARL8B, STX7* and *STXBP5* in the AA and NHW COPD weight loss consensus networks, respectively. Tissue specificity analyses indicated genes were significantly downregulated in the liver ($P<1x10^{-4}$). Further, five of these genes (*KCNK1, ELAVL1, ARL8B, PPP2CA* and *STXBP5)* also have sites recognized by the miRNA520D (GSEA adjusted $P=2.1x10^{-5}$).

**Discussion:**

In the current report, we identified two genes significantly associated with weight loss in COPD: *EFNA2* and *BAIAP2*. Although, *EFNA2* and *BAIAP2* were associated with weight loss in COPD among AA and NHW participants, respectively, they were also members of the COPD weight loss consensus networks. This indicates there is genetic variation in these genes that could influence other genes within the same network. As COPD and weight loss are complex traits, we expect genetic variation in many genes contribute to dysregulation at the pathway level. Importantly, many of the gene-sets enriched in the COPD weight loss consensus networks have important roles in skeletal muscle regeneration and tissue remodeling.

The *EFNA2* gene encodes the membrane-bound protein ephrin-A2. Ephrins interact with Eph receptors via contact-dependent cell-cell signaling regulating developmental processes and adult tissue homeostasis[45]. Interestingly, ephrin-A2 participates in bi-directional signaling by activating eph receptors on neighboring cells as well as its own downstream pathways and has been shown to negatively regulate progenitor cell proliferation[46]. The *EFNA2* gene is a member of the GO SYNAPSE gene-set whose members were represented in both the AA and NHW COPD weight loss consensus networks of genes. We demonstrated using GTeX data the 15 COPD weight loss genes in the SYNAPSE gene-set are downregulated in the liver. The role in the liver is interesting as in metabolic disturbances have been shown to start in the liver before progressing to adipose and skeletal muscle tissues in mouse cancer cachexia models[47,48]. Further, the liver has afferent and efferent neurons which may influence appetite and hormone signals[49]. ~~**Five of the 15 genes (*KCNK1, ELAVL1, ARL8B, PPP2CA* and *STXBP5*) have a binding site for miR-520D. Further, miR-520D has been shown to influence hepatic regulation of low-density lipoprotein (LDL) clearance[50]. Thus, regulation of miR-520D may be a mechanism for therapeutic intervention for cachexia.**~~ However, we were not able test whether genetic variation associated with weight loss in COPD leads to altered gene expression of these 15 genes in the liver. Therefore, additional research into whether expression of these genes is upregulated in the liver with COPD cachexia and weight loss is needed.

The *BAIAP2* gene encodes the insulin-responsive protein of mass 53 kD (IRSp53*)*, which is an adaptor protein primarily known for its role in modulating actin dynamics and membrane protrusions in cell to cell signaling[51]. Impaired skeletal muscle regeneration is one mechanism contributing to skeletal muscle loss in cachexia[52]. IRSp53 can act as a negative regulator of myogenic differentiation influencing the development of skeletal muscle as well skeletal muscle regeneration [52]. Previous research demonstrated COPD patients may exhibit heterogeneous and distinct skeletal muscle molecular biomarker patterns in response to pulmonary rehabilitation[53]. However, IRSp53 was not among the biomarkers investigated in the previous research[53]. It is possible dysregulation of IRSp53 may be one mechanism contributing to impaired ability to regenerate skeletal muscle in cachexia, however, this requires further research. Also, GSEA analyses highlighted *BAIAP2* involvement in Rho signaling. Activation of the Rho signaling pathway is required for the maintenance of myotubes[54]. Interestingly, ephrins such as ephrin-B whose gene was associated with weight loss among AA COPD participants activate eph receptors who exert downstream by regulating Rho GTPase signaling[45].

Our study has many strengths but also limitations. Strengths of the study include a large sample size of COPD participants in multiple cohorts followed longitudinally with weight loss and genotype data available. We also used innovative network methods, going beyond generating a list of genes associated with weight loss in COPD, providing rationales for further mechanistic research. Limitations of the study include heterogeneity in number of visits used to define weight loss in the discovery and replication cohorts. In COPDGene, a self-reported unintentional weight loss greater than 5% in the past year or low BMI collected at a single visit was used. Whereas in ECLIPSE and SPIROMICS, weight loss and BMI were measured at several annual visits. The prevalence of weight loss was higher in ECLIPSE and SPIROMICS which may be due to increased opportunities for observing weight loss over the course of the study. In COPDGene, weight loss was also unintentional whereas in the other two studies participants may have been intentionally trying to lose weight which could inflate the prevalence of weight loss. Self-reported, unintentional weight loss is likely a more conservative measure, however, may also be subject to recall bias. Discovery analyses were performed using more strict criteria to code the weight loss trait which would have biased findings towards the null. **However, the lower number of visits over larger time intervals in COPDGene likely led to misclassification and loss of follow-up of participants with COPD who passed away before weight loss could be recorded by the study. This would have limited our ability to identify some true associations with weight loss in COPD using COPDGene as discovery. Future studies of weight loss in COPD should aim to collect weight measurements more frequently such as the intervals employed in ECLIPSE**

**and SPIROMICS. Further, we also employed a Bonferroni-corrected level of significance to gene-based findings which may have been overly conservative.**

Furthermore, the *UBC* gene encoding Ubiquitin-C was the only gene in common between the AA and NHW COPD weight loss consensus networks. The ubiquitin-proteasome system is fundamental to muscle atrophy in cachexia[55]. However, we previously demonstrated the network analysis method, dmGWAS, was sensitive to hub genes such as UBC, encoding Ubiquitin C[56]. For these reasons, we also performed the network analyses excluding *UBC* and found *EFNA2* and *BAIPA2* were robustly included in each consensus module whether UBC was included in the PPI search space or not. Finally, we analyzed genotyped and imputed SNP data which led to the identification of two gene regions associated with weight loss in COPD rather than specific genetic variants. We maximized the information in the regions through our fine mapping approach. **Although we analyzed a combined set of 4,308 subject with COPD in the three studies combined, the samples sizes became small when stratified by study and ancestry group likely limiting the power to replicate findings in the discovery. For example, the top variant associated with weight loss in the meta-analysis of AA COPD subjects is intronic to the *TBX15* gene. *TBX15* is a member of the T-box family of transcription factors and has been previously associated with waist to hip to ratio[57]. Nonetheless, we were able to maximize information using our integrative approach to discover new etiology for weight loss in COPD.** However, an expanded analysis in these populations of COPD participants using whole-genome sequence with integration with other 'omics data may lead to data identifying specific genetic variants which may guide personalize medicine approaches.

To summarize, *BAIAP2* and *EFNA2* genes were significantly associated with weight loss in COPD among NHW and AA participants. Our integrative network analyses identified COPD weight loss genes enriched with genes involved in skeletal muscle regeneration and tissue remodeling as well as providing rationales for further mechanistic research. Identification of genetic variation contributing to weight loss in COPD due to impaired skeletal muscle regeneration and tissue remodeling may enable discovery of therapies which could enhance response to pulmonary rehabilitation.



**Table 1.** Characteristics of COPD participants included in GWAS of weight loss. Continuous variables (age, BMI, FEV1pp, pack-years of smoking) are represented by means and standard deviations.

| | African Americans | | | Non-Hispanic Whites | | | |
|---|---|---|---|---|---|---|---|
| **Characteristic** N (%) | **COPDGene** | **SPIROMICS** | **p-value*** | **COPDGene** | **ECLIPSE** | **SPIROMICS** | **p-value†** |
| **N** | 401 | 138 | -- | 1380 | 1569 | 820 | -- |
| **Sex** (% Male) | 197 (49.1) | 71 (51.4) | 0.82 | 758 (54.9) | 1, 054 (67.2) | 466 (56.8) | 0.48 |
| **Age** | 58.1±7.5 | 61.4±8.2 | <0.0001 | 64.1±7.9 | 63.7±7.0 | 65.9±7.4 | <0.0001 |
| **BMI** | 28.4±6.6 | 26.7±5.8 | 0.0076 | 28.6±5.9 | 26.8±5.6 | 27.5±5.2 | <0.0001 |
| **Pack Years** | 41.6±22.4 | 42.8±19.5 | 0.56 | 53.7±26.1 | 50.0±27.1 | 57.2±51.0 | <0.0001 |
| **FEV1pp** | 56.4±15.2 | 53.3±16.6 | 0.046 | 54.2±16.2 | 48.3±15.6† | 54.1±16.7 | <0.0001 |
| **Weight Loss**[1] | 68 (17.0) | 50 (36.2) | 0.0082 | 201 (14.6) | 605 (38.6) † | 252 (30.7) | 0.0047 |

COPD: Chronic Obstructive Pulmonary Disease, WL: weight loss, BMI: body mass index, FEV1pp: forced expiratory volume in 1 second expressed as percentage of predicted, GOLD: Global Initiative for Chronic Obstructive Lung Disease;
*P-value generated from chi-square test statistic for categorical variables and paired test statistic for continuous variables comparing African Americans between COPDGene and SPIROMICS;
†P-value generated from chi-square test statistic for categorical variables and one-way ANOVA test statistic for continuous variables comparing Non-Hispanics whites between COPDGene, ECLIPSE, and SPIROMICS;
[1]Weight loss is defined as WL > 5% and/or low BMI at any time point in the study.

**Table 2:** Based on meta-analysis, top 10 genes associated with weight loss in participants with COPD in African American cohorts (COPDGene and SPIROMICS).

| | Meta-analysis | | | | | | COPDGene AA | | | SPIROMICS AA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **GENE** | **CHR** | **START** | **STOP** | **NSNPS** | **ZSTAT** | **P** | **NSNPS** | **ZSTAT** | **P Value** | **NSNPS** | **ZSTAT** | **P Value** |
| ***EFNA2**** | 19 | 1236153 | 1351430 | 252 | 5.6 | 1.40E-08 | 270 | 4.7 | 1.34E-06 | 233 | 3.0 | 1.42E-03 |
| *C19orf24* | 19 | 1225520 | 1329243 | 220 | 5.2 | 1.20E-07 | 237 | 4.5 | 4.00E-06 | 202 | 2.6 | 4.59E-03 |
| ***CIRBP**** | 19 | 1219267 | 1324809 | 221 | 4.6 | 2.10E-06 | 235 | 4.1 | 1.74E-05 | 206 | 2.0 | 2.10E-02 |
| *WI2-2373I1.2* | 7 | 280136 | 384388 | 100 | 3.9 | 4.80E-05 | 81 | 4.6 | 2.16E-06 | 119 | -0.1 | 5.49E-01 |
| ***WDR88**** | 19 | 33572949 | 33717830 | 434 | 3.8 | 6.10E-05 | 435 | 2.2 | 1.43E-02 | 432 | 3.9 | 5.74E-05 |
| ***KCNK1**** | 1 | 233699750 | 233858258 | 505 | 3.8 | 6.30E-05 | 515 | 3.3 | 5.50E-04 | 495 | 2.0 | 2.19E-02 |
| *TET2* | 4 | 106017032 | 106250960 | 477 | 3.7 | 1.30E-04 | 481 | 2.8 | 2.55E-03 | 473 | 2.5 | 7.01E-03 |
| *OR13C8* | 9 | 107281449 | 107382411 | 390 | 3.5 | 1.90E-04 | 391 | 4.4 | 5.86E-06 | 388 | -0.5 | 6.77E-01 |
| *BTNL3* | 5 | 180365845 | 180483727 | 258 | 3.5 | 2.60E-04 | 239 | 3.3 | 4.40E-04 | 276 | 1.2 | 1.18E-01 |
| *MAGI2* | 7 | 77596374 | 79133121 | 5017 | 3.5 | 2.70E-04 | 5142 | 3.4 | 3.26E-04 | 4892 | 1.0 | 1.53E-01 |

***Denotes genes that appear in the consensus network***

**Table 3:** Top 10 genes associated with weight loss in participants with COPD based on meta-analysis in Non-Hispanic White cohorts (COPDGene, ECLIPSE, and SPIROMICS).

| Meta-analysis | | | | | | | COPDGene NHW | | | ECLIPSE | | | SPIROMICS NHW | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GENE | CHR | START | STOP | NSNPS | ZSTAT | P | NSNPS | ZSTAT | P | NSNPS | ZSTAT | P | NSNPS | ZSTAT | P |
| ***BAIAP2**** | 17 | 78958944 | 79141232 | 578 | 4.9 | 5.20E-07 | 581 | 4.8 | 9.83E-07 | 577 | 2.0 | 2.54E-02 | 577 | 1.6 | 5.50E-02 |
| *C8orf48* | 8 | 13374352 | 13475797 | 440 | 4.3 | 8.60E-06 | 434 | 1.5 | 6.28E-02 | 444 | 2.8 | 2.81E-03 | 442 | 3.4 | 3.42E-04 |
| ***AATK**** | 17 | 79034285 | 79189877 | 400 | 4.1 | 1.90E-05 | 392 | 4.6 | 2.50E-06 | 393 | 1.6 | 6.01E-02 | 416 | 0.8 | 2.22E-01 |
| ***ZZEF1**** | 17 | 3857739 | 4096314 | 644 | 3.8 | 7.00E-05 | 642 | 2.7 | 3.79E-03 | 644 | 2.2 | 1.57E-02 | 645 | 1.7 | 4.25E-02 |
| *CACNG6* | 19 | 54444403 | 54565920 | 314 | 3.7 | 1.30E-04 | 315 | 2.3 | 9.99E-03 | 316 | 2.3 | 1.10E-02 | 312 | 1.6 | 4.95E-02 |
| *SEMA5A* | 5 | 8985138 | 9596233 | 1431 | 3.5 | 2.10E-04 | 1395 | 1.9 | 3.09E-02 | 1461 | 3.7 | 9.64E-05 | 1437 | 0.0 | 5.07E-01 |
| *UGT2A3* | 4 | 69744177 | 69867509 | 381 | 3.5 | 2.20E-04 | 380 | 0.9 | 1.87E-01 | 381 | 2.5 | 5.65E-03 | 381 | 2.9 | 1.96E-03 |
| ***RHOB**** | 2 | 20596835 | 20699201 | 290 | 3.5 | 2.70E-04 | 291 | 2.3 | 9.67E-03 | 289 | 1.8 | 4.00E-02 | 291 | 2.0 | 2.55E-02 |
| *ENPP1* | 6 | 132079156 | 132266295 | 329 | 3.4 | 3.20E-04 | 335 | 3.7 | 1.04E-04 | 335 | 1.4 | 8.78E-02 | 317 | 0.6 | 2.62E-01 |
| *ASB5* | 4 | 177084824 | 177248722 | 472 | 3.4 | 3.30E-04 | 473 | 3.6 | 1.48E-04 | 471 | 1.3 | 1.02E-01 | 472 | 0.9 | 1.97E-01 |

****Denotes genes that appear in the consensus network***

**Figure 1A:** Regional Association plot for top weight loss variant region in African American participants from COPDGene near variant, 4:43232667. **1B:** Regional Association plot for top weight loss variant region in Non-Hispanic White COPD participants from COPDGene on chromosome 19. In both panels, lead SNPs denoted in purple. Color of remaining SNPs indicate degree of linkage disequilibrium with the lead SNP, as measured by r2, the squared coefficient of correlation.

**Figure2**: Fine mapping of *EFNA2* region associated with weight loss among African American COPD participants. a) Scatterplot of location versus posterior probabilities with credible set; b) Physical position of EFNA2; c) functional annotation tracks

**Figure 3**: Fine mapping of *BAIAP2* region associated with weight loss among Non-Hispanic White COPD participants. a) Scatterplot of location versus posterior probabilities with credible set; b) Physical position of genes including BAIAP2 in region; c) functional annotation tracks

**Figure 4: COPD weight loss consensus network generated from African American participant analyses.** Node size is proportional to p-value significance where the bigger the node size the smaller the p-value from the gene-based meta-analysis result. Nodes (circles)represents genes that are among the top genes in the consensus network. Edges (lines) are known protein-protein (PPI) interactions. Nodes filled in black represents genes robust to exclusion of UBC from the PPI network.

**Figure 5: COPD weight loss consensus network generated from Non-Hispanic White participant analyses.** Node size is proportional to p-value significance where the bigger the node size the smaller the p-value from the gene-based meta-analysis result. Nodes (circles) represents genes that are among the top genes in the consensus network. Edges (lines) are known protein-protein (PPI) interactions. Nodes filled in black represents genes robust to exclusion of UBC from the PPI network.

**Supplementary Figure 1. COPD GOLD Stage>=2 participants with phenotype and genotype data included in analyses.** A) COPD participants included from the COPDGene study; B) COPD participants included from SPIROMICS study; C) COPD participants included from ECLIPSE study. GWAS – Genome Wide Association Study. In COPDGene, 2 subjects with Visit 2 self-reported weight loss data were excluded for being incongruent with Visit 1 weight.

**Supplementary Table 1:** SNPs associated with COPD weight loss among African American COPD participants from COPDGene (N=401) with P<1E-5.

**Supplementary Table 2**: Top 10 SNPs associated with weight loss among African American COPD participants based on meta-analysis COPDGene and SPIROMICS.

**Supplementary Table 3:** SNPs associated with COPD weight loss among Non-Hispanic White COPD participants from COPDGene with P<1E-5.

**Supplementary Table 4**: Top 10 SNPs associated with weight loss among Non-Hispanic White COPD participants based on meta-analysis of three studies (ECLIPSE, COPDGene and SPIROMICS).

**Supplementary Table 5**: Genes represented in the COPD weight loss consensus network from African American participants with corresponding gene-based results provided.

**Supplementary Table 6**: Gene Set Enrichment Analysis of Consensus Network Genes associated with weight loss among AA COPD participants from COPDGene and SPIROMICS.

**Supplementary Table 7**: Genes represented in the COPD weight loss consensus network from non-Hispanic White participants with corresponding gene-based results provided.

**Supplementary Table 8:** Gene Set Enrichment Analysis of Consensus Network Genes associated with weight loss among NHW COPD participants from COPDGene, SPIROMICS and ECLIPSE.

**Conflict of interest:**
RT-S is a former employee and current shareholder of GSK the sponsor of the ECLIPSE study. She declares. Personal fees from Immunomet, Vocalis Health, Ena Respiratory and Teva. EKS has received institutional grant support from GSK and Bayer. MW has received institutional grant support from Mereo BioPharma, Grifols, Vertex Pharmaceuticals, Bayer AG, ARCUS-Med, Verona, and has been in consultancy roles for AstraZeneca, GlaxoSmithKline, Boehringer Ingelheim, and Takeda. GRW has received institutional grant support from Janssen Pharmaceuticals and Boehringer Ingelheim and has been in consultancy roles for Boehringer Ingelheim, CSL Behring, Janssen Pharmaceuticals, Novartis and Vertex. GRW is also a founder and equity holder of Quantitative Imaging Solutions, a data and image analytics company. WWL reports grants from NIH/NHLBI, non-financial support from Pulmonx and personal fees from Konica Minolta. MHC has received grant funding from Bayer and GSK and speaking or consulting fees from Genentech, AstraZeneca, and Illumina. HBR received institutional grant support from GlaxoSmithKline, AstraZeneca, and Boehringer Ingelheim and consultancy fees from Boehringer Ingelheim, Astellas Pharma and Omniox Inc. MTD reports grants from NIH, Department of Defense, and the American Lung Association; contracted clinical trial support from AstraZeneca, Boehringer Ingelheim, Boston Scientific, Gala, GlaxoSmithKline, Nuvaira, PneumRx/BTG, and Pulmonx; travel from Pulmonx; and consulting fees from AstraZeneca, GlaxoSmithKline, Mereo, Quark, and Teva.

**Funding:**
MNM: NIH R00HL121087, R56HL153460 and the Parker B. Francis Fellowship program. DAL is funded by the National Institute for Health Research (NIHR) University College Hospitals (UCLH) Biomedical Research Centre (BRC), he is an NIHR Senior Investigator. MHC is funded by NIH R01HL135142, R01 HL137927, and R01 HL14714. BDH is funded by NIH K08 HL136928. The COPDGene Study was supported by Award Number U01 HL089897 and Award Number U01 HL089856 from the National Heart, Lung, and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health. COPDGene is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprised of AstraZeneca, Boehringer-Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, Siemens, and Sunovion. SPIROMICS was supported by contracts from the NIH/ NHLBI (HHSN268200900013C, HHSN268200900014C, HHSN268200900015C, HHSN268200900016C, HHSN268200900017C, HHSN268200900018C, HHSN268200900019C, HHSN268200900020C), grants from the NIH/NHLBI (U01 HL137880 and U24 HL141762), and supplemented by contributions made through the Foundation for the NIH and the COPD Foundation from AstraZeneca/MedImmune; Bayer; Bellerophon Therapeutics; Boehringer-Ingelheim Pharmaceuticals, Inc.; Chiesi Farmaceutici S.p.A.; Forest Research Institute, Inc.; GlaxoSmithKline; Grifols Therapeutics, Inc.; Ikaria, Inc.; Novartis Pharmaceuticals Corporation; Nycomed GmbH; ProterixBio; Regeneron Pharmaceuticals, Inc.; Sanofi; Sunovion; Takeda Pharmaceutical Company; and Theravance Biopharma and Mylan.

**Acknowledgements:**
**COPDGene® Investigators – Core Units**
*Administrative Center*: James D. Crapo, MD (PI); Edwin K. Silverman, MD, PhD (PI); Barry J. Make, MD; Elizabeth A. Regan, MD, PhD
*Genetic Analysis Center*: Terri Beaty, PhD; Ferdouse Begum, PhD; Peter J. Castaldi, MD, MSc; Michael Cho, MD; Dawn L. DeMeo, MD, MPH; Adel R. Boueiz, MD; Marilyn G. Foreman, MD, MS; Eitan Halper-Stromberg; Lystra P. Hayden, MD, MMSc; Craig P. Hersh, MD, MPH; Jacqueline Hetmanski, MS, MPH; Brian D. Hobbs, MD; John E. Hokanson, MPH, PhD; Nan Laird, PhD; Christoph Lange, PhD; Sharon M. Lutz, PhD; Merry-Lynn McDonald, PhD; Margaret M. Parker, PhD; Dmitry Prokopenko, Ph.D; Dandi Qiao, PhD; Elizabeth A. Regan, MD, PhD; Phuwanat Sakornsakolpat, MD; Edwin K. Silverman, MD, PhD; Emily S. Wan, MD; Sungho Won, PhD
*Imaging Center*: Juan Pablo Centeno; Jean-Paul Charbonnier, PhD; Harvey O. Coxson, PhD; Craig J. Galban, PhD; MeiLan K. Han, MD, MS; Eric A. Hoffman, Stephen Humphries, PhD; Francine L. Jacobson, MD, MPH; Philip F. Judy, PhD; Ella A. Kazerooni, MD; Alex Kluiber;

David A. Lynch, MB; Pietro Nardelli, PhD; John D. Newell, Jr., MD; Aleena Notary; Andrea Oh, MD; Elizabeth A. Regan, MD, PhD; James C. Ross, PhD; Raul San Jose Estepar, PhD; Joyce Schroeder, MD; Jered Sieren; Berend C. Stoel, PhD; Juerg Tschirren, PhD; Edwin Van Beek, MD, PhD; Bram van Ginneken, PhD; Eva van Rikxoort, PhD; Gonzalo Vegas Sanchez-Ferrero, PhD; Lucas Veitel; George R. Washko, MD; Carla G. Wilson, MS;
*PFT QA Center, Salt Lake City, UT*: Robert Jensen, PhD
*Data Coordinating Center and Biostatistics*, *National Jewish Health, Denver, CO*: Douglas Everett, PhD; Jim Crooks, PhD; Katherine Pratte, PhD; Matt Strand, PhD; Carla G. Wilson, MS
*Epidemiology Core*, *University of Colorado Anschutz Medical Campus, Aurora, CO*: John E. Hokanson, MPH, PhD; Gregory Kinney, MPH, PhD; Sharon M. Lutz, PhD; Kendra A. Young, PhD
*Mortality Adjudication Core:* Surya P. Bhatt, MD; Jessica Bon, MD; Alejandro A. Diaz, MD, MPH; MeiLan K. Han, MD, MS; Barry Make, MD; Susan Murray, ScD; Elizabeth Regan, MD; Xavier Soler, MD; Carla G. Wilson, MS
*Biomarker Core*: Russell P. Bowler, MD, PhD; Katerina Kechris, PhD; Farnoush Banaei-Kashani, Ph.D

**COPDGene® Investigators – Clinical Centers**
*Ann Arbor VA:* Jeffrey L. Curtis, MD; Perry G. Pernicano, MD
*Baylor College of Medicine, Houston, TX*: Nicola Hanania, MD, MS; Mustafa Atik, MD; Aladin Boriek, PhD; Kalpatha Guntupalli, MD; Elizabeth Guy, MD; Amit Parulekar, MD;
*Brigham and Women's Hospital, Boston, MA*: Dawn L. DeMeo, MD, MPH; Alejandro A. Diaz, MD, MPH; Lystra P. Hayden, MD; Brian D. Hobbs, MD; Craig Hersh, MD, MPH; Francine L. Jacobson, MD, MPH; George Washko, MD
*Columbia University, New York, NY*: R. Graham Barr, MD, DrPH; John Austin, MD; Belinda D'Souza, MD; Byron Thomashow, MD
*Duke University Medical Center, Durham, NC*: Neil MacIntyre, Jr., MD; H. Page McAdams, MD; Lacey Washington, MD
*Grady Memorial Hospital, Atlanta, GA: Eric Flenaugh, MD; Silanth Terpenning, MD*
*HealthPartners Research Institute, Minneapolis, MN*: Charlene McEvoy, MD, MPH; Joseph Tashjian, MD
*Johns Hopkins University, Baltimore, MD*: Robert Wise, MD; Robert Brown, MD; Nadia N. Hansel, MD, MPH; Karen Horton, MD; Allison Lambert, MD, MHS; Nirupama Putcha, MD, MHS
*Lundquist Institute for Biomedical Innovationat Harbor UCLA Medical Center, Torrance, CA*: Richard Casaburi, PhD, MD; Alessandra Adami, PhD; Matthew Budoff, MD; Hans Fischer, MD; Janos Porszasz, MD, PhD; Harry Rossiter, PhD; William Stringer, MD
*Michael E. DeBakey VAMC, Houston*, *TX*: Amir Sharafkhaneh, MD, PhD; Charlie Lan, DO
*Minneapolis VA:* Christine Wendt, MD; Brian Bell, MD; Ken M. Kunisaki, MD, MS
*National Jewish Health, Denver, CO*: Russell Bowler, MD, PhD; David A. Lynch, MB

*Reliant Medical Group, Worcester, MA*: Richard Rosiello, MD; David Pace, MD
*Temple University, Philadelphia, PA:* Gerard Criner, MD; David Ciccolella, MD; Francis Cordova, MD; Chandra Dass, MD; Gilbert D'Alonzo, DO; Parag Desai, MD; Michael Jacobs, PharmD; Steven Kelsen, MD, PhD; Victor Kim, MD; A. James Mamary, MD; Nathaniel Marchetti, DO; Aditi Satti, MD; Kartik Shenoy, MD; Robert M. Steiner, MD; Alex Swift, MD; Irene Swift, MD; Maria Elena Vega-Sanchez, MD
*University of Alabama, Birmingham, AL:* Mark Dransfield, MD; William Bailey, MD; Surya P. Bhatt, MD; Anand Iyer, MD; Hrudaya Nath, MD; J. Michael Wells, MD
*University of California, San Diego, CA*: Douglas Conrad, MD; Xavier Soler, MD, PhD; Andrew Yen, MD
*University of Iowa, Iowa City, IA*: Alejandro P. Comellas, MD; Karin F. Hoth, PhD; John Newell, Jr., MD; Brad Thompson, MD
*University of Michigan, Ann Arbor, MI:* MeiLan K. Han, MD MS; Ella Kazerooni, MD MS; Wassim Labaki, MD MS; Craig Galban, PhD; Dharshan Vummidi, MD
*University of Minnesota, Minneapolis, MN*: Joanne Billings, MD; Abbie Begnaud, MD; Tadashi Allen, MD
*University of Pittsburgh, Pittsburgh, PA*: Frank Sciurba, MD; Jessica Bon, MD; Divay Chandra, MD, MSc; Carl Fuhrman, MD; Joel Weissfeld, MD, MPH
*University of Texas Health, San Antonio, San Antonio, TX*: Antonio Anzueto, MD; Sandra Adams, MD; Diego Maselli-Caceres, MD; Mario E. Ruiz, MD; Harjinder Singh

**SPIROMICS**
The authors thank the SPIROMICS participants and participating physicians, investigators and staff for making this research possible. More information about the study and how to access SPIROMICS data is available at www.spiromics.org. The authors would like to acknowledge the University of North Carolina at Chapel Hill BioSpecimen Processing Facility for sample processing, storage, and sample disbursements (http://bsp.web.unc.edu/). We would like to acknowledge the following current and former investigators of the SPIROMICS sites and reading centers: Neil E Alexis, MD; Wayne H Anderson, PhD; Mehrdad Arjomandi, MD; Igor Barjaktarevic, MD, PhD; R Graham Barr, MD, DrPH; Patricia Basta, PhD; Lori A Bateman, MSc; Surya P Bhatt, MD; Eugene R Bleecker, MD; Richard C Boucher, MD; Russell P Bowler, MD, PhD; Stephanie A Christenson, MD; Alejandro P Comellas, MD; Christopher B Cooper, MD, PhD; David J Couper, PhD; Gerard J Criner, MD; Ronald G Crystal, MD; Jeffrey L Curtis, MD; Claire M Doerschuk, MD; Mark T Dransfield, MD; Brad Drummond, MD; Christine M Freeman, PhD; Craig Galban, PhD; MeiLan K Han, MD, MS; Nadia N Hansel, MD, MPH; Annette T Hastie, PhD; Eric A Hoffman, PhD; Yvonne Huang, MD; Robert J Kaner, MD; Richard E Kanner, MD; Eric C Kleerup, MD; Jerry A Krishnan, MD, PhD; Lisa M LaVange, PhD; Stephen C Lazarus, MD; Fernando J Martinez, MD, MS; Deborah A Meyers, PhD; Wendy C Moore, MD; John D

Newell Jr, MD; Robert Paine, III, MD; Laura Paulin, MD, MHS; Stephen P Peters, MD, PhD; Cheryl Pirozzi, MD; Nirupama Putcha, MD, MHS; Elizabeth C Oelsner, MD, MPH; Wanda K O'Neal, PhD; Victor E Ortega, MD, PhD; Sanjeev Raman, MBBS, MD; Stephen I. Rennard, MD; Donald P Tashkin, MD; J Michael Wells, MD; Robert A Wise, MD; and Prescott G Woodruff, MD, MPH. The project officers from the Lung Division of the National Heart, Lung, and Blood Institute were Lisa Postow, PhD, and Lisa Viviano.

**References:**

1. Vestbo, J. *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* **187**, 347-65 (2013).
2. Xu, J., Murphy, S.L., Kochanek, K.D., Bastian, B. & Arias, E. Deaths: Final Data for 2016. *Natl Vital Stat Rep* **67**, 1-76 (2018).
3. Evans, W.J. *et al.* Cachexia: a new definition. *Clin Nutr* **27**, 793-9 (2008).
4. McDonald, M.N. *et al.* It's more than low BMI: prevalence of cachexia and associated mortality in COPD. *Respir Res* **20**, 100 (2019).

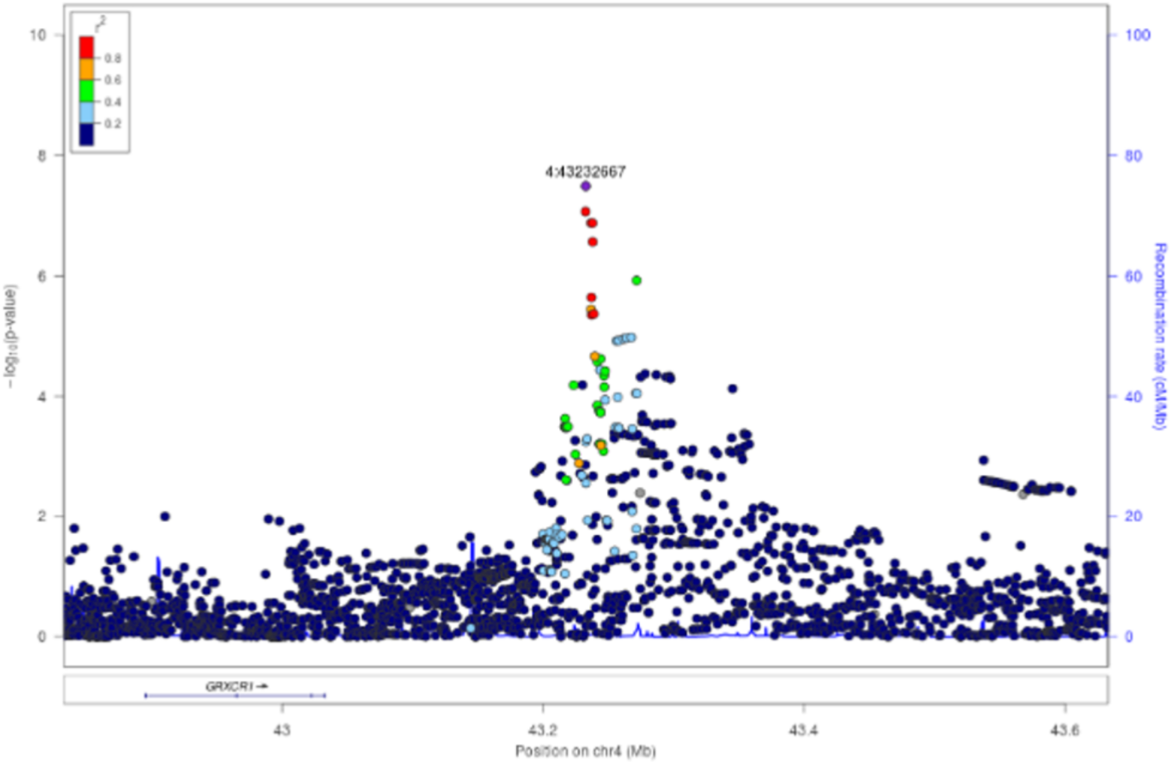
5. Arthur, S.T., Noone, J.M., Van Doren, B.A., Roy, D. & Blanchette, C.M. One-year prevalence, comorbidities and cost of cachexia-related inpatient admissions in the USA. *Drugs Context* **3**, 212265 (2014).
6. von Haehling, S. & Anker, S.D. Prevalence, incidence and clinical impact of cachexia: facts and numbers - update 2014. *Journal of Cachexia, Sarcopenia and Muscle* **5**, 261-263 (2014).
7. Evans, W.J. *et al.* Cachexia: a new definition. *Clinical nutrition (Edinburgh, Scotland)* **27**, 793-799 (2008).
8. McDonald, M.-L.N. *et al.* It's more than low BMI: prevalence of cachexia and associated mortality in COPD. *Respiratory research* **20**, 100 (2019).
9. Remels, A.H., Gosker, H.R., Langen, R.C. & Schols, A.M. The mechanisms of cachexia underlying muscle dysfunction in COPD. *J Appl Physiol (1985)* **114**, 1253-62 (2013).
10. Gosker, H.R., Zeegers, M.P., Wouters, E.F.M. & Schols, A.M.W.J. Muscle fibre type shifting in the vastus lateralis of patients with COPD is associated with disease severity: A systematic review and meta-analysis. *Thorax* **62**, 944-949 (2007).
11. Bassel-Duby, R. & Olson, E.N. Signaling pathways in skeletal muscle remodeling. *Annu Rev Biochem* **75**, 19-37 (2006).
12. Hobbs, B.D. *et al.* Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nature Genetics* (2017).
13. Cho, M.H. *et al.* A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. *Human Molecular Genetics* **21**, 947-957 (2012).
14. Prokopenko, D. *et al.* Whole-Genome Sequencing in Severe Chronic Obstructive Pulmonary Disease. *Am J Respir Cell Mol Biol* **59**, 614-622 (2018).
15. Jackson, V.E. *et al.* Exome-wide analysis of rare coding variation identifies novel associations with COPD and airflow limitation in MOCS3, IFIT3 and SERPINA12. *Thorax* **71**, 501-9 (2016).
16. Mcdonald, M.-L.N. *et al.* Body mass index change in gastrointestinal cancer and chronic obstructive pulmonary disease is associated with Dedicator of Cytokinesis 1. *Journal of Cachexia, Sarcopenia and Muscle*, 1-9 (2017).
17. Tan, B.H. *et al.* Identification of possible genetic polymorphisms involved in cancer cachexia: a systematic review. *Journal of Genetics* **90**, 165-177 (2011).
18. Johns, N. *et al.* Genetic basis of interindividual susceptibility to cancer cachexia: selection of potential candidate gene polymorphisms for association studies. *Journal of Genetics,* **93**, 893-916 (2014).
19. Johns, N. *et al.* New genetic signatures associated with cancer cachexia as defined by low skeletal muscle index and weight loss. *Journal of Cachexia, Sarcopenia and Muscle* **8**, 122-130 (2017).

20. Regan, E.A. *et al.* Genetic epidemiology of COPD (COPDGene) study design. *COPD* **7**, 32-43 (2010).
21. Vestbo, J. *et al.* Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *The European respiratory journal* **31**, 869-873 (2008).
22. Hansel, N.N. *et al.* Design of the Subpopulations and Intermediate Outcome Measures in COPD (SPIROMICS) AIR Study. *BMJ Open Respir Res* **4**, e000186 (2017).
23. Vogelmeier, C.F. *et al.* Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Lung Disease 2017 Report: GOLD Executive Summary. *Respirology* **22**, 575-601 (2017).
24. Cho, M.H. *et al.* Variants in FAM13A are associated with chronic obstructive pulmonary disease. *Nat Genet* **42**, 200-2 (2010).
25. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature genetics* **50**, 1335-1341 (2018).
26. Cho, M.H. *et al.* Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med* **2**, 214-25 (2014).
27. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* **48**, 1279-83 (2016).
28. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).
29. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
30. Ward, L.D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930-4 (2012).
31. Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* **32**, 381-5 (2008).
32. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-7 (2010).
33. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
34. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
35. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).
36. Kichaev, G. *et al.* Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet* **10**, e1004722 (2014).

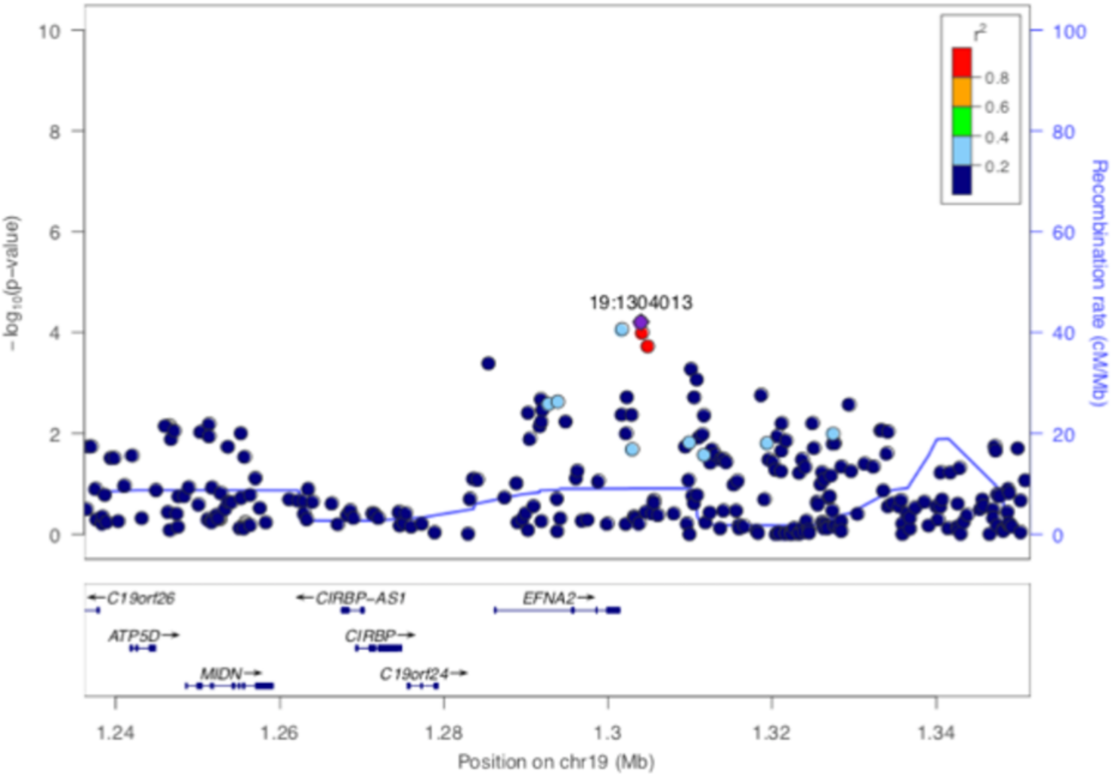
37. Jia, P., Zheng, S., Long, J., Zheng, W. & Zhao, Z. dmGWAS: dense module searching for genome-wide association studies in protein-protein interaction networks. *Bioinformatics* **27**, 95-102 (2011).
38. Wu, J. *et al.* Integrated network analysis platform for protein-protein interactions. *Nat Methods* **6**, 75-7 (2009).
39. Gonzalez-Freire, M. *et al.* The Human Skeletal Muscle Proteome Project: a reappraisal of the current literature. *J Cachexia Sarcopenia Muscle* **8**, 5-18 (2017).
40. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-504 (2003).
41. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-50 (2005).
42. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nature communications* **8**, 1826 (2017).
43. Consortium, G.T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5 (2013).
44. Miyahara, A., Okamura-Oho, Y., Miyashita, T., Hoshika, A. & Yamada, M. Genomic structure and alternative splicing of the insulin receptor tyrosine kinase substrate of 53-kDa protein. *J Hum Genet* **48**, 410-4 (2003).
45. Lisabeth, E.M., Falivelli, G. & Pasquale, E.B. Eph receptor signaling and ephrins. *Cold Spring Harb Perspect Biol* **5**(2013).
46. Holmberg, J. *et al.* Ephrin-A2 reverse signaling negatively regulates neural progenitor proliferation and neurogenesis. *Genes Dev* **19**, 462-71 (2005).
47. Halle, J.L. *et al.* Tissue-specific dysregulation of mitochondrial respiratory capacity and coupling control in colon-26 tumor-induced cachexia. *Am J Physiol Regul Integr Comp Physiol* **317**, R68-R82 (2019).
48. Khamoui, A.V., Tokmina-Roszyk, D., Rossiter, H.B., Fields, G.B. & Visavadiya, N.P. Hepatic proteome analysis reveals altered mitochondrial metabolism and suppressed acyl-CoA synthetase-1 in colon-26 tumor-induced cachexia. *Physiol Genomics* **52**, 203-216 (2020).
49. Jensen, K.J., Alpini, G. & Glaser, S. Hepatic nervous system and neurobiology of the liver. *Compr Physiol* **3**, 655-65 (2013).
50. Salerno, A.G. *et al.* LDL Receptor Pathway Regulation by miR-224 and miR-520d. *Front Cardiovasc Med* **7**, 81 (2020).
51. Scita, G., Confalonieri, S., Lappalainen, P. & Suetsugu, S. IRSp53: crossing the road of membrane and actin dynamics in the formation of membrane protrusions. *Trends Cell Biol* **18**, 52-60 (2008).
52. Misra, A. *et al.* Insulin receptor substrate protein 53kDa (IRSp53) is a negative regulator of myogenic differentiation. *Int J Biochem Cell Biol* **44**, 928-41 (2012).

53. Kneppers, A.E.M. *et al.* Distinct skeletal muscle molecular responses to pulmonary rehabilitation in chronic obstructive pulmonary disease: a cluster analysis. *Journal of Cachexia, Sarcopenia and Muscle* (2019).
54. Wallace, M.A. *et al.* Striated muscle activator of Rho signalling (STARS) is a PGC-1alpha/oestrogen-related receptor-alpha target gene and is upregulated in human skeletal muscle after endurance exercise. *J Physiol* **589**, 2027-39 (2011).
55. Sanders, K.J.C., Kneppers, A.E.M., van de Bool, C., Langen, R.C.J. & Schols, A.M.W.J. Cachexia in chronic obstructive pulmonary disease: New insights and therapeutic perspective. *Journal of Cachexia, Sarcopenia and Muscle* **7**, 5-22 (2016).
56. McDonald, M.L. *et al.* Beyond GWAS in COPD: probing the landscape between gene-set associations, genome-wide associations and protein-protein interaction networks. *Hum Hered* **78**, 131-9 (2014).
57. Heid, I.M. *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* **42**, 949-60 (2010).
58. von Haehling, S., Morley, J.E., Coats, A.J.S. & Anker, S.D. Ethical guidelines for publishing in the Journal of Cachexia, Sarcopenia and Muscle: update 2019. *J Cachexia Sarcopenia Muscle* **10**, 1143-1145 (2019).

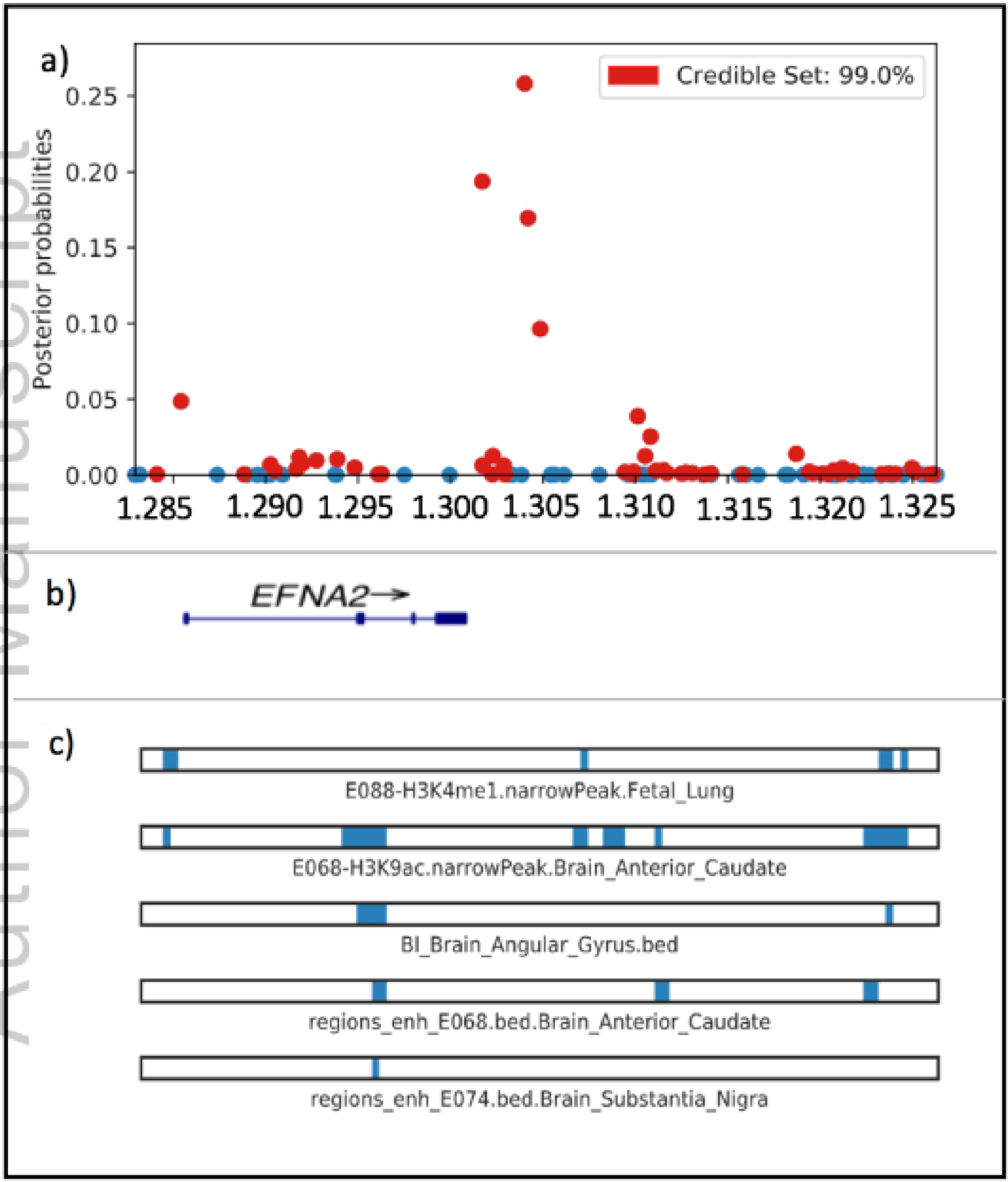
**A**



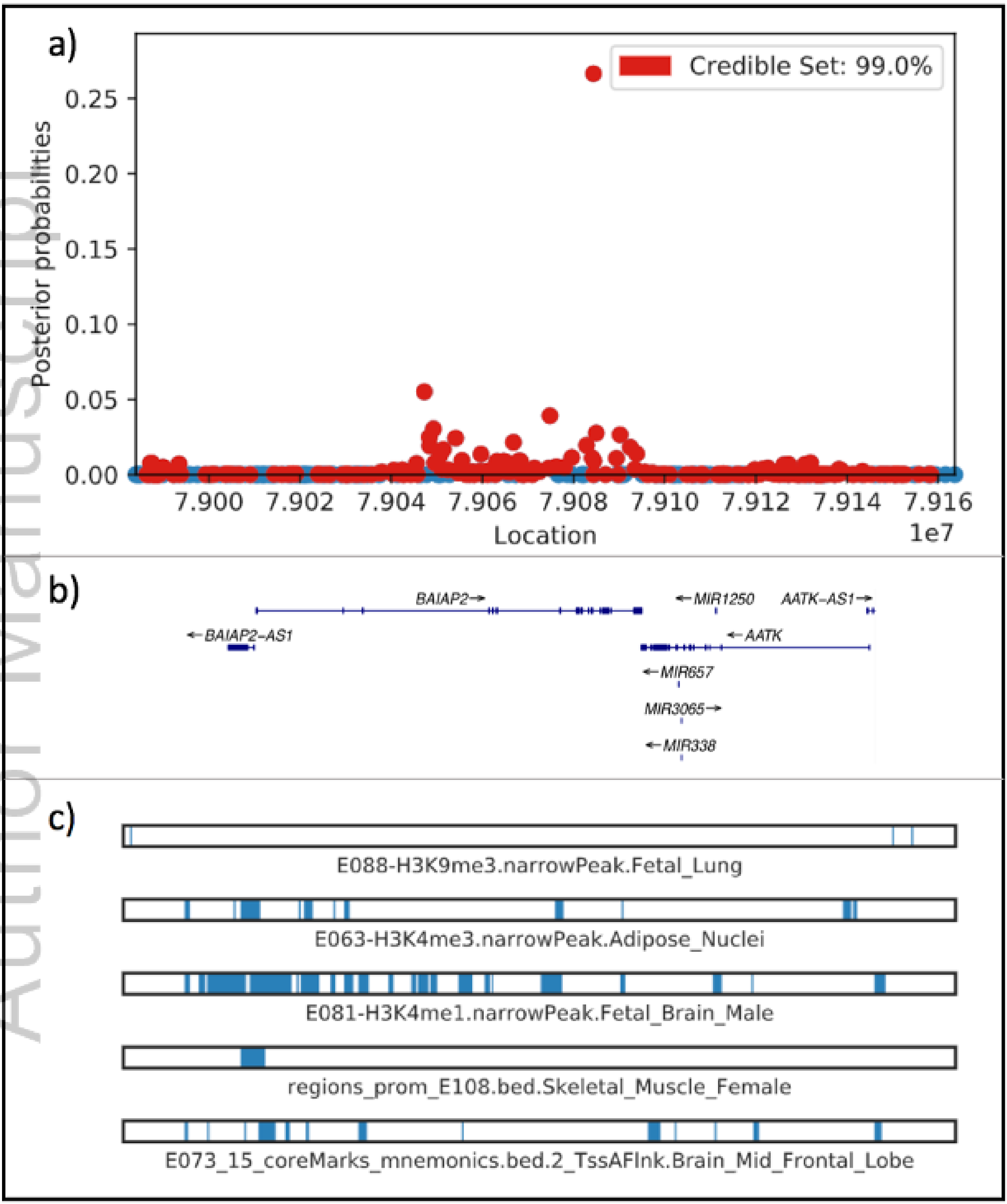

**B**



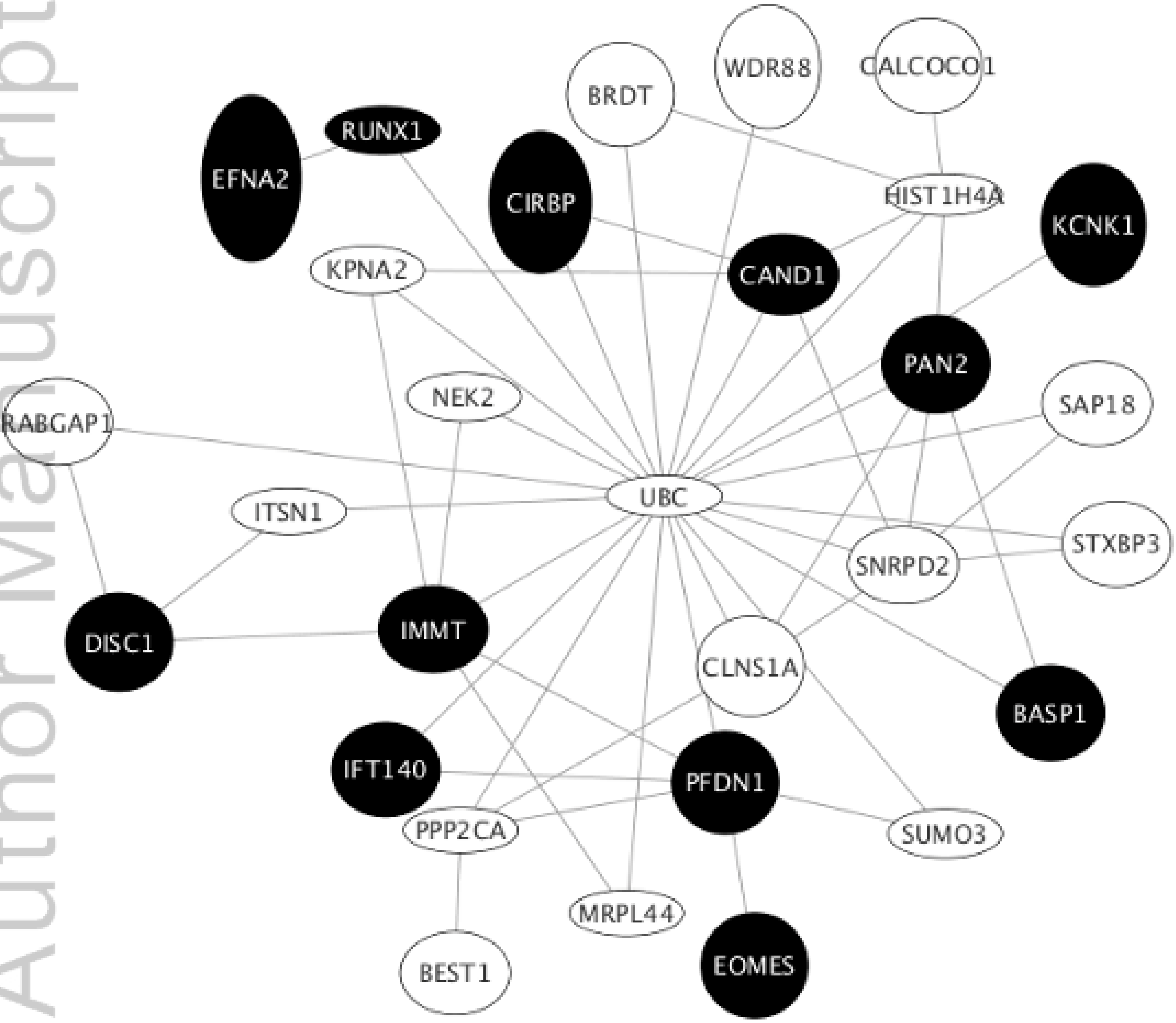

JCSM_12782_Figure1.png



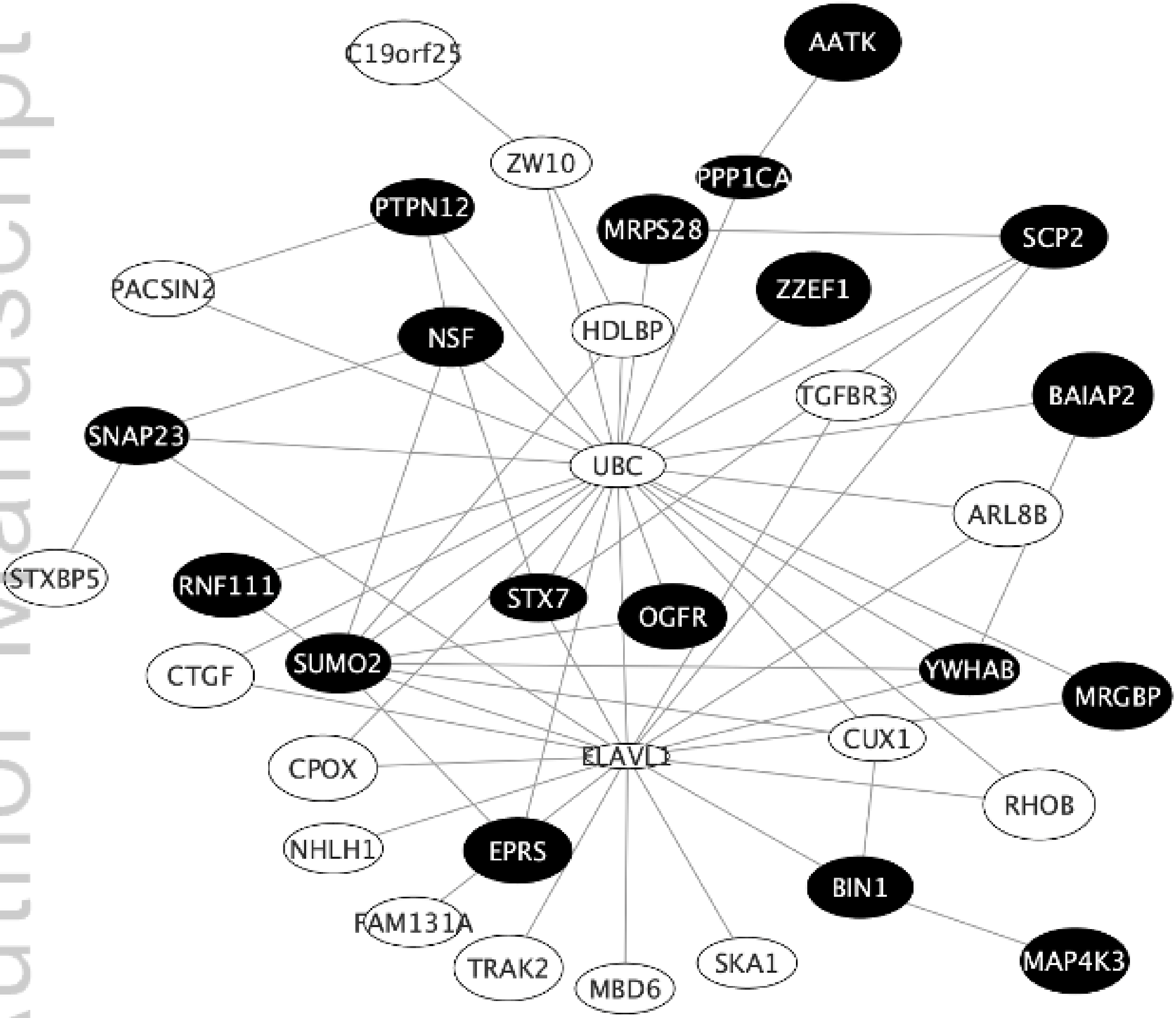

JCSM_12782_Figure2.png




JCSM_12782_Figure3.png




JCSM_12782_Figure4.png




JCSM_12782_Figure5.png