UNIVERSITY OF CALIFORNIA

Los Angeles

Multi-Omic Analysis of COVID-19 Severity

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Science

in Physiological Science

by

Roza Lorin Kirmizi

2023

ABSTRACT OF THE THESIS

Multi-Omic Analysis of COVID-19 Severity

by

Roza Lorin Kirmizi

Master of Science in Physiological Science

University of California, Los Angeles, 2023

Professor Xia Yang, Chair

The COVID-19 pandemic was caused by the severe acute respiratory syndrome coronavirus 2 (SARS CoV-2) virus. COVID-19 has been associated with multiorgan dysfunctions, including pulmonary, cardiovascular, metabolic, neurological, and psychiatric disorders, some of which can last weeks or months after viral contagion. Individual genome-wide association study (GWAS) studies have revealed specific genetic risk loci for COVID-19 severity. However, the small sample sizes and limited replication of studies prevent a comprehensive understanding of the mechanisms underlying COVID-19 severity and complications. Here we integrated the summary statistics from European and Chinese GWAS with tissue-specific expression quantitative loci (eQTLs) from GTEx and molecular pathways using Mergeomics to identify the tissue-specific molecular networks and key drivers of COVID-19 severity.

We focused on tissues relevant to COVID complications, such as the lung, heart, blood, brain, intestine, and skin. Across the datasets, we discovered replicated pathways involved in viral transcription and activity, protein processing, and cell cycle and cancer. We also predicted regulatory genes such as *RARRES3, MIF-AS1, CXCL3, RNASE3, LY86, LCP2, SLC11A1, PSMB8, GBP4, ISG15, UQCRC2, and NCKAP1L* and those for the major histocompatibility complex (MHC) Class I and II molecules. These pathways have downstream effects related to pulmonary, cardiovascular, autoimmune, skin dysfunctions, and various cancers. Identifying the pathways and regulatory genes will guide the discovery of therapeutic and pharmacological treatments for the severe illness of COVID-19 and subsequent complications.

The thesis of Roza Lorin Kirmizi is approved.

Eric Deeds

Pearl Quijada

Xia Yang, Committee Chair

University of California, Los Angeles

2023

Dedication

This thesis is dedicated to my family, as they have encouraged my academic pursuits and have always been supportive. This is also dedicated to my mom, who would have been proud of my academic accomplishments.

TABLE OF CONTENTS

LIST OF FIGURES

SUPPLEMENTARY MATERIALS

Supplementary 1: HGI A2 Study Abbreviations. The study names, abbreviations, and case/control count for all the studies in the Multi-European GWAS dataset.

Supplementary Table 2: List of all the tissue-specific eQTLs used for the MSEA and meta MSEA. Abbreviations for each eQTL are also included.

Supplementary Table 3: Top most consistent pathways for the Chinese cohort. The adaptive immune system was the most consistent, seen across 31 tissues. The topmost consistent pathways are listed with their corresponding tissues and count. The table is organized alphabetically, first by pathway, then by tissue.

Supplementary Table 4: Count of all the pathways found in the UK single MSEA run. There were a total of 465 enriched pathways. The largest pathway count was for HATs Actylate Histones.

Supplementary Table 5: The total count of the modules seen in the single tissue eQTL Italian and Spanish MSEA. There was a total of 226 enriched modules.

Supplementary Table 6: List of all the Unique and Consistent pathways across the single tissue eQTLs for the Multi European datasets. A)There were 44 unique enriched pathways and 423 consistent pathways across 15 tissue eqtls. B) When filtering for an FDR significance of 0.001, the top significant modules are seen. Abbreviations are located in Supplementary 2.

Supplementary Table 7: Total Meta MSEA results. a) Breakdown of the pathways and tissue number. b) All the consistent modules across tissues. Abbreviations for the tissues are found in Supplementary 2.

Supplementary Table 8: Results of the KDA for all 4 cohorts, the enriched tissues, and the top key drivers with their corresponding modules. Table is organized alphabetically, first by cohort, then by tissue. Abbreviations of tissues are found in supplementary 2.

ACKNOWLEDGEMENTS

I want to thank Dr. Xia Yang for agreeing to be my PI and accepting me into the lab. She has always provided guidance and had been a great help throughout my master's. I would also like to thank all the lab members, who have been very helpful and supportive, and my committee members, Dr. Eric Deeds and Dr. Pearl Quijada.

Introduction

COVID-19 is a new coronavirus disease that led to a major worldwide outbreak in 2019. It is caused by the virus known as the severe acute respiratory syndrome coronavirus 2 (SARS CoV-2) and has spread to pandemic levels. As of February 9, 2023, there have been 689,829,953 million cases worldwide and 6,885,930 million deaths. In the US, there have been 107,127,223 million cases and 1,165,538 million deaths (1).

Although most COVID-19 cases involve mild symptoms, some patients experience severe respiratory dysfunction, dyspnea, hypoxia, and multiorgan failure, leading to hospitalization, intensive care unit (ICU) admission, invasive mechanical ventilation (IMV), and sometimes death. On average, severe and critical cases make up 14% and 6%, respectively (2). The frequency of severe cases varies depending on age and comorbidities such as: age 65 or older, immunosuppression, diabetes, obesity, hypertension, as well as chronic kidney, cardiac, pulmonary, neurologic, and liver disease (2–5).

COVID-19 has also been found to affect multiple organs or tissues both acutely and chronically, with certain symptoms ongoing for weeks or months after viral infection. The SARS CoV-2 binds to the angiotensin-converting enzyme 2 (ACE2) receptor, which is present in tissues including, but not limited to, the lungs, heart, brain, liver, kidneys, intestines, and vascular endothelial cells. Once the virus gains access to the body, it triggers the innate immune system, and in turn leads to the activation of inflammatory responses as part of the antiviral response. However, both severe cases and cases that lead to mortality from COVID-19 show an abnormal immune response where acute pro-inflammatory cytokines (including IL-6, IL-1, TNF- αm, and interferon) lead to an

unchecked circulation of T-Cells, neutrophils, and macrophages. This is called the "cytokine storm" and eventually causes damage to the tissues in the vicinity of the immune response through the destruction of endothelial cell-to-cell interactions, and thereby the vascular barrier, capillaries, and situations of diffuse alveolar damage (DAD) (6). This extreme inflammatory response can lead to acute respiratory distress syndrome (ARDS), systemic inflammatory response syndrome, vasoconstriction, coagulation, edema, deep vein thrombosis (DVT), kidney injury, gastrointestinal issues, heart injury, stroke, multiorgan damage, and even death (3,7,8).

Genome-wide association studies (GWAS) have been conducted in numerous human populations to search for genetic risk factors that may confer susceptibility to severe COVID. These genes play a wide range of roles in cilia dysfunction (*TYK2, DNAH7*), cardiovascular disease (*DES, SPEG*), thromboembolic disease (*STXBP5*), mitochondrial dysfunction (*CLUAP1 and TOMM7*), and the innate and adaptive immune system (*IFNAR2, OAS1, OAS2, and OAS3*)(9). These studies have brought insights into the genetically vulnerable populations and the associated mechanisms that the target genes typically work within, which may partially explain the determining factors in COVID-19 severity and the downstream complications.

Despite the insights gained from GWAS, multiple limitations in these studies hinder the interpretation and translational use of the findings. First, the existing GWAS suffer from small sample size, the top loci and candidate genes identified from these studies have limited replication, and it is unclear through which tissues the top candidate genes act. We hypothesize that, like many other complex diseases, COVID-19 severity involves numerous genetic loci that are far beyond a few to

genome-wide significant loci in individual studies, and that these genetic risks converge on coherent pathways and networks in multiple tissues and organ systems to together confer disease severity. To test this hypothesis, we integrated independent GWAS from multiple populations with tissue-specific gene regulation data to identify the convergent and population-specific pathways or networks affected by genetic risks of COVID-19 severity in individual tissues. We further utilize a network approach to identify key drivers that likely regulate the severity of COVID-19 illness.

Material and Methods

Overview of Study Design

An integrative genomics approach, specifically Mergeomics, was utilized to analyze multiple large-scale human genetic and genomic datasets to highlight the genetic pathways, networks, and regulators of COVID-19 illness severity (Figure 1)(10–12). The datasets used included the summary statistics of four GWAS from Chinese, Italian and Spanish, UK, and multi-European populations taken from the NHGRI-EBI GWAS Catalog or the Human Genome Initiative (HGI) (13,14). Additionally, to provide functional relevance in a tissue-specific manner to the above GWAS datasets, functional genomic data in the form of expression quantitative trait loci (eQTLs) were utilized from 35 diverse human tissues and immune cell types, to map GWAS SNPs to genes in a tissue/cell specific manner (15,16). The step in the Mergeomics pipeline to achieve this mapping is called Marker Dependency Filtering (MDF), which in addition to SNP-gene mapping also filters and removes redundant SNPs from each GWAS dataset that has high linkage or a high level of redundancy ($r^2>0.7$). To be able to gain high level functional insight to the

mapped genes, we utilized an array of canonical pathways derived from resources such as KEGG, Biocarta and Reactome, whereby we input tissue-specific genes mapped from each GWAS into the Marker Set Enrichment Analysis (MSEA), to identify the canonical pathways that are enriched for stronger genetic associations with COVID severity from a genetic standpoint in each tissue. After MSEA is conducted for each population, all four populations were meta-analyzed to conduct a Meta-MSEA, which informs on the shared pathways/modules that are significant to severity of COVID-19. Lastly, to pinpoint the key regulators of such biological pathways in a tissue-specific manner, we utilized tissue specific Bayesian gene regulatory networks, which we input along with our significant biological pathways (FDR<0.05), to highlight the what we term “key driver (KD) genes” in the Key Driver Analysis (KDA) step of Mergeomics.

GWAS Data Collection

Chinese GWAS

This dataset was mined from the NHGRI-EBI GWAS Catalog (13) from study *GCST90014052* (17). The reported trait being studied was severe COVID-19 infection, which was defined using the guidelines for COVID-19 diagnosis and treatment (Trial Version 7) released by the National Health Commission of the People’s Republic of China based on the course of illness (18). Participants ranged in age from 19 to 70 years old (mean ± SD, 46.7 ± 13.5) and selected based on exhibiting no comorbidities, which were classified as hypertension, coronary heart disease, diabetes, chronic obstructive pulmonary disease, malignancy, surgical history, chronic kidney

disease, cerebrovascular disease, immunodeficiency disease, chronic hepatitis, and tuberculosis. The population included 65 Chinese ancestry cases, and 138 Chinese ancestry controls from Wuhan, China.

Whole Genome-wide sequencing of the 203 samples was conducted using the MGI2000 PE100 platform with 100-bp paired-end reads. The data was processed using Sentieon Genomics and then mapped to the reference genome hg38 using the BWA algorithm. After duplicate marking, Indel realignment, and base quality score recalibration (BQSR), per-sample variants were called using the Haplotyper algorithm in genomic VCF (GVCF) mode. Cohort Variant Call Format (VCF) was produced with the GVCFtyper algorithm, followed by variant quality score recalibration utilizing Genome Analysis Toolkit by selecting the truth-sensitivity-filter-level to 99.0 for both SNPs and Indels. Variants used for further analysis were selected depending on a PASS flag and quality score ≥ 100. The summary statistics for the GWAS dataset was that of 5,056,041 SNPs.

Italian and Spanish GWAS

This dataset was mined from the NHGRI-EBI GWAS Catalog (13) and is the analysis of severe COVID-19 infection with respiratory failure taken from study *GCST90000255* (19). The study recruited 1980 patients with severe COVID-19, defined as hospitalization with respiratory failure in the ICU, and confirmed COVID-19 across pandemic epicenters in Italy and Spain. Whole blood samples were taken for DNA extraction. 2381 control participants from Italy and Spain were recruited as controls.

DNA extraction was performed using a Chemagic 360 (PerkinElmer). Global Screening Array (GSA), version 2.0 (Illumina), was used for genotyping. Imputation was performed

on genome build GRCh38 using the Michigan Imputation Server and 194,512 haplotypes generated by the Trans-Omics for Precision Medicine (TOPMed) program. After quality control, the final datasets consisted of 835 patients and 1255 control participants from Italy and 775 patients and 950 control participants from Spain. In total, there were 3815 European individuals from Italy and Spain (1,610 European ancestry cases, 2,205 European ancestry controls). A total of 8,965,091 SNPs were included in the Italian cohort and 9,140,716 SNPs in the Spanish cohort. Statistical analysis involved the use of the PLINK logistic-regression framework for dosage data (PLINK, version 1.9). Genome-wide association tests were conducted with adjustments to control for potential population stratification and an additional analysis that corrected for age and sex. The meta-analysis tool METAL was utilized to conduct a fixed-effects meta-analysis with the use of 8,582,968 variants that were common to both the Italian and Spanish data sets.

UK GWAS

This dataset was mined from the NHGRI-EBI GWAS Catalog (13) and is taken from the study *GCST90013414* (9). The dataset studied individuals from the UK with critical COVID-19 illnesses. The population included 10056 European (U.K.) with 1,676 European ancestry cases and 8,380 European ancestry controls. Initially, the cohort included 2,636 COVID-19 patients recruited to the GenOMICC study. An additional 135 patients were recruited through the International Severe Acute Respiratory Infection Comprehensive Clinical Characterisation Collaboration (ISARIC4C). Both studies had individuals with confirmed COVID-19 that were in the ICU requiring continuous cardiorespiratory monitoring. Genotyping was performed using the Illumina Global

Screening Array v.3.0 and Illumina NovaSeq 6000 whole-genome sequencing (WGS). Samples were aligned using the HG38 human reference genome and converted to GVCF using the DRAGEN pipeline. Variants were called with the GATK GenotypeGVCFs tool, filtered to minimum depth 8× (95% sensitivity for heterozygous variant detection) merged and annotated with allele frequency with bcftools. Quality control was conducted using PLINK and data were filtered using an exclusion of samples with a call rate of <95%, selection of variants with call rate of >99% and MAF > 1% and the final selection of samples using a call rate of >97%, leaving 2,790 individuals and 479,095 variants for further analyses. Kinship and ancestry inferences were calculated following the UK Biobank and 1 Million Veteran Program using GenomeStudio. This left 2,718 individuals considered unrelated up to the third degree. Unrelated individuals from the 1,000 Genome Project dataset were calculated using the same steps as the kinship analysis and both datasets were merged using the common SNPs. Next, the data was pruned using PLINK with a window of 1,000 markers, a step size of 50 and a r2 of 0.05, giving an output of 92,017 markers. Imputation was calculated using the TOPMed reference panel. The imputed dataset was filtered for monogenic and low imputation quality scores (r2 < 0.4) using BCFtools 1.9. To perform GWAS analyses, files in VCF format were further filtered for r2 > 0.9 and converted to BGEN format using QCtools 1.358. UK Biobank imputed variants with imputation score >0.9 and overlapping the set of 5,981,137 variants were extracted and merged with GenOMICC data into a single BGEN file containing cases and controls using QCtools 1.3. The GWAS summary statistics include 4,700,000 SNPs.

Multi-European GWAS

This data set was taken from the COVID-19-hg GWAS meta-analyses round 5 (A2 all) (14). The large dataset compromised a compilation of European datasets (Supplementary Table 1). This dataset also utilized GenOMICC, however, it only consisted of 1676 participants with severe COVID-19 from a later date that wasn't ancestry matched. The population included individuals with very severe respiratory confirmed COVID-19 from countries around Europe. Very severe respiratory confirmed COVID-19 was defined as hospitalized laboratory-confirmed SARS-CoV-2 infection with the primary reason for admission being COVID-19 and patients being on respiratory support, on either intubation, continuous positive airway pressure (CPAP), or bilevel positive airway pressure (BiPAP). The population included 5,101 individuals with severe COVID-19 and 1,383,241 European ancestry controls. Processing of genotyping data used the Ricopili pipeline (14,20). GWAS analyses were run using Scalable and Accurate Implementation of GEneralized mixed model (SAIGE) on chromosomes 1–22 and X. Other software such as PLINK was also used. Study-specific summary statistics were processed for meta-analysis. Each GWAS was examined for potential false positives, inflation, deflation, and allele frequency plots. Standard error values were utilized to find studies that deviated from the expected trend. These summary statistics that passed manual inspection were then utilized in the meta-analysis. Variants with an allele frequency of >0.1% and an imputation score of >0.6 were carried forward from each study. Variants and alleles were lifted over to genome build GRCh38 and adjusted to gnomAD 3.0 genomes by finding matching variants by strand flipping or switching the ordering of alleles. If multiple matching variants were included, the best match was chosen according to the minimum fold change in absolute allele frequency. Meta-analysis was performed using the inverse-

variance-weighted (IVW) method on variants that were present in at least two-thirds of the studies contributing to the phenotype analysis. The GWAS summary statistics contained 9,856,860 SNPs.

Marker Dependency Filtering (MDF) and Mapping SNPs to genes

Summary statistics from each GWAS was reformatted using the chromosome and base pair information from the HG38 and HG19 reference genome to find SNP IDs (RSIDs). RSIDs and a -log10 of the P values were inputted into the Mergeomics pipeline to correct for linkage disequilibrium (LD) between GWAS SNPs using marker dependency filtering (MDF). A marker LD dependency file is selected depending on the population to correct for marker dependencies and remove redundant SNPs that had LD of r2>0.7 with a selected SNP. European populations are run with a marker LD dependency file of CEU LD70 (Utah residents with Northern and Western European ancestry from the CEPH collection), and the Chinese population with marker LD dependency file CHB LD70 (Chinese Han, Beijing). This was completed by GWAS SNPs being compared against other SNPs for LD and COVID-19 Severity association. When SNPs with a LD of r2>0.7 are compared with other SNPs, selection is dependent on the strongest association to COVID-19 Severity. A repetition of this step was conducted until the rest of the SNPs were no longer in LD based on the r2>0.7 cut-off. The output of the non-redundant SNPs was utilized for downstream analyses. To gain tissue specific insight to gene regulation, we used expression quantitative trait loci (eQTLs) curated from 35 tissues from the GTEx consortium version 9 eQTLs that were combined into one immune cells tissue from the DICE-DB 1 database (16). The immune cells used were B cells (naïve), monocytes

(classical and non-classical), NK cells (CD56dim CD16+), and T cells (CD4, memory TREGCD4; naïve; CD4, naïve [activated]; CD4, naïve TREG; CD4, TFH; CD4, TH1; CD4, TH1/17; CD4, TH17; CD4, TH2; CD8, naïve; CD8, naïve [activated])(16). Thirty-five different tissues were tested to induce tissue specificity. We mapped the SNPs from the different cohorts to all available tissue/cell-type eQTL sets to retrieve tissue/cell-specific SNP to gene mapping based on the regulatory associations between SNPs and genes in each tissue or cell type (Supplementary Table 2). Both LD filtering and eQTL mapping were completed within the MDF step of the analysis.

Marker Set Enrichment Analysis (MSEA)

To highlight pathways affected by SNPs associated with COVID-19 illness severity within each GWAS and each tissue/cell type, the output file from MDF was used as input into MSEA to assess the enrichment of the LD-filtered COVID-19 GWAS SNPs through MDF within each biological pathway for stronger association to COVID-19 severity compared to SNPs mapped to random sets of genes of equal gene number as the given pathway. The pathways used were from the KEGG, Reactome, and BioCarta databases (21–23). The results from MSEA informed on the critical biological processes of COVID-19 severity. MSEA utilizes a chi-square-like enrichment statistic at multiple quantile thresholds and summarizes enrichment statistics across quantiles to ensure stability of the enrichment scores. Ten thousand permuted gene sets were generated for each pathway. Enrichment statistics from the permutations step were used to approximate a Gaussian null distribution, and enrichment P-values for each pathway were determined based on the true enrichment statistics of the pathway against its corresponding null

distribution from permuted random gene sets. Benjamini-Hochberg (BH) false discovery rate (FDR) was then estimated across all pathways. Statistical significance was determined at an FDR < 5%.

Meta-MSEA

To search for consistent biological pathways across GWAS datasets we utilized the meta-MSEA function of Mergeomics. Meta-P values were calculated using Stouffer's Z score method and were based on the P-values from MSEA runs for individual GWAS cohorts. Meta-FDR was calculated using the Benjamini-Hochberg method.

Key Driver Analysis (KDA)

To identify potential regulators of disease-associated pathways and networks, a key driver analysis (KDA) was conducted. All of the significant pathways from the MSEA or Meta-MSEA were inputted into the KDA. MSEA results are overlaid with tissue-specific Bayesian Networks that were selected based on matching tissue type with that of the tissue used in the MDF. KDA utilizes a chi-square like statistic to find top key drivers which are linked to a significantly larger number of disease-associated pathway genes than compared to the expected number for a randomly selected gene within the Bayesian network. This was done by overlaying the tissue-specific severe COVID-19 associated genes with the corresponding tissue Bayesian molecular network to see if a particular subnetwork was enriched for the pathway genes associated with severe illness. The edge weights (representing the reliability measures of network edges) that linked a hub to its neighbors are summarized into node strengths within the vicinity excluding the hub itself.

Statistical significance of the disease-gene enriched hubs was estimated through permutation of the gene labels 10000 times and approximating the P-value dependent on the null distribution. An FDR < 5% was used to find top KDs. Bayesian tissue networks utilized were for adipose, blood, liver, brain, skin, colon, heart, and arteries (15,24–29). Network visualization was conducted using Cytoscape.

Results

Tissue-specific MSEA identified pathways enriched for COVID-19 severity SNPs within and across-tissues

To understand the contribution of individual tissues to severe COVID-19 outcomes, we ran MSEA for each tissue and cell-type with available eQTLs to link GWAS SNPs to tissue-specific gene expression. We found pathways unique to specific tissues and pathways that are more system-wide in that they are commonly shared across tissues within each GWAS cohort.

MSEA for the Chinese cohort identified the importance of adipose tissue

The Chinese cohort had 593 enriched pathways (133 enriched pathways when /duplicates are removed) with an FDR<0.05 (Figure 2 and Figure 3). There were significant pathway results across all 35 tissues tested, highlighting a system-wide tissue contribution to COVID-19 severity, with the largest number of pathways (102) identified from the adipose subcutaneous tissue (Figure 3). Numerous pathways were consistently identified from adipose tissue (subcutaneous and visceral omentum), whole blood, artery tibial, and esophagus mucosa, tissues reflecting the immune system and endothelial

functions. These pathways ranged from cell signaling and processes to adaptive immune system, inflammatory response, protein synthesis, transcription/translation, autoimmune responses, and viral infection (Supplementary 3). The most consistent pathways across tissues in terms of significance were those relating to immune system activation and signaling, cell adhesion molecules (CAMs), and Type I diabetes mellitus (highlighting autoimmunity), with the pathway "adaptive immune system" appearing in 31 tissues (Figure 3).

There were 34 unique tissue-specific pathways seen for 14 tissues. The most significant of these unique pathways was for the "the role of GTSE1 in G2/M progression after G2 checkpoint" pathway (involved with the cell cycle) (FDR= 5.39E-05) and the "MAPK6 MAPK4 signaling" pathway (involved in cell motility and tumor suppression) (5.35E-05), both from the adipose subcutaneous (Table 1).

MSEA for the UK cohort identified the importance of the esophagus and artery tissues

The MSEA analysis of the UK dataset yielded 465 significantly enriched pathways across 32 tissues (123 enriched pathways when duplicates are removed) (Figure 2 and 4). Overall, the pathways are involved in viral infection, leishmania infection, chemokine signaling, innate immune system, adaptive immune system, vesicle transport, RHO GTPase activity, cell signaling, cell death, reproductive processes, phototransduction pathway, neuronal/synaptic activity, and long-term depression (Supplementary 4). The most significant pathway was for "COPII mediated vesicle transport" identified when using the esophagus mucosa tissue (Table 2).

There were 14 unique (only seen once and not including duplicated pathways) pathways seen in the artery (tibial), immune cells, brain (cerebral hemisphere and cerebellum), and cultured fibroblasts. These modules are for antimicrobial peptides, fertilization, glioma, HIV infection, the MEF2D pathway, peptide ligand binding receptors, reproduction, RHO GTPase effectors, immune response, the phototransduction cascade, and hormone, light, and neurotransmitter receptors. The most significant unique module was for the “phototransduction cascade” from cultured fibroblasts (FDR= 0.004280066) (Table 3).

There were 49 consistent pathways across the enriched tissues, reaching 452 pathways in total for each tissue (when aren’t removed). The most consistent pathway was “Hats acetylate histones” identified when eQTLs from immune cells, brain (cerebellum and cerebral hemisphere), colon (sigmoid and transverse), lung, EBV transformed lymphocytes, small intestine terminal ileum, whole blood, cultured fibroblasts, artery tibial, adipose (visceral omentum and subcutaneous), and skin were used (Supplementary 4).

MSEA for the Italian and Spanish GWAS implicated the importance of digestive, brain, and heart tissues

The MSEA for the Italian and Spanish GWAS dataset identified 226 enriched pathways (92 enriched pathways when duplicates are removed) across 25 tissues (Figure 2 and 5). The top 5 tissues with the most enriched modules were the esophagus muscularis (31 pathways), cerebellum (CER), heart atrial appendage (HAA), nucleus accumbens basal ganglia (NABG), and cerebellar hemisphere (CH) (Figure 5). Over all the tissues, there were 226 enriched pathways in total. These pathways involve cell processes, hormones,

neurotransmitters, transcription, translation, and immune processes (Supplementary 5). The most consistent pathways were the WNT signaling pathway, activation of NMDA receptors and postsynaptic events, GPCR ligand binding, melanogenesis, and chromatin modifying enzymes. The WNT signaling pathway and activation of NMDA receptors and postsynaptic events were seen across eight different tissues (Supplementary 5). There were 3 unique pathways, which included cargo concentration in the ER (skin exposed lower leg), phototransduction cascade (cells cultured fibroblasts), and cell cycle mitotic (substantia nigra) (Table 4).

MSEA for the multi-European cohort showed the importance of synaptic signaling and brain and digestive tissues

The MSEA analysis of the multi-European dataset yielded 467 enriched pathways (115 enriched pathways when duplicates are removed) across 34 tissues (Figure 2 and Supplementary 6). These pathways were involved with the adaptive immune system, innate immune system, cytokine signaling, viral infection, Leishmania infection, neuronal/synaptic signaling, cell death, cell cycle, RHO GTPases, tumor suppression, DNA repair, vesicle transport, Alzheimer's disease, phototransduction pathways, and reproduction (Supplementary 6).

There were 73 consistent pathways seen over multiple tissues (Supplementary 6). The top 5 most consistent pathways are those of GPCR ligand binding, HATs acetylate histones, neurotransmitter receptors and postsynaptic signal transmission, transmission across chemical synapses, and chromatin modifying enzymes. There were 44 unique enriched pathways observed for tissues including brain (amygdala and cerebellum),

esophagus (muscularis), cultured fibroblasts, skin (sun-exposed lower leg), adipose (visceral, omentum, and subcutaneous), whole blood, and EBV-transformed lymphocytes.

Comparison across MSEA results for all cohorts highlights consistent and unique tissue and pathway involvement

Across all populations and tissues, there were 20 tissues that exhibited pathway enrichment across all populations, including adipose (visceral omentum), artery (aorta, coronary, and tibial), brain (caudate basal ganglia, cerebellar hemisphere, cerebellum, hippocampus, and nucleus accumbens basal ganglia), cultured fibroblasts, EBV-transformed lymphocytes, colon (sigmoid and transverse), esophagus (muscularis), heart (atrial appendage and left ventricle), lung, immune cells, skin (not sun exposed suprapubic and sun exposed lower leg), and small intestine terminal ileum. These results implicate the multi-tissue involvement in COVID-19 severity. The tissues that had the most pathways enriched were for esophagus muscularis (178 pathways), visceral omentum adipose tissues (149 pathways), and subcutaneous adipose tissue (142 pathways). However, when using subcutaneous adipose eQTLs, no enriched pathways were identified for the Italian and Spanish population.

Across all 4 populations, there were 9 consistent pathways for apoptosis, HIV infection, HIV life cycle, interferon signaling, Leishmania infection, neuronal system, programmed cell death, signaling by WNT, and TCF dependent signaling in response to WNT (Figure 2), highlighting the importance of these pathways in COVID-19 severity across cohorts.

We also found 56 pathways unique to individual populations. There were 15 for the Chinese population, namely 5 pathways in adipose (subcutaneous and visceral omentum), 2 in brain (frontal cortex ba9 and cerebellum), 1 in esophagus mucosa, 1 in heart left venticle, 1 in skin not sun exposed suprapubic, 3 in whole blood, and 2 in lung. There were 11 unique enriched pathways for the Italian and Spanish dataset, namely 6 in the brain (cerebellum, hypothalamus, nucleus accumbens basal ganglia, cerebellar hemisphere), 1 in artery aorta, 2 in heart (left ventricle and atrial appendage), and 2 in esophagus muscularis. There were 18 unique enriched pathways for the UK dataset, including 1 in the adipose visceral omentum, 2 in the artery tibial, 13 in the esophagus (muscularis and mucosa), and 2 in immune cells. There were 12 unique enriched pathways for the multi-European dataset, including 1 in the adipose subcutaneous, 3 in the amygdala, 1 in cultured fibroblasts, 6 in esophagus muscularis, and 1 in whole blood (Table 7).

Meta-MSEA identified pathways enriched for COVID-19 severity SNPs across cohorts

Meta-MSEA was used to further assess consistencies across cohorts, yielding 741 modules across 26 tissues that were significantly enriched when all cohorts were considered. The tissues with the most enriched pathways were the cerebellum, aorta, esophagus muscularis, and cultured fibroblasts (Figure 6). Enriched modules included SARS CoV 2 infection, HIV infection, adaptive and innate immune processes, anti-immune response pathways, cytokine and interleukin signaling, auto-immune response, Alzheimer's disease, asthma, cancer, cell adhesion molecules, cardiomyopathy, cell

signaling, Leishmania infection, long term depression, neuronal processes and signaling, vesicular transport, cell membrane transport, cell death, fatty acid synthesis, DNA repair, tumor suppression, and phototransduction (Supplementary 7).

The tissue eQTL with the most enriched pathways was the cerebellum, which exhibited 58 enriched pathways involved in SARS CoV 2 infection, innate and adaptive immune response, and interferon and chemokine signaling, vesicle transport, synaptic signaling, cell life cycle and death, epigenetic processes, Alzheimer's Disease, GTPases pathways, auto-immune processes, antigen processing and presentation, and neuronal processes (Figure 7).

There were 103 consistent pathways across tissues. These pathways are involved in SARS CoV 2 Infection, HIV infection, antigen presenting pathways, autoimmune processes, intestinal immune network for IGA production, chemokine signaling, epigenetic processes, secretin-receptor processes, vesicle transport, synaptic signaling, neuronal activity, long-term depression, GTPase activity, phototransduction pathways, water homeostasis, sensory perception, DNA repair, recombination, cell death, cell signaling, tumor suppression, melanogenesis, and cancer (Supplementary 7). The most consistent pathway was that of HATS acetylate histones, which was identified across 15 tissues, with the digestive, brain, and immune tissues showing the highest significance (Figure 8).

There were 35 significant pathways across tissues that were unique to individual tissues, such as the CCR3 pathway when the whole blood eQTLs were used, SLC transporter disorders from the esophagus muscularis eQTLs, post-translational protein modification, infectious disease, and asthma when using brain eQTLs (Figure 9).

Key driver analysis (KDA) identified potential key drivers (KDs) of COVID-19 severity pathways within and across-tissues

KDs for the COVID-19 pathways from the Chinese Cohort

KDA for the significant pathways from the Chinese population identified a total of 81 KDs for artery (aorta, tibial), adipose (subcutaneous, visceral omentum), colon (sigmoid), and the liver. The consistent top KDs across tissues included *CD74, HLA-B, HLA-DMA, HLA-DPB1, HLA-DQB1, HLA-DRA, HLA-DRB1, HLA-DRB5*, and *MIF-AS1*.

The COVID-19 severity pathways regulated by the top KDs included interferon signaling, cell adhesion molecules (CAMs), adaptive immune system, and cytokine signaling in immune system. The tissues that had the most key drivers were the adipose subcutaneous, whose KDs included: *CD74, HLA-B, HLA-DMA, HLA-DQB1, HLA-DRB1*, and *MYO1F*, and the artery tibia, whose KDs included: *CD74, HLA-DQB1, HLA-DRA, HLA-DRB1, HLA-DRB5,* and *MIF-AS1* (Supplementary 8 and Figure 10a).

KDA for the COVID-19 severity pathways in the UK Cohort

The top KDs for the COVID-19 severity pathways identified from the UK cohort were identified for arteries (aorta, coronary, tibial), cultured fibroblasts, and esophagus gastroesophageal junction. These included *HLA-DQA1, HLA-DQB1, HLA-DRA, HLA-DRB6, MIF-AS1,* and *RP11-259G18.3*, which were KDs for pathways CAMs and interferon gamma signaling, and cytokine signaling. The tissue that had the most KDs was the esophagus, whose KDs included: *HLA-DQA1, HLA-DQB1, HLA-DRA, HLA-DRB6, MIF-AS1,* and *RP11-259G18.3* (Supplementary 8 and Figure 10b).

KDs for the COVID-19 severity pathways from the Italian and Spanish cohort

There were 29 top KDs across 5 tissues: artery (aorta), esophagus (gastroesophageal junction, mucosa, muscularis), and heart atrial appendage. Three KDs *HLA-DQA2, HLA-DRB5,* and *RP11-259G18.3* were consistent across tissues and pathways such as costimulation of *CD28* family, cell death, vesicle transport, WNT signaling, cancer, long term depression, and cell signaling. The tissue with the most KDs was heart atrial appendage whose KDs included HLA-DQA2 and HLA-DRB5 (Supplementary 8 and Figure 10c).

KDs for COVID-19 severity pathways from the multi-European cohort

We identified 29 KDs form COVID-19 severity pathways from the multi-European cohort from adipose tissues (subcutaneous, visceral omentum), fibroblasts, and heart atrial appendage. Top KDs seen consistently across multiple tissues were *RP11-259G18.3, CD74, HLA-B, HLA-DQB1, HLA-DRA, HLA-DRB1, IRF7, IRF9,* and *OAS2*. These KDs orchestrated genes in the following pathways enriched for COVID-19 severity SNPs: adaptive immune system, interferon/cytokine signaling, water homeostasis, vesicle transport, synaptic/postsynaptic signaling, and visual phototransduction. The tissue with the most KDs was the esophagus whose KDs included: *RP11-259G18.3, CD74, HLA-B, HLA-DQB1, HLA-DRA, HLA-DRB1, IRF7, IRF9,* and *OAS2* (Supplementary 8 and Figure 10d).

Across data visualizations for the Esophagus (Gastroesophageal Junction and Mucosa), Artery (Aorta and Tibial) and Adipose (Subcutaneous and Visceral Omentum), HLAs were consistent KDs

The KDA of the meta-analysis of all 4 cohorts showed 14 tissues with 29 KDs (Table 8). Of these the esophagus (gastroesophageal junction and mucosa), artery (aorta and tibial) and adipose (subcutaneous and visceral omentum) had the most KDs.

Next, we visualized the top KDs for the immune system from the following tissues: the esophagus (gastroesophageal junction and mucosa), arteries (aorta and tibial) and adipose (subcutaneous and visceral omentum).

In the esophagus, top KDs including MHC class II genes *(HLA- DQA1, DQB1, DRB6, DRA), RP11-259G18.3, MIF-AS1,* and *IGHV1-56* were visualized. Although the HLA genes were highly connected in the network, *RP11-259G18.3, MIF-AS1,* and *IGHV1-56* were KDs in their own subnetworks. *IGHV1-56* was also a KD unique to esophagus mucosa. (Figure 11)

For the two artery tissues (tibial and aorta), top KDs included MHC class II molecules (*HLA-DRB1, DQB1, DRA, DPB1*), *CD74, MIF-AS1, RP11-259G18.3, CXCL3,* and *RARRES3*. KDs *HLA, CD74,* and *RARRES3* were connected, while *MIF-AS1, RP11-259G18.3*, and *CXCL3* were KD in their own subnetworks. *RARRES3* was only present in the aorta artery network, while *MIF-AS1* and *CXCL3* were only seen in the tibial artery network (Figure 12)

In the visceral omentum adipose network, MHC class II genes (*HLA- DQA1, DQB1, DRB1, DMA*), MHC class I gene *HLA-B, RNASE3, CD47, LY86, LCP2*, and *SLC11A1* were all key drivers and all highly connected. All these SNPs shared networks. The KD

unique to this tissue was *LCP2*. The adipose KD network also contained KDs which were genes reported as the GWAS loci in the original GWAS summary statistics: GWAS gene *CCR2* was connected to KD *RNASE3* and GWAS gene *HLA-G* was connected to KD *HLA-B* in the visceral omentum adipose network (Figure 13).

The subcutaneous adipose tissue had the largest KD subnetwork, containing 12 KDs, including those for MHC class II genes (*HLA- DQA1, DQB1, DRB1, DMA, DMB*), MHC class I gene (*HLA-B*), *CD47, PSMB8, GBP4, ISG15, UQCRC2*, and *NCKAP1L*. *UQCRC2* was the only KD in its own network, while the other KDs shared networks. Of the KDs, the ones that were not seen in other tissues were for *PSMB8, GBP4, ISG15*, and *UQCRC2*. There were also nodes that were reported to exhibit significant associations in the original GWAS: *OAS1, OAS2* (nodes connecting to KD *ISG15*), *CCR2* (node connecting KD *HLA-DMB*), and *HLA-G* (node connecting to KD *HLA-B*). *IFNAR1*, which is in the same family as a previously reported GWAS gene *IFNAR2*, is also in the subcutaneous adipose top KD network (Figure 14).

Discussion

The four severe COVID-19 GWAS datasets had previously reported GWAS hits implicating diverse processes. The UK dataset reported GWAS loci associated with antiviral activity (*IFNAR2, OAS1, OAS2*, and *OAS3*), and inflammatory response (*TYK2, DPP9, CCR2*). The Italian and Spanish datasets reported genes involved in the interaction with the SARS-CoV-2 cell-surface receptor (*SLC6A20*), intracellular cargo trafficking (*LZTFL1*), microtubule transport of autophagosomes (*FYCO1*), immune response (*CCR9, CXCR6 and XCR1*), and the ABO blood group. The Chinese data set

from Wuhan had reported genes important for immune response (*FOXP4-AS1*) and the ABO blood group. The multi-European dataset reported SNPs for intracellular cargo trafficking (*LZTFL1*), immune response (*SLC6A20*), ABO, and antiviral activity (*IL10RB, IFNAR2 and OAS1*) (11). The Italian and Spanish study showed association signals spanning the ABO blood group locus, with blood-group-specific analysis showing risk being higher in blood type A compared to other blood groups, with a protective effect in blood type O compared to other types. This study also found association signals spanning *SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, and XCR1*(12). Another Chinese GWAS analysis study found loci associated with COVID-19 severity on 11q23.3 and 11q14.2. These loci are involved in inflammation and immune cell functions and include genes *CTSC, CADM1, REXO2, and ZBTB16* (13). Across these GWAS, *LZTFL1, CCR9, IFNAR2, CXCR6, SLC6A20, OAS1, OAS2*, and *OAS3* were replicated in more than one cohort, whereas the others were not.

As genome-wide loci are highly dependent on population size and only capture very limited genetic heritability of complex diseases or traits like COVID-19 severity, here we carried out a pathway/network-based analysis that is complementary to the original GWAS that considers individual SNPs. This approach has been previously reported to reveal more comprehensive biological insights and uncover more missing heritability. Our pathway/network-based analysis for the four GWAS studies of COVID-19 severity recapitulated the importance of both the innate and adaptive immune responses, signaling for cytokines, chemokines, interferons, antigen-presenting pathways, and anti-inflammatory response. It is believed that the severity of COVID-19 is linked to how the immune system responds to the virus. The immune system responds to COVID-19 by

activating TLR receptors which in turn turns on IFN production, activates antigen presentation and eventually can lead to uncontrolled inflammation at later stages of disease. There is also a response with chemokines, leading to the activation of T cells which eventually causes an increase in IFN, leading to the activation of pro-inflammatory macrophages. All of these accelerate the severity and course of the disease (30).

In addition to inflammatory pathways, we also found autoimmune response pathways, seen in pathways like allograft rejection, type 1 diabetes mellitus, etc. It has been reported that autoimmune diseases tend to exhibit an exaggerated immune response called the cytokine storm. Cytokine storms are also seen in COVID-19 patients with severe illness (31). There has also been previous research that has investigated the autoimmune response and its association with COVID-19, where some individuals develop autoimmune diseases, including type 1 diabetes mellitus, systemic lupus erythematosus, and rheumatoid arthritis (32,33).

We also uncovered pathways related to antimicrobial, parasitic, and viral infection, more particularly SARS-CoV-2, HIV, and Leishmania infection. HIV and Leishmania are different diseases from COVID-19; however, they share some similarities regarding the immune response. Leishmania infection is characterized by an immune response of uncontrolled activation of T lymphocytes, natural killer cells and macrophages, and the possibility of increased levels of proinflammatory cytokines (34). HIV, on the other hand, seems to share common mechanisms in immune response with COVID-19. Both viruses cause similarities in the pro-inflammatory cytokine response (cytokine storm),

and T lymphocyte deficiency seen with high expression of T cell inhibitory receptors, including programmed cell death-1 (PD-10) (35)

In addition to the immune/inflammation pathways that agree with the original GWAS discoveries, our pathway analysis also uncovered additional cell processes such as vesicle transport, water homeostasis, cell signaling, cell cycle/death, and cell transport. Pathways involved with DNA/RNA, such as translation, transcription, and epigenetics, were also found. Previous studies have found alterations in DNA methylation as individuals with severe COVID-19 progress through illness, particularly in the “apoptotic execution pathway” (36). These essential cellular processes may be important for cell survival and therefore affect COVID severity.

Surprisingly, our analysis also revealed pathways involved with neuronal and synaptic activity, particularly with AMPA and NMDA signaling and long-term neuronal depression, which were not implicated in the original GWAS discoveries. There has been some previous work that has shown that COVID-19 affects the central nervous system (CNS) and the olfactory sensory neurons (37,38), and genetic vulnerabilities associated with CNS processes may predispose patients to more severe COVID outcomes.

Our tissue-specific pathway/network analysis also revealed tissue importance and tissue-specific processes, which could not be retrieved in standard GWAS analysis. We suggest that a multitude of tissues contribute to COVID severity, but the ones that had the most significant pathway results were for adipose, artery, esophagus, cerebellum and the heart. In particular, the use of adipose tissue eQTLs informed significant pathways across all GWAS populations. Obesity is a known risk factor for COVID-19

and dysregulated adipose tissue occurs in obesity, leading to systemic inflammation and altered immune response. There is research that shows that adipose tissue is correlated with COVID-19 severity. In lethal cases, viruses' nucleocapsid antigen has been found in a sizeable proportion of adipocytes of subcutaneous fat (39). Similarly, another study found that SARS-CoV-2 infects fat tissue and the surrounding pre-adipocytes and macrophages, which can lead to a cytokine storm (40). High levels of adipocytes could be associated with the level of severity of illness in COVID-19. The large number of immune and inflammatory pathways from adipose tissue supports the contribution of adipose tissue to COVID-19 severity through immune dysregulation.

COVID-19 has also been shown to affect the brain. Individuals who had COVID-19 were shown to have decreased grey matter volume (41). An individual with a severe case of COVID-19 exhibited cerebellar dysfunction, with brain magnetic resonance imaging (MRI) depicting edema of the cerebellar hemisphere (42). There is an apparent "invasion of the brain" in these cases, where the SARS-CoV-2 virus has been found in the brain regions of those who passed from the disease, with the olfactory bulb being the predicted entry site of COVID-19 into the brain (43,44). Our work has shown pathways involved in the olfactory senses. Furthermore, this could explain why there is a loss of smell for people with COVID-19. There is also a change in immune response, with there being a decrease in microglia cells, a pathway also revealed in our findings, in the cerebellum of individuals who have passed from COVID-19 (45).

Despite limited research on the causal association of heart, artery, or esophagus on the severe illness of COVID-19, complications in these tissues after COVID-19 severity have been reported. There is a correlation between individuals hospitalized for severe

COVID-19 symptoms and experiencing long-term sequelae, such as a cardiovascular event, arterial and venous thromboses, and esophageal hypersensitivity (46–48). It is hypothesized that due to the digestive symptoms that occur in COVID-19 illness, that the digestive systems are probably involved. Some studies have shown that angiotensin-converting enzyme 2, which is how SARS-CoV-2 gets into the body, was not only expressed in the lungs, but also in the upper esophagus, and other digestive areas like the small intestine, liver, and colon, tissues that have had enriched pathways in our results (49,50).

The esophagus also has its own immune response where esophageal epithelial cells and underlying immune cells can trigger an inflammatory response characterized by increased expression of Th2 cytokines, such as thymic stromal lymphopoietin (TSLP), IL-5, and IL-13 and activation of dendritic cells (51). Therefore, the esophagus may also play a role in disease severity through immune modulation.

When conducting key driver analysis, top key drivers were implicated in the potential regulatory roles of the MHC system (*CD74, PSMB8, HLA-B, HLA-DMA, HLA-DPB1, HLA-DQA1, HLA-DQB1, HLA-DRA, HLA-DRB1, HLA-DRB5, HLA-DMB, HLA-DQA2, HLA-DRB6, HLA-DQB1-AS1, HLA-H*), cell proliferation (*MIF-AS1*), metastasis suppression (*RARRES3*), host antiviral response (*RTP4, GPB4,* and *ISG15*), collagen production (*COL5A1*) protein processing (*FAU, RPL10A,* and *RPL3*), anti-microbial activity (*RNASE3*), mitochondrial processes (*UQCRC2*), inflammatory responses (*ALOX5AP, CXCL3, LY86, LCP2, SLC11A1,* and *DPEP2*), alcohol processing (*ADH1B*), lymphocyte processes (*IGHV1-46,* and *NCKAP1L*), amino acid transportation (*LAT2*), neurotransmitter release (*KCNK13*) and a pseudogene (*RP11-259G18.3*).

The HLA key drivers can be broken down into 2 categories: MHC class I and Class II. HLAs play an important role in the immune system and are involved in processes such as antigen presentation to T cells and acting as ligands for receptors seen on natural killer (NK) cells. MHC Class I include *HLA-A, HLA-B,* and *HLA-C*. The other HLAs belong in class II. It has been found that different alleles of HLA are associated with COVID-19 severity (52). *HLA-B, HLA-DRB1*, and *HLA-DQB1* are associated with COVID-19 susceptibility and severity (53–55). Diseases associated with HLA include MHC-peptides and autoimmune diseases, some of which were pathways significant in our MSEA results. In addition, some HLAs are associated with Alzheimer's and Parkinson diseases (56,57). The other top key drivers are involved in macrophage activity, T Cell Activity, and protein synthesis.

In our network visualization, we found the top KDs are highly connected in the esophagus, artery, and adipose tissues. Most of the KDs were HLA I or II genes, in particular, the KDs mostly fell into HLA alleles of HLA-DQ , DR, DP, and DM alleles. Other interesting KDs were involved with MHC processes (*CD74* and *PSMB8*), cell proliferation (*MIF-AS1*), metastasis suppression (*RARRES3*), anti-microbial activity (*RNase3*), inflammatory response (*CXCL3, LY86, LCP2,* and *SLC11A1*), host antiviral response (*GPB4* and *ISG15*). *CD74*, a KD replicated across tissues and COVID-19 severity pathways, is closely involved in antigen presentation in the immune response and is involved in pathways related to the innate immune system, inflammation, and cell proliferation (58). It plays a role in both MHC and non-MHC functions. It is also seen as a cell membrane receptor for macrophage migration inhibitory factor (MIF) (59).

MIF-AS1 is a MIF, which is associated with COVID-19 infection or disease severity (60). Although MIF-AS1 has not directly been found to be associated with COVID-19, it has been found to be associated with breast cancer cell proliferation along with *CD74* (58,61). Another KD linked to COVID-19 is *ISG15*, which exacerbates inflammation in SARS-CoV-2 infection (62). In the immune response of individuals with severe COVID-19, there was a higher level of inflammatory macrophages expressing *CXCL3* in comparison to moderate cases (63). In addition, *LCP2* is considered a novel biomarker for COVID-19, and *SLC11A1* is shown in studies as a possible therapeutic for COVID-19 (64,65). *LY86* has been found as a most influential gene in a study linking COVID-19 and sepsis (66). Our network analysis revealed the central roles of these KDs in regulating COVID-19 severity pathway, thereby highlighting them as potential therapeutic targets.

This study has some strengths in regard to the way the analysis is being conducted. Previous studies mainly focused on individual genetic loci and genes with limited replication and significant missing heritability due to small sample sizes, lack of tissue specific analysis, and lack of omics integration. Our study overcomes these previous limitations by focusing on tissue-specific aggregation of genetic association across cohorts thereby improving power and interpretability, and capturing a more holistic view of disease mechanism. We also acknowledge the following limitations. The datasets used were published in 2021 and are not viral variant specific. SARs-CoV2 variants, although the same virus, have different levels of disease severity and contagiousness. Thus, the results may not be an indicator for the newer variants after 2021. Future steps should involve newer COVID-19 datasets, especially those with more diversity in viral

variants and genetic diversity in human populations. Transcriptomic, metabolomic, and proteomic datasets should also be included to offer additional molecular insights, as Mergeomics allows for multi-omics data analysis.

Table 1: Unique pathways found across single tissue-specific eQTLs for the Chinese population. The most unique modules were for the adipose subcutaneous tissue. The most significant unique module was for the generation of second messenger molecules. Table is organized alphabetically, first by tissue, then by pathway.

Tissue	Pathway	FDR
Adipose Subcutaneous	APOPTOSIS	1.19E-02
	HEDGEHOG LIGAND BIOGENESIS	1.82E-04
	HEDGEHOG ON STATE	2.47E-04
	MAPK6 MAPK4 SIGNALING	5.35E-05
	METABOLISM OF AMINO ACIDS AND DERIVATIVES	4.00E-02
	MYD88 INDEPENDENT TLR4 CASCADE	3.20E-02
	POTENTIAL THERAPEUTICS FOR SARS	7.96E-03
	PROGRAMMED CELL DEATH	1.30E-02
	PROTEIN LOCALIZATION	2.85E-03
	PROTEIN PROTEIN INTERACTIONS AT SYNAPSES	1.73E-02
	REGULATION OF RAS BY GAPS	1.25E-04
	THE ROLE OF GTSE1 IN G2 M PROGRESSION AFTER G2 CHECKPOINT	5.39E-05
	TOLL LIKE RECEPTOR CASCADES	4.19E-02
	TOLL LIKE RECEPTOR TLR1 TLR2 CASCADE	1.53E-02
Adipose Visceral Omentum	DEVELOPMENTAL BIOLOGY	1.67E-02
	INFECTIOUS DISEASE	1.68E-02
	SIGNALING BY RECEPTOR TYROSINE KINASES	1.53E-02
Artery Aorta	PROCESSING OF CAPPED INTRON CONTAINING PRE MRNA	1.90E-02
Artery Tibial	ANTIGEN PRESENTATION FOLDING ASSEMBLY AND PEPTIDE LOADING OF CLASS I MHC	4.94E-02
Brain Caudate Basal Ganglia	HIV LIFE CYCLE	3.03E-02
Brain Cerebellum	FATTY ACID METABOLISM	2.38E-02
	LYSOSOME	2.05E-03
Brain Cortex	ASTHMA	8.07E-08
Brain Frontal Cortex Ba9	METABOLISM OF NUCLEOTIDES	8.51E-03

Esophagus Mucosa	TIGHT JUNCTION	1.73E-02
Heart Left Ventricle	GLYCEROPHOSPHOLIPID BIOSYNTHESIS	4.25E-03
Lung	DISEASES ASSOCIATED WITH O GLYCOSYLATION OF PROTEINS	2.76E-03
	DRUG METABOLISM CYTOCHROME P450	2.77E-03
Skin Not Sun Exposed Suprapubic	GENERATION OF SECOND MESSENGER MOLECULES	3.78E-11
	THE CITRIC ACID TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT	3.48E-02
Skin Sun Exposed Lower Leg	NEURONAL SYSTEM	2.59E-02
Whole Blood	CELLULAR RESPONSE TO HYPOXIA	5.86E-03
	DOWNSTREAM SIGNALING EVENTS OF B CELL RECEPTOR BCR	2.08E-02
	TNFR2 NON CANONICAL NF KB PATHWAY	5.23E-03

Table 2: The top most significant pathways for the UK single MSEA analysis. The most significant enriched modules were found in the esophagus mucosa, immune cells, esophagus muscularis, and cerebellum.

Tissue	Pathway	FDR
Esophagus Mucosa	COPII MEDIATED VESICLE TRANSPORT	1.67E-07
Immune	CHROMATIN MODIFYING ENZYMES	1.95E-07
Esophagus muscularis	APOPTOTIC EXECUTION PHASE	1.89E-06
Immune	HATS ACETYLATE HISTONES	2.01E-06
Esophagus Mucosa	INTRA GOLGI TRAFFIC	2.66E-06
Brain Cerebellum	COPII MEDIATED VESICLE TRANSPORT	4.02E-06

Table 3: Unique pathways found across single tissue-specific eQTLs for the UK population. The most unique modules were for the esophagus muscularis tissue. The most significant unique module was for the transport of bile salts and organic acids metal ions and amine compounds. Table is organized alphabetically by tissue. Abbreviations for tissues are in **Supplementary table 2.**

Tissue	Pathway	FDR
AVO	ANTIMICROBIAL PEPTIDES	3.13E-02
AT	TRAFFICKING OF GLUR2 CONTAINING AMPA RECEPTORS	1.06E-02

	MEF2D PATHWAY	4.16E-02
	GLIOMA	4.92E-02
CER	RHO GTPASE EFFECTORS	1.87E-02
CH	HIV LIFE CYCLE	1.33E-02
	HIV INFECTION	3.11E-02
CCF	THE PHOTOTRANSDUCTION CASCADE	4.28E-03
EGJ	FATTY ACID METABOLISM	4.12E-02
EM	CDC42 GTPASE CYCLE	5.30E-04
	RAC1 GTPASE CYCLE	1.45E-03
	ANTIGEN PRESENTATION FOLDING ASSEMBLY AND PEPTIDE LOADING OF CLASS I MHC	6.08E-03
	MHC CLASS II ANTIGEN PRESENTATION	6.79E-03
	SARS COV 2 INFECTION	8.33E-03
	CARGO CONCENTRATION IN THE ER	9.79E-03
	DISEASES OF DNA REPAIR	1.06E-02
	INSULIN SIGNALING PATHWAY	2.27E-02
	HDR THROUGH SINGLE STRAND ANNEALING SSA	2.29E-02
	SIGNALING BY RHO GTPASES MIRO GTPASES AND RHOBTB3	2.42E-02
	REGULATION OF TP53 ACTIVITY THROUGH PHOSPHORYLATION	2.43E-02
	CLASS I MHC MEDIATED ANTIGEN PROCESSING PRESENTATION	2.82E-02
	DOWNREGULATION OF TGF BETA RECEPTOR SIGNALING	2.84E-02
	SARS COV 2 ACTIVATES MODULATES INNATE AND ADAPTIVE IMMUNE RESPONSES	3.48E-02
	FOCAL ADHESION	3.50E-02
	OOCYTE MEIOSIS	3.54E-02
	TRIGLYCERIDE METABOLISM	3.55E-02
	HOMOLOGOUS DNA PAIRING AND STRAND EXCHANGE	4.01E-02
	TGF BETA RECEPTOR SIGNALING ACTIVATES SMADS	4.02E-02
	LONG TERM POTENTIATION	4.30E-02
	SIGNALING BY TGF BETA RECEPTOR COMPLEX	4.88E-02
EMUSC	TRANSPORT OF BILE SALTS AND ORGANIC ACIDS METAL IONS AND AMINE COMPOUNDS	4.97E-04
	TRANSPORT OF INORGANIC CATIONS ANIONS AND AMINO ACIDS OLIGOPEPTIDES	1.33E-03
	SLC MEDIATED TRANSMEMBRANE TRANSPORT	1.40E-03

	TYPE I DIABETES MELLITUS	3.42E-03
	GRAFT VERSUS HOST DISEASE	3.55E-03
	DISORDERS OF TRANSMEMBRANE TRANSPORTERS	4.86E-03
	SLC TRANSPORTER DISORDERS	4.87E-03
	TRANSPORT OF SMALL MOLECULES	4.89E-03
	AUTOIMMUNE THYROID DISEASE	4.92E-03
	ALLOGRAFT REJECTION	6.73E-03
	ANTIGEN PROCESSING AND PRESENTATION	9.23E-03
	ASTHMA	1.55E-02
	RHOV GTPASE CYCLE	1.81E-02
	GPCR PATHWAY	1.81E-02
	NOS1 PATHWAY	1.94E-02
	ENDOCYTOSIS	2.09E-02
	RHOU GTPASE CYCLE	2.37E-02
	CREB PATHWAY	2.41E-02
	INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION	2.93E-02
	SYSTEMIC LUPUS ERYTHEMATOSUS	3.44E-02
	NON SMALL CELL LUNG CANCER	3.45E-02
	AT1R PATHWAY	3.72E-02
	NATURAL KILLER CELL MEDIATED CYTOTOXICITY	4.27E-02
	GENERATION OF SECOND MESSENGER MOLECULES	4.30E-02
	MEIOTIC RECOMBINATION	4.88E-02
	DEATH PATHWAY	4.89E-02
IC	ROS AND RNS PRODUCTION IN PHAGOCYTES	3.61E-02
	FERTILIZATION	3.64E-02

Table 4: Unique modules found across single tissue-specific eQTLs for the Italian and Spanish dataset. The most significant unique module was for cargo concentration in the ER found in the skin exposed lower leg tissue. Table is organized by smallest to largest FDR.

Tissues	**Pathway**	**FDR**
Skin Exposed Lower Leg	CARGO CONCENTRATION IN THE ER	8.54E-03
Cells Cultured Fibroblasts	THE PHOTOTRANSDUCTION CASCADE	1.96E-02

Brain Substantia Nigra	CELL CYCLE MITOTIC	3.05E-02

Table 5: The top most significant enriched pathway of the single tissue eQTL from the Italian and Spanish dataset. Of the enriched tissue-specific eQTLs the RHO GTPase was the most significant. Table is organized by smallest to largest FDR.

Tissues	Pathway	FDR
Brain Caudate Basal Ganglia	RHO GTPASE CYCLE	2.19E-05
Brain Caudate Basal Ganglia	RAC1 GTPASE CYCLE	2.60E-05
Heart Atrial Appendage	TRANSPORT TO THE GOLGI AND SUBSEQUENT MODIFICATION	3.47E-05
Brain Nucleus Accumbens Basal Ganglia	LONG TERM DEPRESSION	4.75E-05
Brain Caudate Basal Ganglia	CDC42 GTPASE CYCLE	9.13E-05
Colon Sigmoid	WNT SIGNALING PATHWAY	1.12E-04
Brain Caudate Basal Ganglia	NEUROACTIVE LIGAND RECEPTOR INTERACTION	1.22E-04
Brain Cortex	CDC42 GTPASE CYCLE	1.24E-04
Immune	HATS ACETYLATE HISTONES	1.30E-04

Table 6: All of the enriched pathways across the Chinese (C), UK, Italian and Spanish (IS), and Multi-European (ME) datasets. The table is organized in alphabetical order by tissue name. In total, there were 1751 enriched pathways across all populations. The esophagus muscularis had the greatest number of enriched pathways at 178. The totals of each tissue across populations are presented in the total column.

Tissue	*C*	*UK*	*IS*	*ME*	*Total*
Adipose Subcutaneous	102	17	0	23	142
Adipose Visceral Omentum	85	22	8	34	149
Artery Aorta	21	32	5	24	82
Artery Coronary	3	4	5	1	13
Artery Tibial	31	32	4	2	69
Brain Amygdala	2	0	0	15	17
Brain Anterior Cingulate Cortex Ba24	1	0	0	0	1
Brain Caudate Basal Ganglia	15	10	8	18	51
Brain Cerebellar Hemisphere	11	9	14	9	43
Brain Cerebellum	20	24	28	46	118

Brain Cortex	16	1	6	0	23
Brain Frontal Cortex Ba9	13	0	0	9	22
Brain Hippocampus	4	7	5	21	37
Brain Hypothalamus	3	0	3	0	6
Brain Nucleus Accumbens Basal Ganglia	13	9	15	21	58
Brain Putamen Basal Ganglia	5	0	0	0	5
Brain Substantia Nigra	2	0	1	1	4
Cells Cultured Fibroblasts	14	18	1	38	71
Cells EBV-Transformed Lymphocytes	3	3	5	15	26
Colon Sigmoid	16	19	11	22	68
Colon Transverse	3	17	6	21	47
Esophagus Gastroesophageal Junction	18	23	12	0	53
Esophagus Mucosa	23	64	0	0	87
Esophagus Muscularis	9	80	31	58	178
Heart Atrial Appendage	17	20	28	30	95
Heart Left Ventricle	15	13	13	18	59
Immune	13	5	11	2	31
Liver	5	0	0	0	5
Lung	17	17	1	1	36
Skin Not Sun Exposed Suprapubic	18	8	2	15	43
Skin Sun Exposed Lower Leg	18	7	1	17	43
Small Intestine Terminal Ileum	1	2	2	2	7
Whole Blood	69	2	0	4	75
Grand Total	*593*	*465*	*226*	*467*	*1751*

Table 7: The unique pathways seen when the Chinese (C), UK, Italian and Spanish (IS), and Multi-European (ME) are combined for MSEA. Table is sorted alphabetically for the population, tissue, and pathways. The Chinese population had 9 tissues and 15 unique pathways, The multi-European population had 5 tissues and 12 unique pathways, the Italian and Spanish population had 8 tissues and 11 unique pathways, and the UK population had 5 tissues and 18 unique pathways. Tissue abbreviations are found in **Supplementary 2.**

Pop	Tissue	Pathway
C	AS	MAPK6 MAPK4 SIGNALING
		MYD88 INDEPENDENT TLR4 CASCADE
		REGULATION OF RAS BY GAPS
		THE ROLE OF GTSE1 IN G2 M PROGRESSION AFTER G2 CHECKPOINT
	AVO	DEVELOPMENTAL BIOLOGY
	CER	LYSOSOME
	FC	METABOLISM OF NUCLEOTIDES

	EM	TIGHT JUNCTION
	HLV	GLYCEROPHOSPHOLIPID BIOSYNTHESIS
	L	DISEASES ASSOCIATED WITH O GLYCOSYLATION OF PROTEINS
		DRUG METABOLISM CYTOCHROME P450
	SNSES	THE CITRIC ACID TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT
	WB	CELLULAR RESPONSE TO HYPOXIA
		DOWNSTREAM SIGNALING EVENTS OF B CELL RECEPTOR BCR
		TNFR2 NON CANONICAL NF KB PATHWAY
IS	AA	ACTIVATION OF ATR IN RESPONSE TO REPLICATION STRESS
	CH	SENSORY PERCEPTION
	CER	ADIPOCYTOKINE SIGNALING PATHWAY
		REGULATION OF LIPID METABOLISM BY PPARALPHA
		SIGNALING BY RETINOIC ACID
	HYP	METABOLISM OF RNA
	NABG	CILIUM ASSEMBLY
	EMUSC	ECM RECEPTOR INTERACTION
		NICOTINATE AND NICOTINAMIDE METABOLISM
	HAA	GLYCINE SERINE AND THREONINE METABOLISM
	HLV	FATTY ACYL COA BIOSYNTHESIS
ME	AS	CASPASE MEDIATED CLEAVAGE OF CYTOSKELETAL PROTEINS
	A	DNA REPAIR
		EXTENSION OF TELOMERES
		G2 M DNA DAMAGE CHECKPOINT
	CCF	CLASS A 1 RHODOPSIN LIKE RECEPTORS
	EMUSC	AMINO ACID TRANSPORT ACROSS THE PLASMA MEMBRANE
		MUCOPOLYSACCHARIDOSES
		RNA POLYMERASE II TRANSCRIBES SNRNA GENES
		SLC TRANSPORTER DISORDERS
		TRAFFICKING OF GLUR2 CONTAINING AMPA RECEPTORS
		TRANSPORT OF INORGANIC CATIONS ANIONS AND AMINO ACIDS OLIGOPEPTIDES
	WB	NKT PATHWAY
UK	AVO	ANTIMICROBIAL PEPTIDES
	AT	GLIOMA
		MEF2D PATHWAY
	EM	FOCAL ADHESION
		INSULIN SIGNALING PATHWAY
		LONG TERM POTENTIATION
		TRIGLYCERIDE METABOLISM
	EMUSC	AT1R PATHWAY

		CREB PATHWAY
		DEATH PATHWAY
		ENDOCYTOSIS
		NATURAL KILLER CELL MEDIATED CYTOTOXICITY
		NON-SMALL CELL LUNG CANCER
		NOS1 PATHWAY
		SLC TRANSPORTER DISORDERS
		TRANSPORT OF INORGANIC CATIONS ANIONS AND AMINO ACIDS OLIGOPEPTIDES
	IC	FERTILIZATION
		ROS AND RNS PRODUCTION IN PHAGOCYTES

Table 8: Meta key driver analysis including all four populations. The results are for the tissue eQTLs and the top KDs for each. There were 14 tissues that had top key driver results, with a total of 350 results. Tissue abbreviations are found in **Supplementary 2.**

Row Labels	AA	AC	AS	AT	AVO	CCF	CS	EGJ	EM	EMUSC	HAA	HLV	SELL	SNES	Grand Total
CD74	5		6	9	4		3								27
COL5A1										1					1
CXCL3				1											1
GBP4			2												2
HLA-B			6		4										10
HLA-DMA			3		4										7
HLA-DMB			6												6
HLA-DPB1	4			6											10
HLA-DQA1			4		5			7	3	8	3				30
HLA-DQA2											9	6			15
HLA-DQB1	4	1	6	4	4			5	7	8					39
HLA-DQB1-AS1														1	1
HLA-DRA	7			9				7	7	8			5	8	51
HLA-DRB1			5	2	5										12
HLA-DRB5											12	8			20
HLA-DRB6								6	5	7					18
HLA-H														3	3
IGHV1-46									1						1
ISG15			1												1
LCP2					1										1
LY86					1										1
MIF-AS1				1				1	1	1			1		5
NCKAP1L			1												1
PSMB8			3												3
RARRES3	1														1
RNASE3					1										1
RP11-259G18.3	15	2		11		11		7	13	21					80
SLC11A1					1										1
UQCRC2			1												1
Grand Total	**36**	**3**	**44**	**43**	**30**	**11**	**3**	**33**	**37**	**54**	**24**	**14**	**6**	**12**	**350**

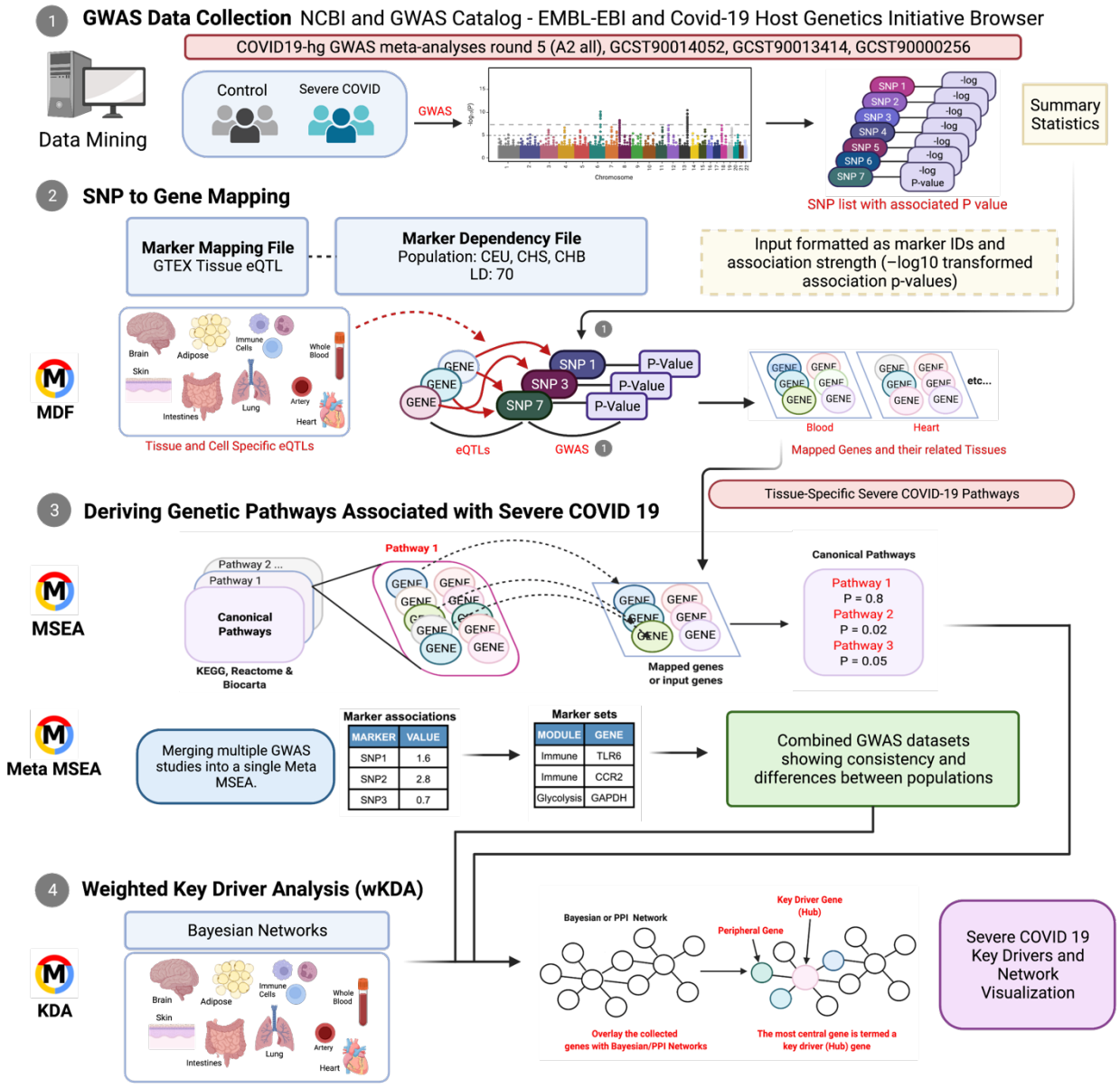


Figure 1: Overview of study. 1) All data were mined from the NCBI GWAS Catalog or Covid-19 Host Genetics Initiative Browser. Summary statistics for the 4 datasets were mined and edited for input into MDF. 2) GWAS SNPs are mapped onto genes using single tissue-specific eQTLs or combined tissue-specific eQTLs. Genes found are then linked to canonical pathways and co-expression modules. 3) MSEA of single and combined tissue eQTLs for each of the 4 populations are carried out independently for pathway/co-expression module enrichment. Meta-MSEA of the combined cohorts for pathway/co-expression module enrichment. Similar modules are then categorized into independent supersets and input into a wKDA. 4) wKDA implemented Bayesian networks independently for key driver gene identification and network visualization.

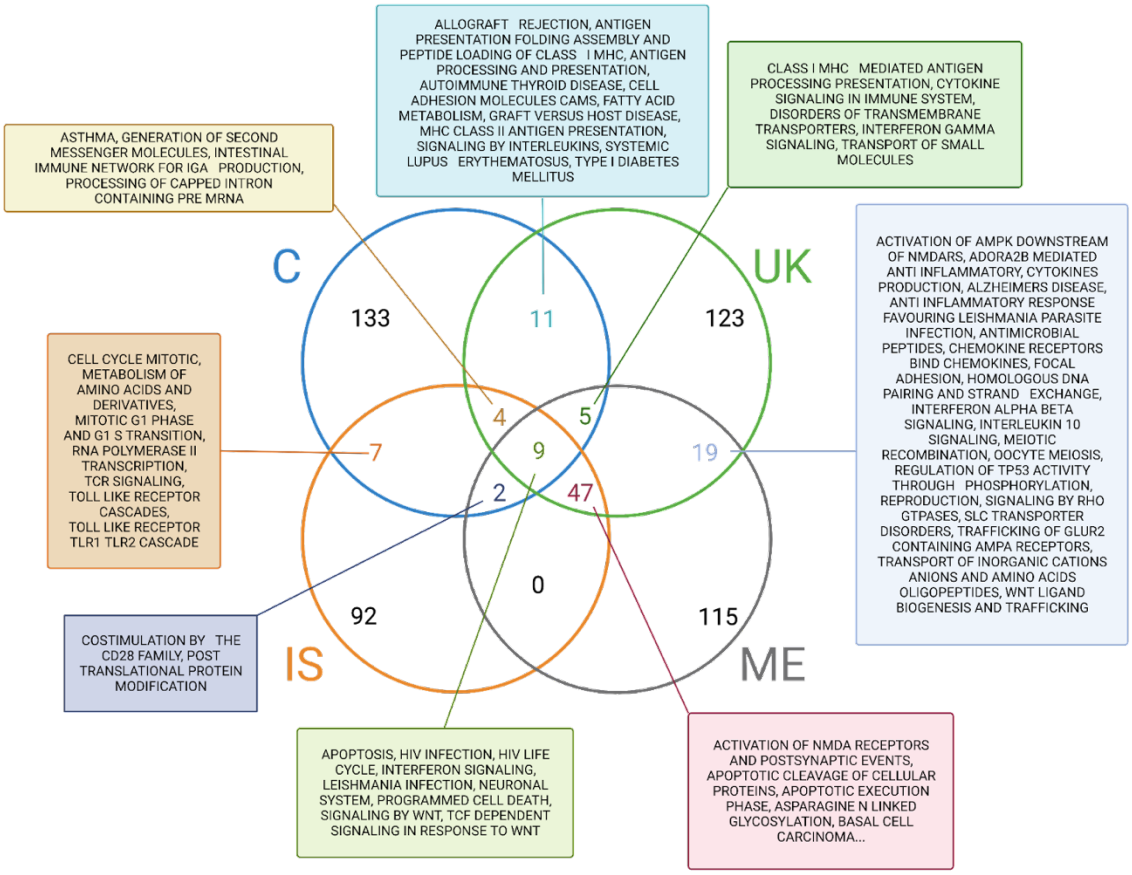


Figure 2: Venn Diagram of the unique and consistent (not repeated) enriched modules for the MSEA single tissue eQTLs. There were 463 enriched modules (not repeated for their own population) across four population datasets. The Chinese (C), Italian and Spanish (IS), Multi-European (ME), and UK populations had 133, 92, 115, and 123 non-repeated enriched modules, respectively. Seven enriched modules were shared between C and IS, 9 between C and ME, 11 between C and UK, 0 between IS and ME, 4 between IS and UK, and 19 between ME and UK. When looking at similarities between three of the populations, C, IS, and UK had 4 similar enriched modules, C, ME, and UK had 5 similar enriched modules, IS, ME, and UK, had 47 similar enriched modules, and C, IS, and ME had 2 similar enriched modules. Across all four datasets, there were 4 shared enriched modules.

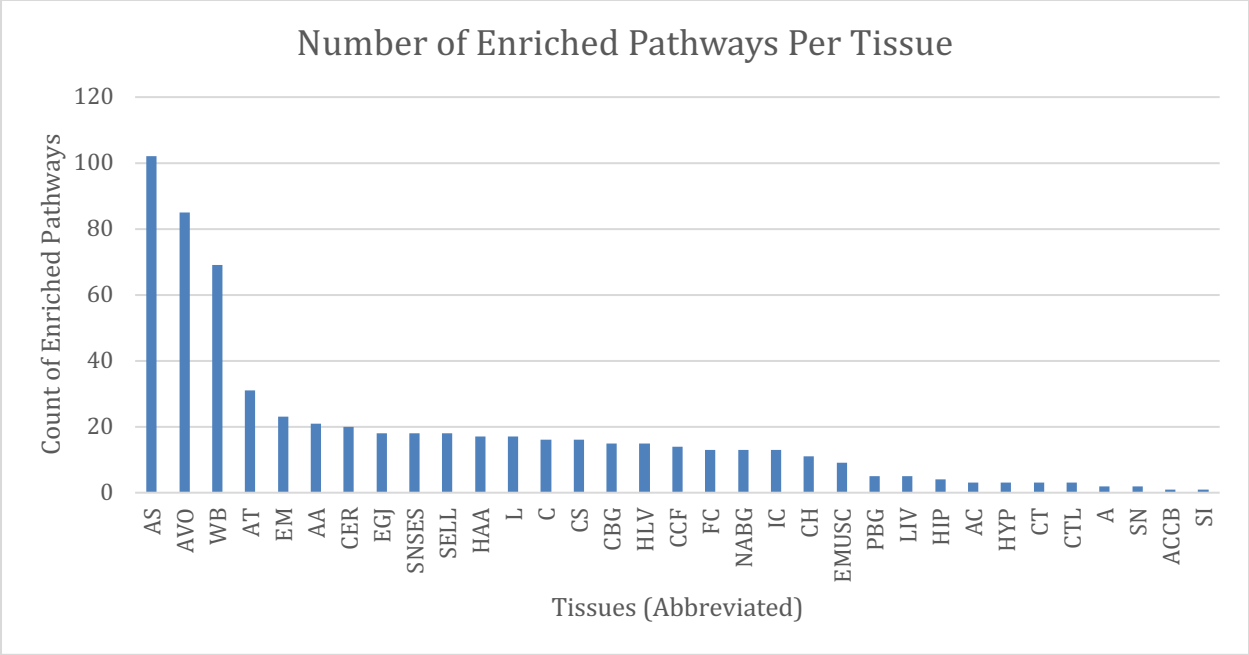


Figure 3: Bar graph of tissue eQTLs with their enriched pathway count for the Chinese dataset. The X-axis is organized from highest to lowest enriched pathway count. Chinese MSEA results showed a total of 593 enriched modules across 33 tissues. The GTEx adipose subcutaneous eQTL had the most enriched modules at 102. Tissue-specific eQTLs with only 1 enriched module were for brain anterior cingulate cortex Ba24 and small intestine terminal ileum. Abbreviations for tissues found in **Supplementary 2**.

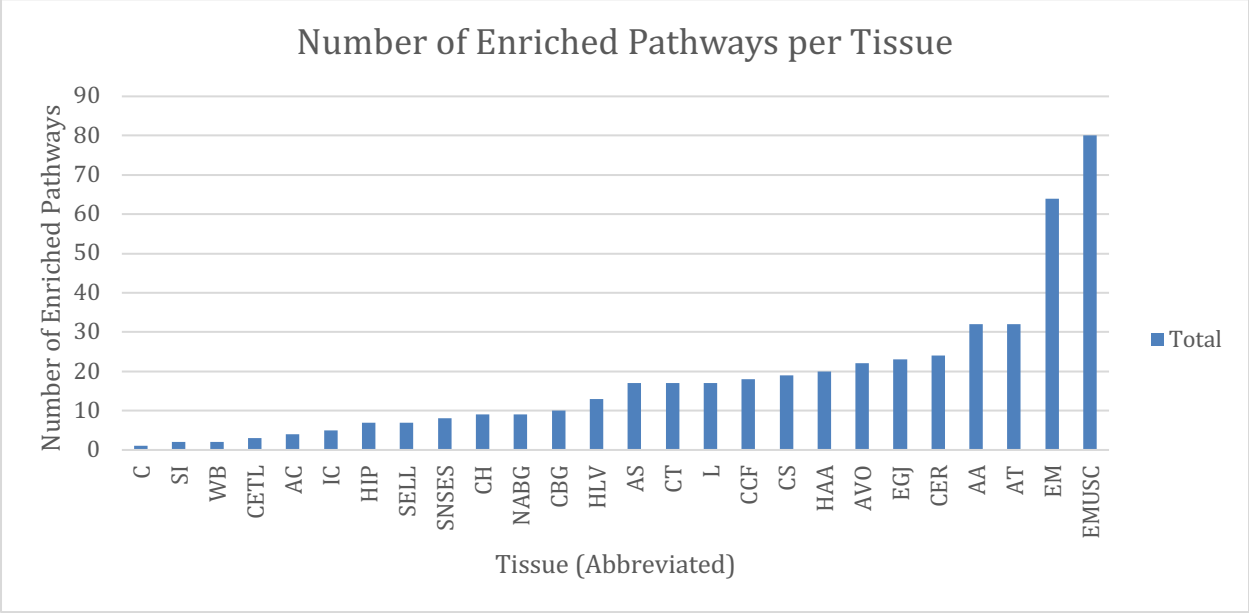


Figure 4: Bar graph of number of enriched pathways per tissue for the UK dataset. UK MSEA results showed a total of 465 enriched pathways across 32 tissues. The GTEx esophagus muscularis eQTL had the most enriched modules at 80. Tissue-specific eQTLs with

only 1 enriched module were for brain cortex. X-axis is organized by tissue abbreviation from lowest the highest pathway count. Abbreviations for tissues are found in **Supplementary 2**.

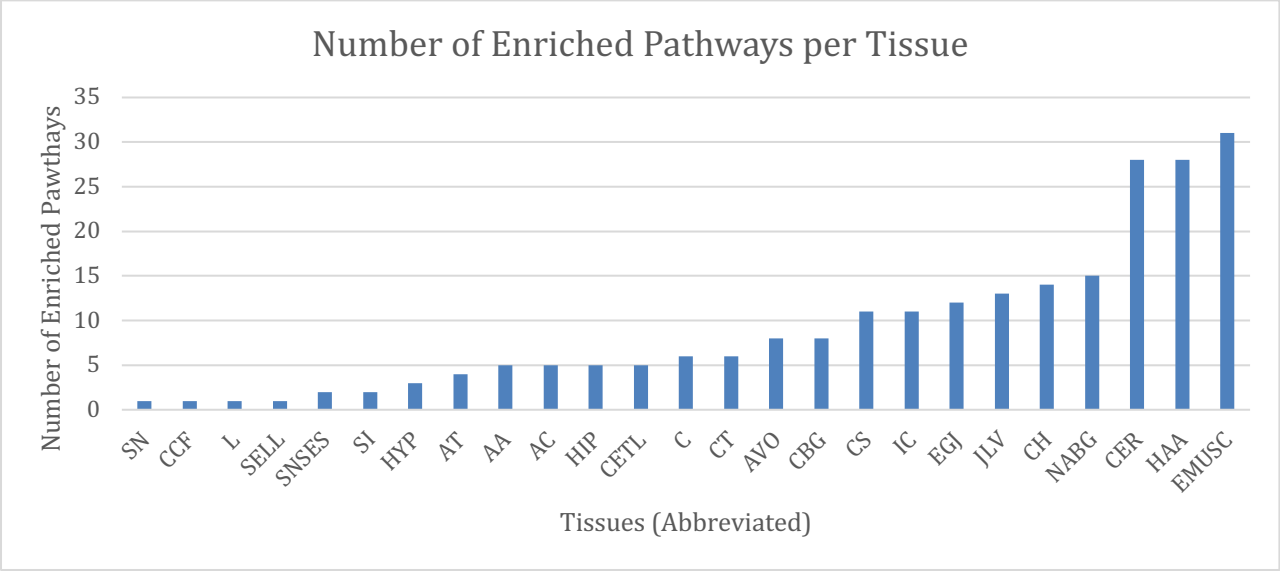


Figure 5: Bar graph of tissue eQTLs with their enriched module count for the Italian and Spanish dataset. Italian and Spanish MSEA results showed a total of 226 enriched modules across 25 tissues. The GTEx ssophagus muscularis eQTL had the most enriched modules at 31. Tissue-specific eQTLs with only 1 enriched module were for brain substantia nigra, cultured fibroblasts, lung, and skin exposed lower leg. The x-axis is organized by tissue abbreviation from lowest to highest pathway count. Abbreviations for tissues are found in **Supplementary 2**.

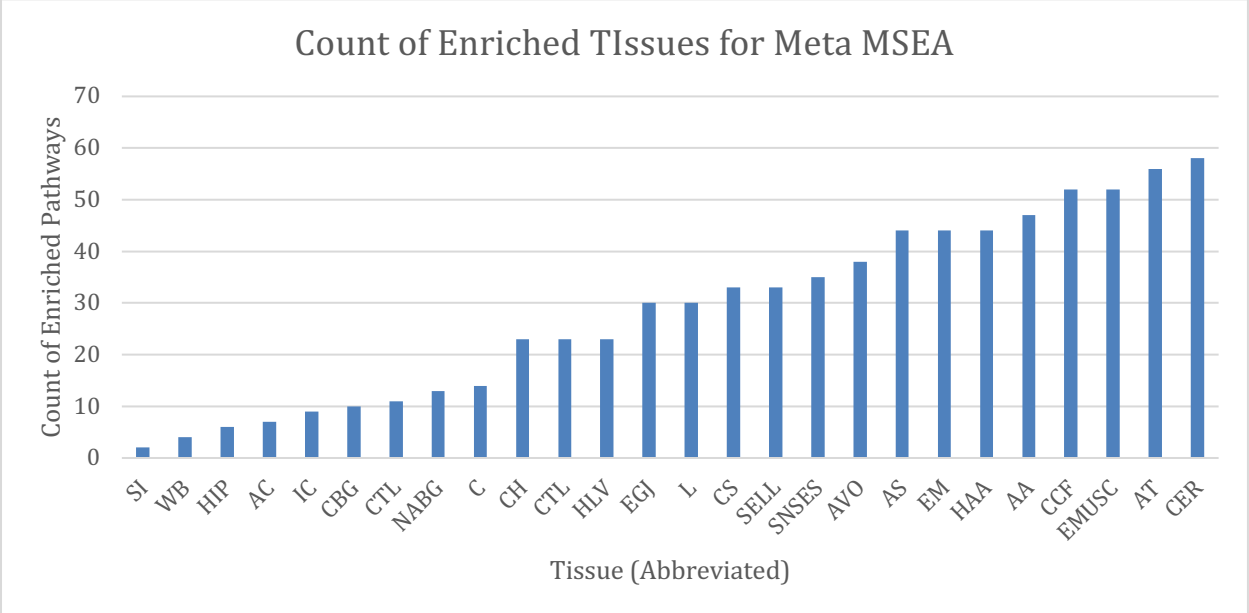


Figure 6: Bar graph of tissue eQTLs with their enriched module count for the Meta MSEA. Meta MSEA results showed a total of 741 enriched modules across 26 tissues. The GTEx cerebellum eQTL had the most enriched modules at 58. Tissue-specific eQTLs with only 2

enriched modules was for the small intestine. The x-axis is organized by tissue abbreviation from lowest to highest pathway count. Abbreviations for tissues are found in **Supplementary 2**.

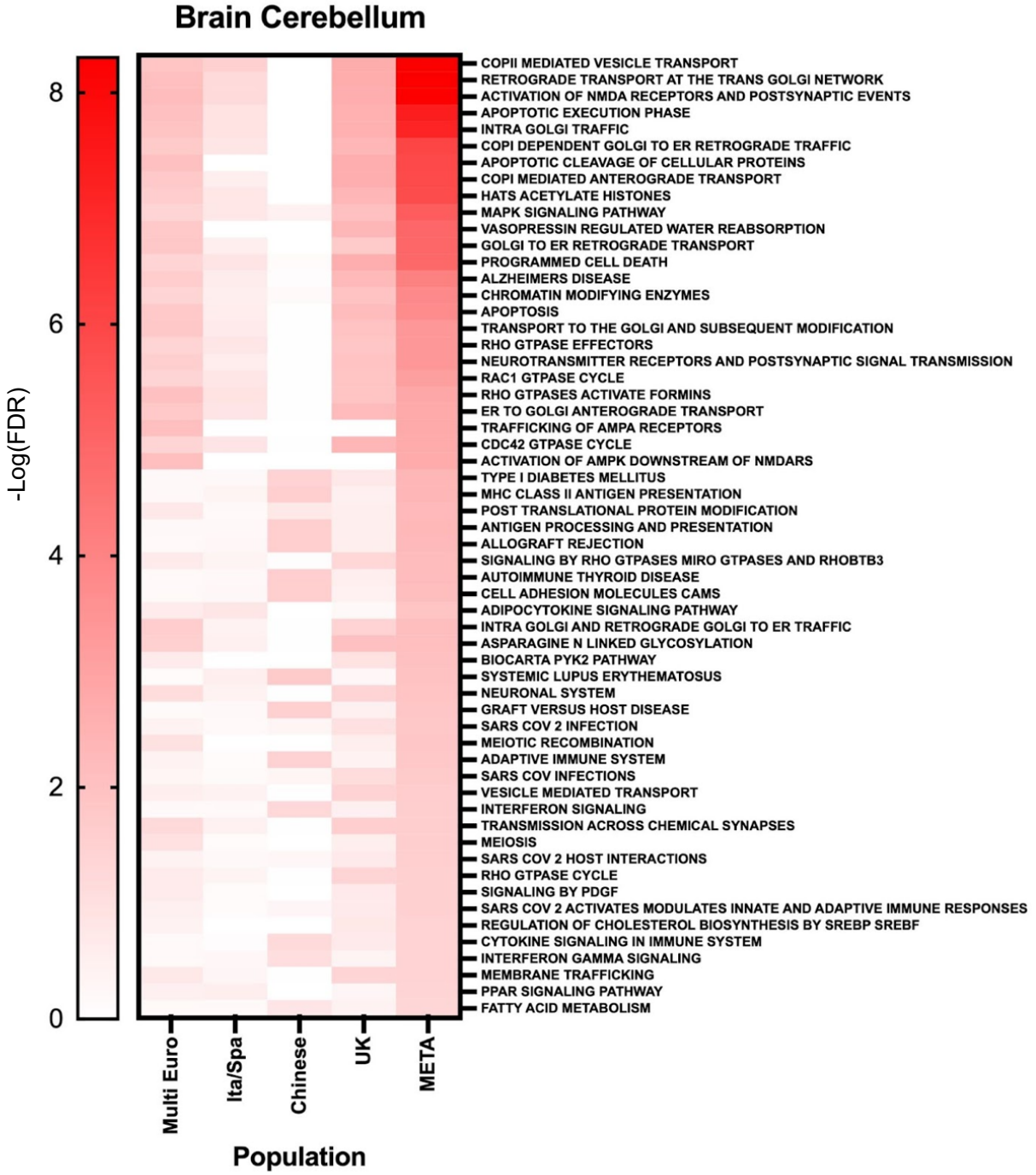


Figure 7: Heat map of the cerebellum from the Meta MSEA for Multi European, Italian/Spanish, Chinese, and UK populations for Severe COVID-19. Heat map for the Meta results of all four populations and their -Log(FDR) for each enriched pathway in the cerebellum. The most significant pathway was for COPII mediated vesicle transport. The y-axis lists all the enriched pathways in order of their -Log(FDR) from smallest to largest.

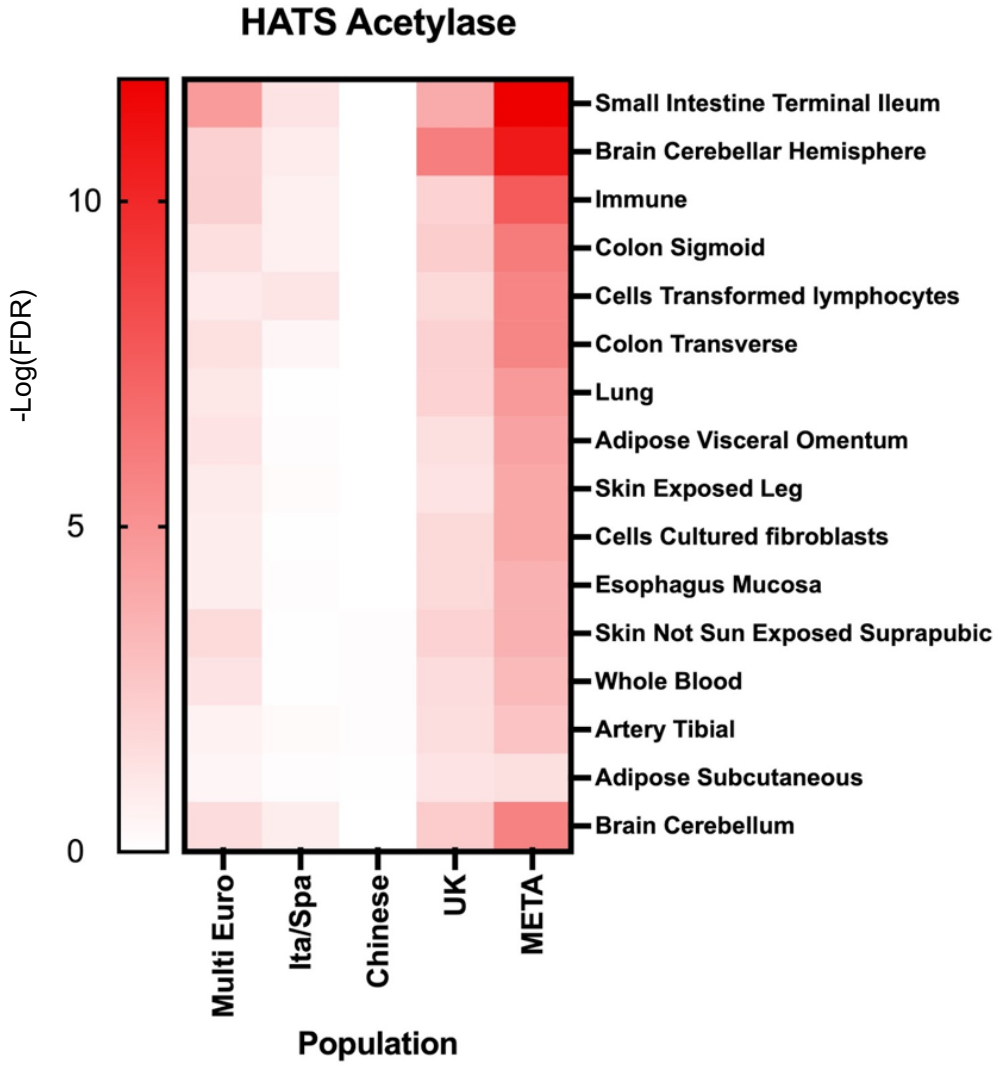


Figure 8: Heat map of the HATS Acetylase pathway from the Meta MSEA for Multi European, Italian/Spanish, Chinese, and UK populations for Severe COVID-19. Heat map of the Meta results for all four populations and their -Log(FDR) for each tissue that had the enriched pathway HATS Acetylase. The most significant -LOG(FDR) for the pathway was found in the small intestine terminal ileum. The y-axis lists all the tissues for HATS Acetylase in order of their -Log(FDR) from smallest to largest.

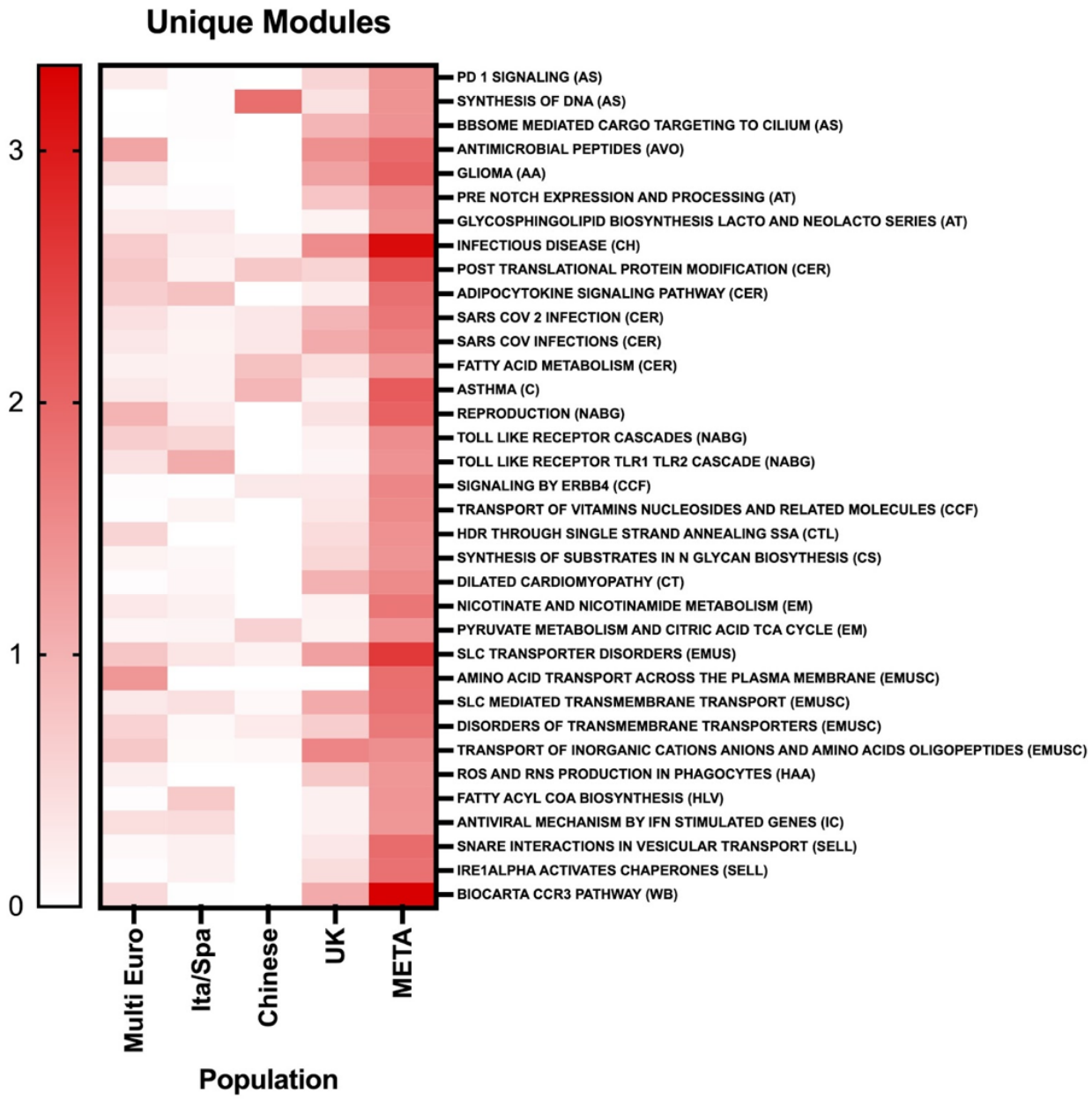
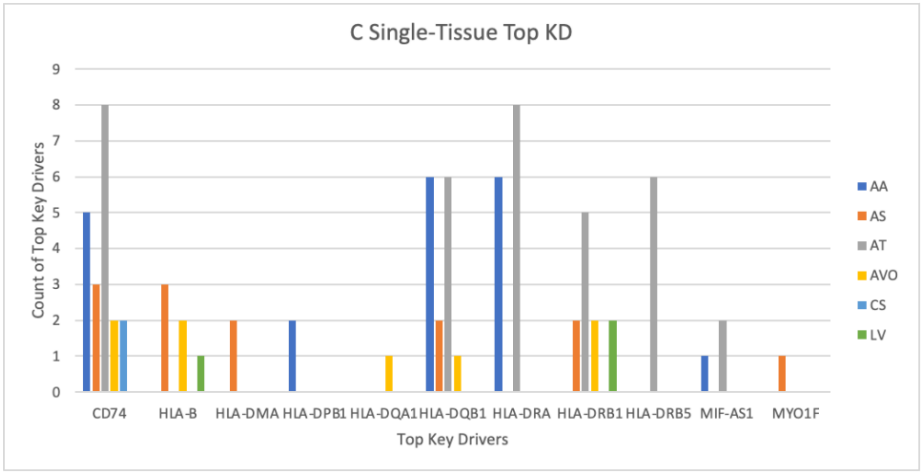
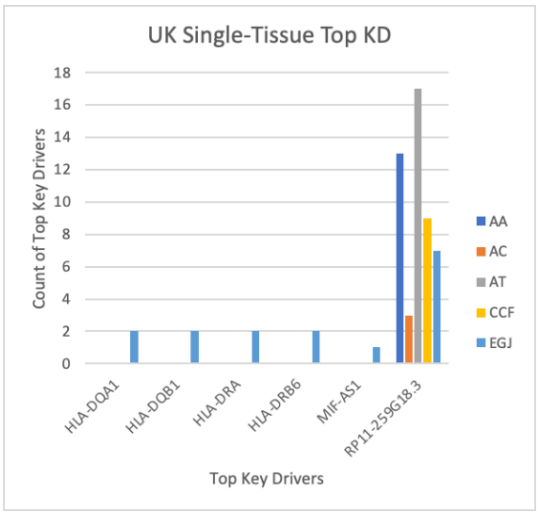


Figure 9: Heat map of all the tissue-specific pathways from the Meta MSEA for Multi European, Italian/Spanish, Chinese, and UK populations for Severe COVID-19. Heat map of the Meta results all four populations and their -Log(FDR) for each unique pathway found across all tissues. Unique pathway labels have their corresponding tissue abbreviation in parentheses. The most significant -Log(FDR) were found for the CCR3 pathway in whole blood and the infectious disease pathway found in the cerebral hemisphere. The y-axis lists all the unique pathways with their abbreviated tissues in order of their -Log(FDR) from smallest to largest.

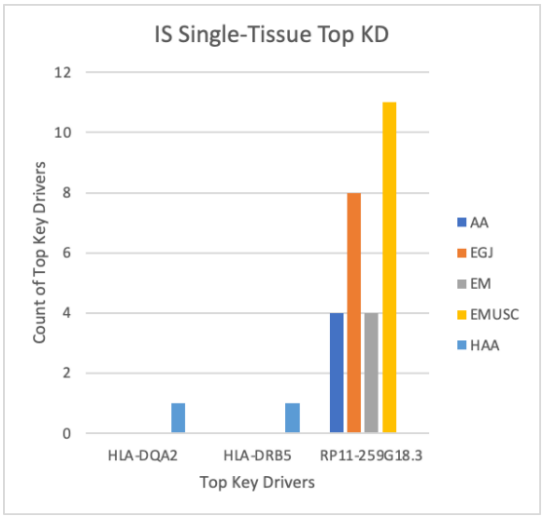
a.



b.



c.



d.

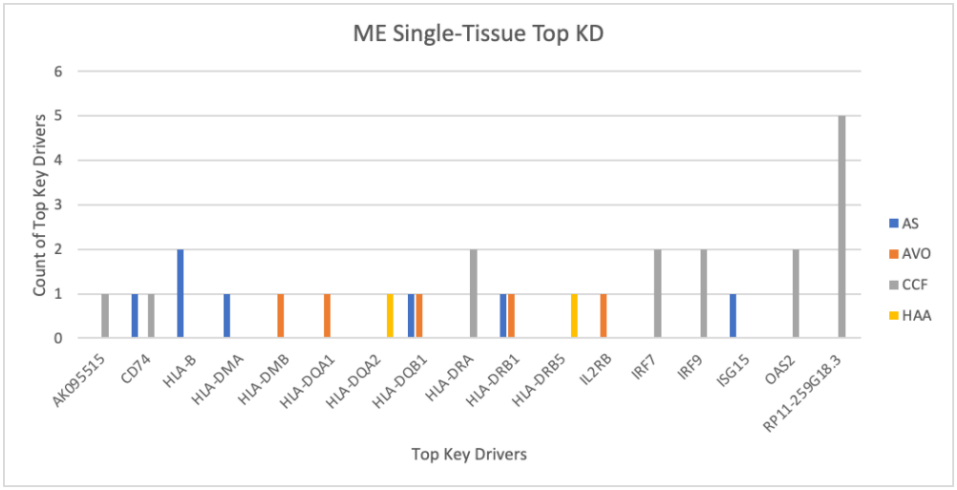


Figure 10: Single tissue eQTL key driver analysis with the key drivers for each tissue. A) There were 11 top key drivers over 6 tissues for the Chinese dataset. b) There were 6 top key drivers over 5 tissues for the UK dataset. c) There were 3 top key drivers over 5 tissues for the Italian and Spanish dataset. d) There were 17 top key drivers over 4 tissues for the Multi European dataset. Tissues were abbreviated based on the table found in **Supplementary 2**.

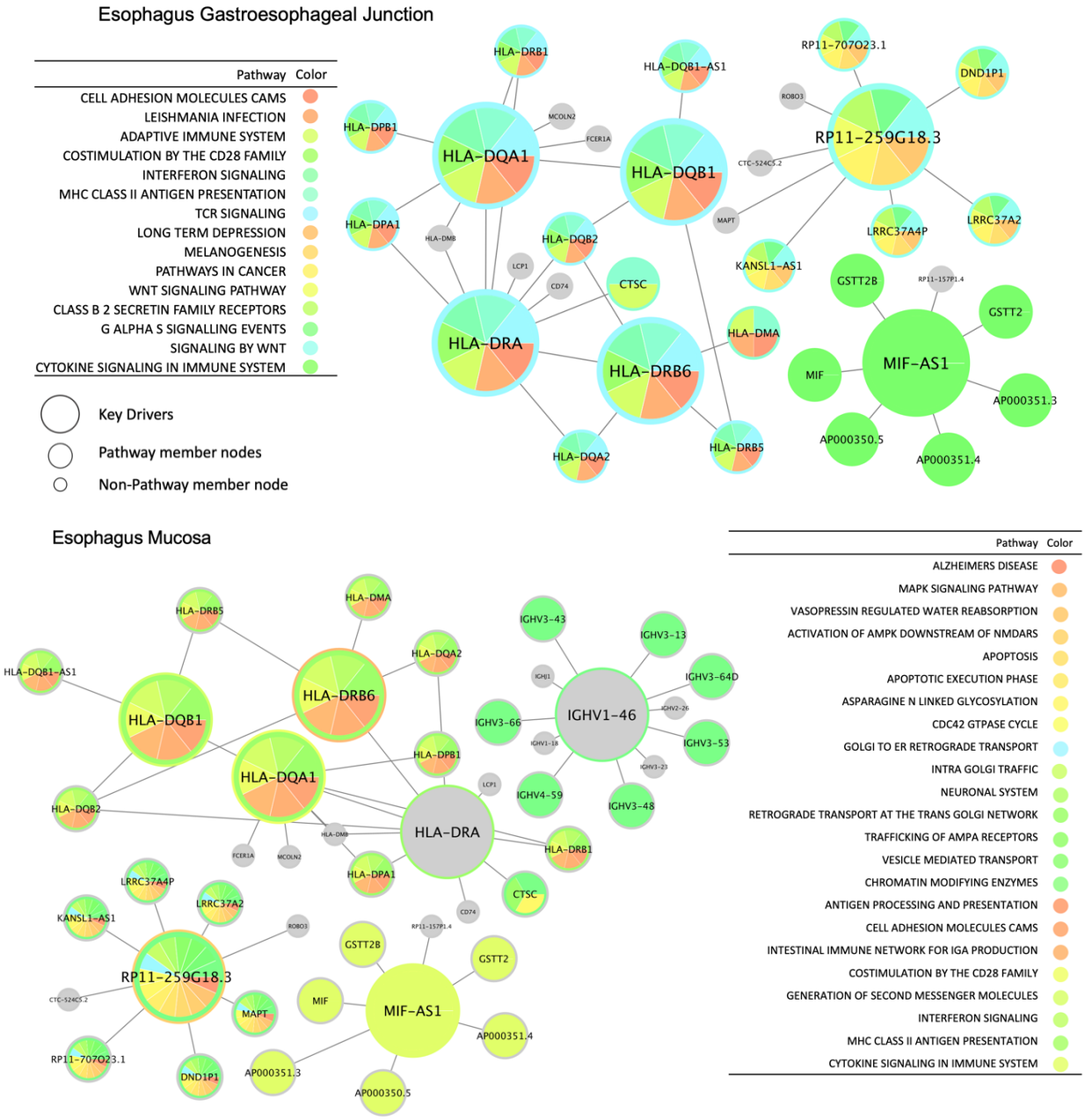


Figure 11: Network visualization of Esophagus Mucosa (EM) and Gastroesophageal Junction (EGJ) for the Meta MSEA. Key drivers (KDs) of the networks are represented by the largest circles, with mid-size circles representing nodes that are part of a pathway but are the peripheral genes that are involved but aren't critical, and the smaller gray circles are the nodes that are a non-pathway members but are present in the Bayesian network. The colors represent the pathways that are involved and are labeled in the tables. Circles with more than one color are multiple pathway members and gray circles are non-pathway members. The border on the shapes represents the pathway for the KD. The EGJ meta results yielded 6 KD, with 5 in a network. EM meta results showed 7 KDs with 4 in a network.

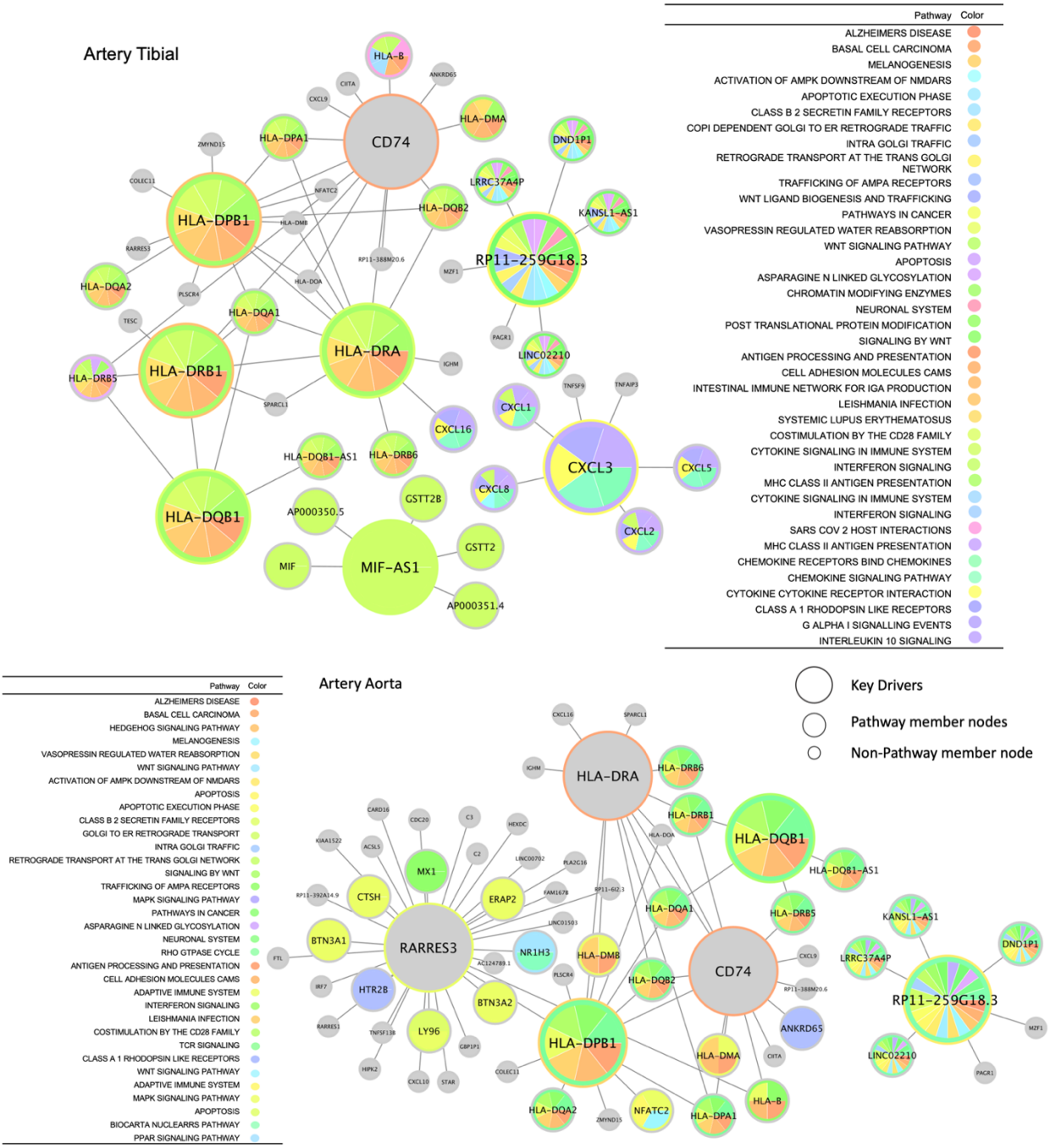


Figure 12: Network visualization of Artery Tibial (AT) and Aorta (AA) for the Meta MSEA. Key drivers (KDs) of the networks aere represented as the largest circles, with mid-size circles representing nodes that are part of a pathway but are the peripheral genes that are involved but aren't critical, and the smaller gray circles are the nodes that are a non-pathway member but were present in the Bayesian network. The colors represent the pathways that are involved and are labeled in the tables. Circles with more than one color are multiple pathway members and gray circles are non-pathway members. The border on the shapes represents the pathway for the KD. The AT meta results yielded 8 KD, with 5 in a network. AA meta results showed 6 KDs with 5 in a network.

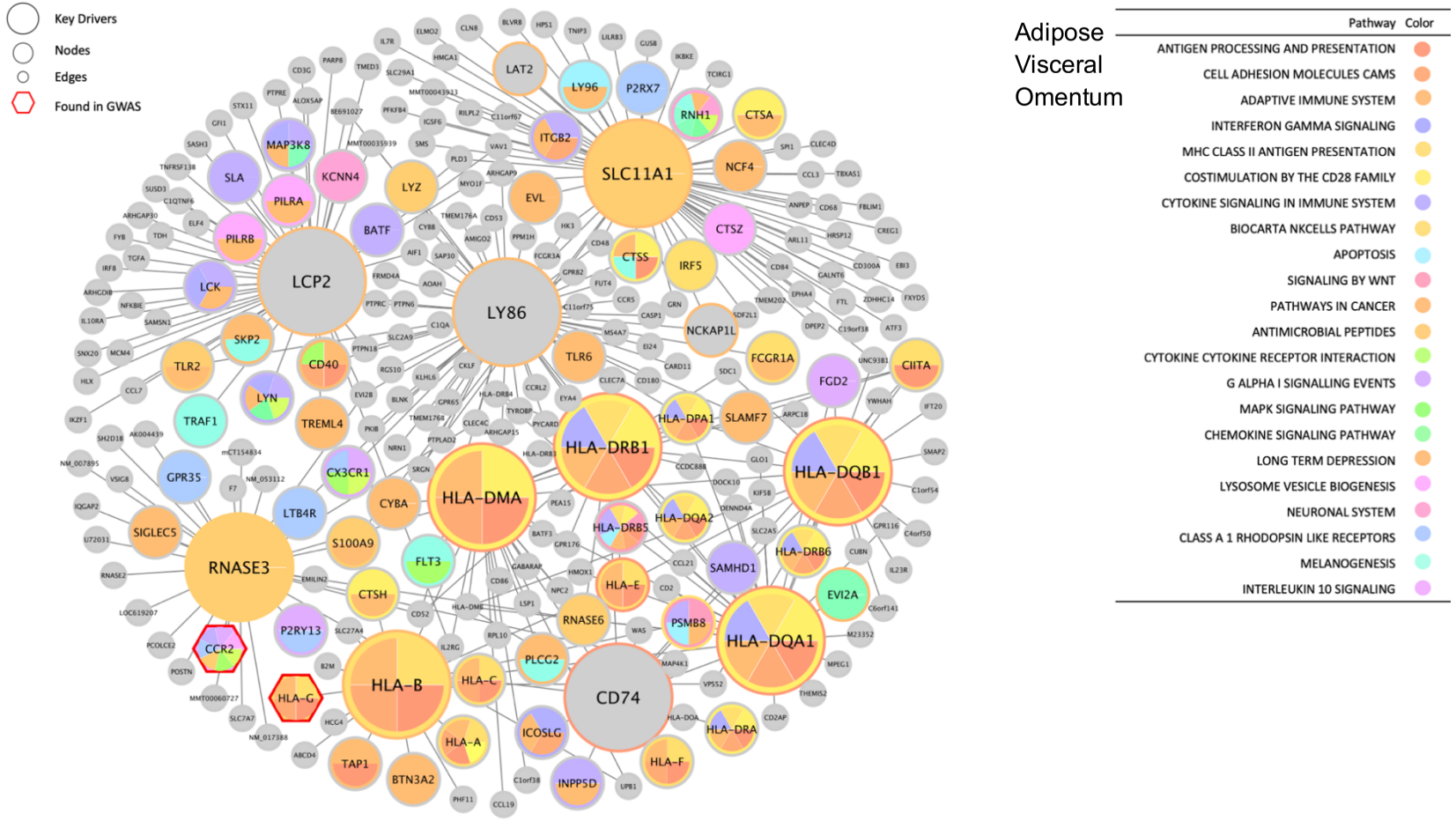


Figure 13: Network visualization of Adipose Visceral Omentum (AVO) for the Meta MSEA. Key drivers (KDs) of the networks are represented as the largest circles, with mid-size circles representing nodes that are part of a pathway but are the peripheral genes that are involved but aren't critical, and the smaller gray circles are the nodes that are non-pathway members but were present in the Bayesian network. The colors represent the pathways that are involved and are labeled in the tables. Circles with more than one color are multiple pathway members and gray circles are non-pathway members. Hexagons with a red border indicates that the highlighted SNP was also found in the results for one of the four GWAS datasets. The border on the shapes represents the pathway for the KD. The AVO meta results yielded 10 KD, all of which were connected to a network. CCR2 and HLA-G were SNPs that were also found in the original results the GWAS were taken from.

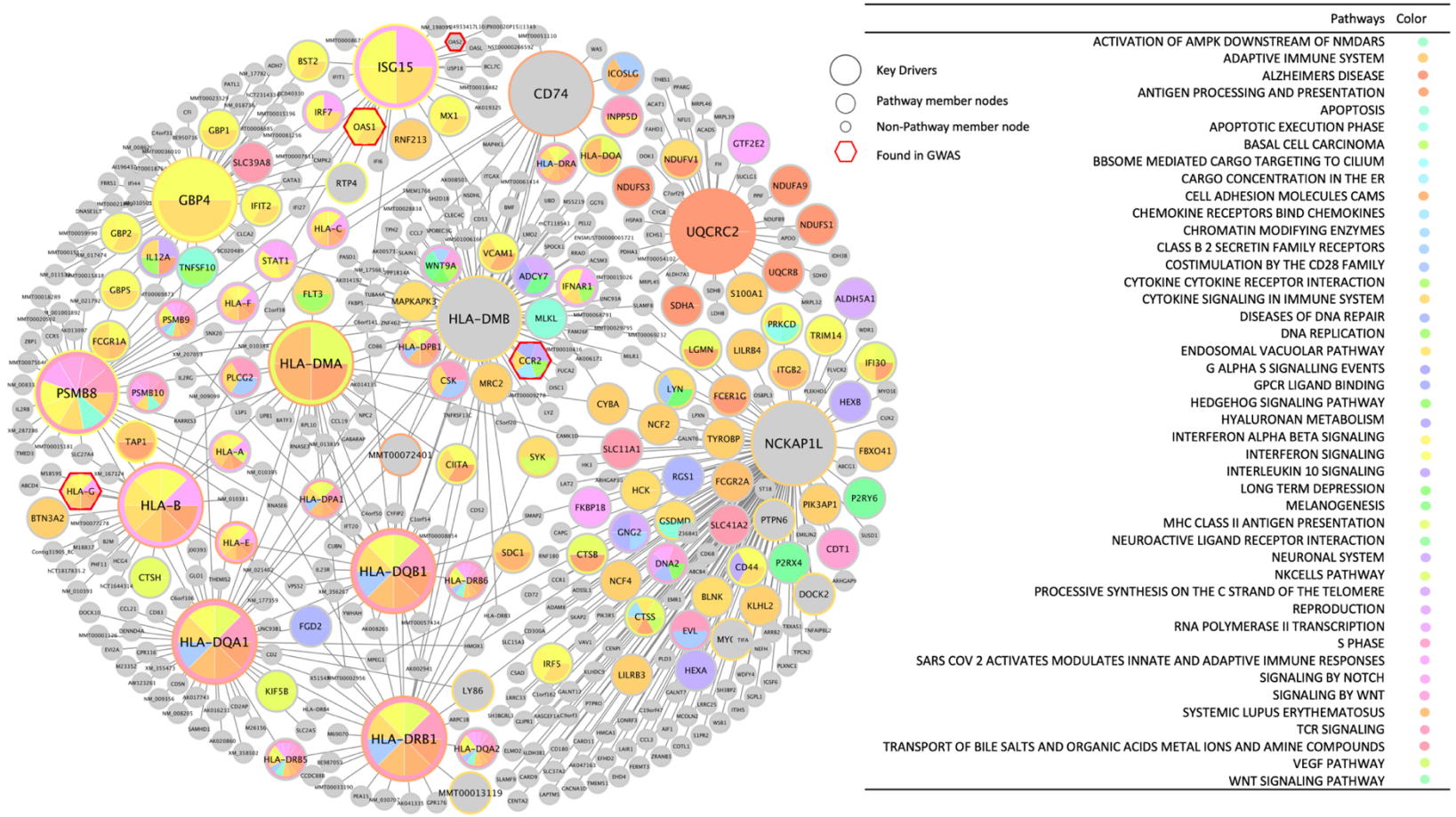


Figure 14: Network visualization of Adipose Subcutaneous (AS) for the Meta MSEA. Key drivers (KDs) of the networks are represented as the largest circles, with mid-size circles representing nodes that are part of a pathway but are the peripheral genes that are involved but aren't critical, and the smaller gray circles are the nodes that are non-pathway members but were present in the Bayesian network. The colors represent the pathways that are involved and are labeled in the tables. Circles with more than one color are multiple pathway members and gray circles are non-pathway members. Hexagons with a red border indicate that the highlighted SNP was also found in the results for one of the four GWAS datasets. The border on the shapes represents the pathway for the KD. The AS meta results yielded 10 KDs, 9 of which were connected to a network. OAS1, OAS2, CCR2 and HLA-G were SNPs that were also found in the original results the GWAS were taken from.

Bibliography

1. “WHO Team. Global Overview. 2021. ,” n.d.
2. Liu X, Zhou H, Zhou Y, Wu X, Zhao Y, Lu Y, et al. Risk factors associated with disease severity and length of hospital stay in COVID-19 patients. Journal of Infection. 2020 Jul;81(1):e95–7.
3. Baj J, Karakuła-Juchnowicz H, Teresiński G, Buszewicz G, Ciesielka M, Sitarz R, et al. COVID-19: Specific and Non-Specific Clinical Manifestations and Symptoms: The Current State of Knowledge. J Clin Med. 2020 Jun 5;9(6):1753.
4. Fricke-Galindo I, Falfán-Valencia R. Genetics Insight for COVID-19 Susceptibility and Severity: A Review. Front Immunol. 2021;12:622176.
5. Stokes EK, Zambrano LD, Anderson KN, Marder EP, Raz KM, El Burai Felix S, et al. Coronavirus Disease 2019 Case Surveillance — United States, January 22–May 30, 2020. MMWR Morb Mortal Wkly Rep. 2020 Jun 19;69(24):759–65.
6. Ragab D, Salah Eldin H, Taeimah M, Khattab R, Salem R. The COVID-19 Cytokine Storm; What We Know So Far. Front Immunol. 2020 Jun 16;11.
7. Mokhtari T, Hassani F, Ghaffari N, Ebrahimi B, Yarahmadi A, Hassanzadeh G. COVID-19 and multiorgan failure: A narrative review on potential mechanisms. J Mol Histol. 2020 Dec 4;51(6):613–28.
8. Jain U. Effect of COVID-19 on the Organs. Cureus. 2020 Aug 3;12(8):e9540.
9. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic mechanisms of critical illness in COVID-19. Nature. 2021 Mar 4;591(7848):92–8.
10. Chen YW, Diamante G, Ding J, Nghiem TX, Yang J, Ha SM, et al. PharmOmics: A species- and tissue-specific drug signature database and gene-network-based drug repositioning tool. iScience. 2022 Apr;25(4):104052.
11. Arneson D, Bhattacharya A, Shu L, Mäkinen VP, Yang X. Mergeomics: a web server for identifying pathological pathways, networks, and key regulators via multidimensional data integration. BMC Genomics. 2016 Dec 9;17(1):722.
12. Ding J, Blencowe M, Nghiem T, Ha S min, Chen YW, Li G, et al. Mergeomics 2.0: a web server for multi-omics data integration to elucidate disease networks and predict therapeutics. Nucleic Acids Res. 2021 Jul 2;49(W1):W375–87.
13. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res. 2023 Jan 6;51(D1):D977–85.
14. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. Nature. 2021 Dec;600(7889):472–7.
15. Aguet F, Anand S, Ardlie KG, Gabriel S, Getz GA, Graubert A, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science (1979). 2020 Sep 11;369(6509):1318–30.
16. “Database of Immune Cell Eqtls, Expression, Epigenomics.” DICE, 2018, dice-database.org/.
17. Wu P, Chen D, Ding W, Wu P, Hou H, Bai Y, et al. The trans-omics landscape of COVID-19. Nat Commun. 2021 Jul 27;12(1):4543.

18. Arunachalam PS, Wimmers F, Mok CKP, Perera RAPM, Scott M, Hagan T, et al. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. Science. 2020 Sep 4;369(6508):1210–20.
19. Severe Covid-19 GWAS Group. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. New England Journal of Medicine. 2020 Oct 15;383(16):1522–34.
20. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. European Journal of Human Genetics. 2020 Jun 13;28(6):715–8.
21. Fabregat A, Sidiropoulos K, Viteri G, Forner O, Marin-Garcia P, Arnau V, et al. Reactome pathway analysis: a high-performance in-memory approach. BMC Bioinformatics. 2017 Dec 2;18(1):142.
22. Nishimura D. BioCarta. Biotech Software & Internet Report. 2001 Jun;2(3):117–20.
23. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 2007 Dec 23;36(Database):D480–4.
24. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, et al. Genetics of gene expression and its effect on disease. Nature. 2008 Mar 27;452(7186):423–8.
25. Derry JMJ, Zhong H, Molony C, MacNeil D, Guhathakurta D, Zhang B, et al. Identification of genes and networks driving cardiovascular and metabolic phenotypes in a mouse F2 intercross. PLoS One. 2010 Dec 14;5(12):e14319.
26. Wang SS, Schadt EE, Wang H, Wang X, Ingram-Drake L, Shi W, et al. Identification of Pathways for Atherosclerosis in Mice. Circ Res. 2007 Aug 3;101(3).
27. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, et al. Mapping the Genetic Architecture of Gene Expression in Human Liver. PLoS Biol. 2008 May 6;6(5):e107.
28. Tu Z, Keller MP, Zhang C, Rabaglia ME, Greenawalt DM, Yang X, et al. Integrative Analysis of a Cross-Loci Regulation Network Identifies App as a Gene Regulating Insulin Secretion from Pancreatic Islets. PLoS Genet. 2012 Dec 6;8(12):e1003107.
29. Zhu J, Lum PY, Lamb J, GuhaThakurta D, Edwards SW, Thieringer R, et al. An integrative genomics approach to the reconstruction of gene networks in segregating populations. Cytogenet Genome Res. 2004;105(2–4):363–74.
30. Li Q, Wang Y, Sun Q, Knopf J, Herrmann M, Lin L, et al. Immune response in COVID-19: what is next? Cell Death Differ. 2022 Jun 17;29(6):1107–22.
31. Marcuzzi A, Melloni E, Zauli G, Romani A, Secchiero P, Maximova N, et al. Autoinflammatory Diseases and Cytokine Storms-Imbalances of Innate and Adaptative Immunity. Int J Mol Sci. 2021 Oct 18;22(20).
32. Sharma C, Bayry J. High risk of autoimmune diseases after COVID-19. Nat Rev Rheumatol. 2023 Apr 12;
33. Liu Y, Sawalha AH, Lu Q. COVID-19 and autoimmune diseases. Curr Opin Rheumatol. 2021 Mar;33(2):155–62.
34. Pikoulas A, Piperaki ET, Spanakos G, Kallianos A, Mparmparousi D, Rentziou G, et al. Visceral leishmaniasis and COVID-19 coinfection - A case report. IDCases. 2022;27:e01358.

35. Demoliou C, Papaneophytou C, Nicolaidou V. SARS-CoV-2 and HIV-1: So Different yet so Alike. Immune Response at the Cellular and Molecular Level. Int J Med Sci. 2022;19(12):1787–95.
36. Bradic M, Taleb S, Thomas B, Chidiac O, Robay A, Hassan N, et al. DNA methylation predicts the outcome of COVID-19 patients with acute respiratory distress syndrome. J Transl Med. 2022 Nov 12;20(1):526.
37. Swain O, Romano SK, Miryala R, Tsai J, Parikh V, Umanah GKE. SARS-CoV-2 Neuronal Invasion and Complications: Potential Mechanisms and Therapeutic Approaches. J Neurosci. 2021 Jun 23;41(25):5338–49.
38. Natale NR, Lukens JR, Petri WA. The nervous system during <scp>COVID</scp> -19: Caught in the crossfire. Immunol Rev. 2022 Oct 30;311(1):90–111.
39. Basolo A, Poma AM, Bonuccelli D, Proietti A, Macerola E, Ugolini C, et al. Adipose tissue in COVID-19: detection of SARS-CoV-2 in adipocytes and activation of the interferon-alpha response. J Endocrinol Invest. 2022 May 15;45(5):1021–9.
40. Martínez-Colón GJ, Ratnasiri K, Chen H, Jiang S, Zanley E, Rustagi A, et al. SARS-CoV-2 infection drives an inflammatory response in human adipose tissue through infection of adipocytes and macrophages. Sci Transl Med. 2022 Dec 7;14(674):eabm9151.
41. Kamasak B, Ulcay T, Nisari M, Gorgulu O, Akca V, Alpaslan M, et al. Effects of COVID-19 on brain and cerebellum: a voxel based morphometrical analysis. Bratisl Lek Listy. 2023;124(6):442–8.
42. Fadakar N, Ghaemmaghami S, Masoompour SM, Shirazi Yeganeh B, Akbari A, Hooshmandi S, et al. A First Case of Acute Cerebellitis Associated with Coronavirus Disease (COVID-19): a Case Report and Literature Review. The Cerebellum. 2020 Dec 31;19(6):911–4.
43. Meinhardt J, Radke J, Dittmayer C, Franz J, Thomas C, Mothes R, et al. Olfactory transmucosal SARS-CoV-2 invasion as a port of central nervous system entry in individuals with COVID-19. Nat Neurosci. 2021 Feb 30;24(2):168–75.
44. Xydakis MS, Albers MW, Holbrook EH, Lyon DM, Shih RY, Frasnelli JA, et al. Post-viral effects of COVID-19 in the olfactory system and their implications. Lancet Neurol. 2021 Sep;20(9):753–61.
45. Serrano GE, Walker JE, Tremblay C, Piras IS, Huentelman MJ, Belden CM, et al. SARS-CoV-2 Brain Regional Detection, Histopathology, Gene Expression, and Immunomodulatory Changes in Decedents with COVID-19. J Neuropathol Exp Neurol. 2022 Sep;81(9):666–95.
46. Song JH, Langworthy J, Abbas AE, Parkman HP, Malik Z. S2041 COVID-19 Associated With Esophageal Hypersensitivity After Hiatal Hernia Repair and Fundoplication: A Case Report. American Journal of Gastroenterology. 2020 Oct;115(1):S1071–S1071.
47. Wiemken TL, McGrath LJ, Andersen KM, Khan F, Malhotra D, Alfred T, et al. Coronavirus Disease 2019 Severity and Risk of Subsequent Cardiovascular Events. Clinical Infectious Diseases. 2023 Feb 8;76(3):e42–50.
48. Whiteley W, Wood A. Risk of arterial and venous thromboses after COVID-19. Lancet Infect Dis. 2022 Aug;22(8):1093–4.
49. Wang MK, Yue HY, Cai J, Zhai YJ, Peng JH, Hui JF, et al. COVID-19 and the digestive system: A comprehensive review. World J Clin Cases. 2021 Jun 6;9(16):3796–813.

50. Nie K, Yang YY, Deng MZ, Wang XY. Gastrointestinal insights during the COVID-19 epidemic. World J Clin Cases. 2020 Sep 26;8(18):3934–41.
51. Kaymak T, Hruz P, Niess JH. Immune system and microbiome in the esophagus: implications for understanding inflammatory diseases. FEBS J. 2022 Aug 13;289(16):4758–72.
52. Ishak A, Mehendale M, AlRawashdeh MM, Sestacovschi C, Sharath M, Pandav K, et al. The association of COVID-19 severity and susceptibility and genetic risk factors: A systematic review of the literature. Gene. 2022 Aug;836:146674.
53. Yu X, Ho K, Shen Z, Fu X, Huang H, Wu D, et al. The Association of Human Leukocyte Antigen and COVID-19 in Southern China. Open Forum Infect Dis. 2021 Sep 1;8(9).
54. Warren RL, Birol I. HLA predictions from the bronchoalveolar lavage fluid samples of five patients at the early stage of the wuhan seafood market COVID-19 outbreak. ArXiv. 2020 Apr 15;
55. Fakhkhari M, Caidi H, Sadki K. HLA alleles associated with COVID-19 susceptibility and severity in different populations: a systematic review. The Egyptian journal of medical human genetics. 2023;24(1):10.
56. Wang ZX, Wan Q, Xing A. HLA in Alzheimer's Disease: Genetic Association and Possible Pathogenic Roles. Neuromolecular Med. 2020 Dec;22(4):464–73.
57. Yu E, Ambati A, Andersen MS, Krohn L, Estiar MA, Saini P, et al. Fine mapping of the HLA locus in Parkinson's disease in Europeans. NPJ Parkinsons Dis. 2021 Sep 21;7(1):84.
58. Al Ssadh H, Al Abdulmonem W, Rasheed Z, Madar IH, Alhoderi J, Eldeen SKN, et al. Knockdown of CD-74 in the Proliferative and Apoptotic Activity of Breast Cancer Cells. Open Access Maced J Med Sci. 2019 Oct 14;7(19):3169–76.
59. Su H, Na N, Zhang X, Zhao Y. The biological function and significance of CD74 in immune diseases. Inflammation Research. 2017 Mar 17;66(3):209–16.
60. Shin JJ, Fan W, Par-Young J, Piecychna M, Leng L, Israni-Winger K, et al. MIF is a common genetic determinant of COVID-19 symptomatic infection and severity. QJM. 2023 Mar 27;116(3):205–12.
61. Ding J, Wu W, Yang J, Wu M. Long non-coding RNA MIF-AS1 promotes breast cancer cell proliferation, migration and EMT process through regulating miR-1249-3p/HOXB8 axis. Pathol Res Pract. 2019 Jul;215(7):152376.
62. Cao X. ISG15 secretion exacerbates inflammation in SARS-CoV-2 infection. Nat Immunol. 2021 Nov 20;22(11):1360–2.
63. Chua RL, Lukassen S, Trump S, Hennig BP, Wendisch D, Pott F, et al. COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. Nat Biotechnol. 2020 Aug 26;38(8):970–9.
64. Vastrad B, Vastrad C, Tengli A. Identification of potential mRNA panels for severe acute respiratory syndrome coronavirus 2 (COVID-19) diagnosis and treatment using microarray dataset and bioinformatics methods. 3 Biotech. 2020 Oct 11;10(10):422.
65. Sfera A, Osorio C, Maguire G, Rahman L, Afzaal J, Cummings M, et al. COVID-19, ferrosenescence and neurodegeneration, a mini-review. Prog Neuropsychopharmacol Biol Psychiatry. 2021 Jul;109:110230.

66. Lu L, Liu LP, Gui R, Dong H, Su YR, Zhou XH, et al. Discovering common pathogenetic processes between COVID-19 and sepsis by bioinformatics and system biology approach. Front Immunol. 2022;13:975848.