

Expression Level Analysis of ACE2 Receptor Gene in African-American and Non-African-American COVID-19 Patients

Marion N. Nyamari¹, Kauthar M. Omar¹, Ayorinde F. Fayehun², Oumaima Dachi³, Billiah Kemunto Bwana⁴ and Olaitan I. Awe^{*5,6}

Corresponding author: laitanawe@gmail.com

¹School of Pure and Applied Sciences, Pwani University, Kilifi, Kenya

²Biotechnology Unit, Institute of Child Health, University of Ibadan, Ibadan, Oyo State, Nigeria

³National School of Arts and Crafts, University of Hassan II Casablanca, Casablanca, Morocco

⁴Department of Biological Sciences, School of Pure and Applied Sciences, University of Embu, Embu, Kenya

⁵Department of Computer Science, University of Ibadan, Ibadan, Oyo State, Nigeria

⁶African Society for Bioinformatics and Computational Biology, Cape Town, South Africa

Abstract

Background: The COVID-19 pandemic caused by SARS-CoV-2 has spread rapidly across the continents. While the incidence of COVID-19 has been reported to be higher among African-American individuals, the rate of mortality has been lower compared to that of non-African-Americans. ACE2 is involved in COVID-19 as SARS-CoV-2 uses the ACE2 enzyme to enter host cells. Although the difference in COVID-19 incidence can be explained by many factors such as low accessibility of health insurance among the African-American community, little is known about ACE2 expression in African-American COVID-19 patients compared to non-African-American COVID-19 patients. The variable expression of genes can contribute to this observed phenomenon.

Methodology: In this study, transcriptomes from African-American and non-African-American COVID-19 patients were retrieved from the sequence read archive and analyzed for ACE2 gene expression. HISAT2 was used to align the reads to the human reference genome, and HTseq-count was used to get raw gene counts. EdgeR was utilized for differential gene expression analysis, and enrichR was employed for gene enrichment analysis.

Results: The datasets included 14 and 33 transcriptome sequences from COVID-19 patients of African-American and non-African-American descent, respectively. There were 24,092 differentially expressed genes, with 7,718 upregulated (log fold change > 1 and FDR 0.05) and 16,374 downregulated (log fold change < -1 and FDR 0.05). The ACE2 mRNA level was found to be considerably downregulated in the African-American cohort (p-value = 0.0242, p-adjusted value = 0.038).

Conclusion: The downregulation of ACE2 in the African-American cohort could indicate a correlation to the low COVID-19 severity observed among the African-American community.

Keywords: ACE2 gene, SARS-CoV-2, Transcriptomics, African-American, Non-African-

American

Introduction

Coronavirus Disease 2019 (COVID-19) is a viral ailment caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Wuhan city, China was the first place where the virus was detected, before it spread to other countries throughout the world, causing a global pandemic [1]. Fever, dry cough, and malaise are common COVID-19 symptoms [1]. Difficulty breathing, chest pain as well as Acute Respiratory Distress Syndrome (ARDS) have been reported in severe cases [1].

As of August 1, 2020, the WHO recorded about 18 million COVID-19 cases and 600,000 deaths globally [2]. The total number of COVID-19 cases in the European region and United States were about 3 million and 4 million confirmed cases, out of a population of 741 million and 327 million respectively [3]. The rate of mortality among the African-American community was reported to be 25 percent lower than white and native Americans while Asian Americans had the lowest rates [4]. Areas with a large African-American population however, reported that COVID-19 patients had increased in-hospital mortality, caused by more comorbid conditions being present [5]. Older residents, people in dense housing such as the projects and African-Americans with heart disease and lacking health insurance had a doubled chance of dying from COVID-19 [6], [7]. Patient outcomes were improved by starting dialysis early and using ACE inhibitors prior to admission [5].

Asia with about 60% of the world population recorded about 28.6% of cases and 22.7% of deaths (303 deaths/million) as of 1st of April 2022 [8]. According to the same report by worldometer, North America (4.7% of world population) appears to have been the worst hit with 19.7% and 23.4% of cases and deaths respectively (3,914 deaths/million). This is closely followed by Europe (9.6% of the world population) with 36.7% of cases and 28.7% of global cases and deaths respectively (2,374 deaths/million). Study reports have shown that next-generation sequences can be used to screen for disorders in newborns [9], for the study of pathogen evolution using African SARS-CoV-2 sequences [18], prostate cancer biomarker discovery [10], malaria/CoVID-19 biomarker discovery [58], SARS-CoV-2 variants classification [59], HIV-1 evolution in sub-Saharan Africa [60], analysis of RNA-seq and ChIP-seq data [61], and ebola virus comparative genomics [62]. We will be using next-generation sequencing data for our investigation in this study.

Zeberg and his colleagues showed that the major genetic risk to COVID-19 is caused by a 50-kilobase pair genomic segment inherited from the Neanderthals, which is present in 50% and 16% of people in south Asia and Europe respectively [12]. Gene allele distribution could also explain some of the differences in COVID-19 infection incidence among people in different geographic

regions [3]. The authors however suggest that due to a lack of sufficient data, more genetic investigations are needed to corroborate this notion.

Angiotensin Converting Enzyme 2 (ACE2) hydrolyzes Ang-II to produce Ang- (1-7) - a vasodilator and anti-inflammatory agent, by. ACE2 promotes a protective anti-inflammatory response in the lung by lowering oedema, vascular permeability, and pulmonary cell damage. ACE2 promotes a protective anti-inflammatory response in the lung by lowering oedema, pulmonary cell damage and vascular permeability [16], [39], [50]. When the expression of ACE2 receptors is reduced, its substrate Ang-II, accumulates leading to increased vascular permeability and pulmonary oedema [3].

While ACE supports vasoconstriction and inflammation through Ang-II induction, ACE2 promotes vasodilation and anti-inflammatory responses through Ang-II breakdown. Thus, ACE and ACE2 play opposite roles in the generation and degradation of Ang-II and RAS tissue inflammatory balance, and ACE/ACE2 imbalance is believed to play a role in COVID-19 pathogenesis [17].

Genetics is believed to have a substantial control on plasma ACE levels [38]. There is a genetic deletion/insertion polymorphism of ACE enzyme, which is linked to half of the variation in the total plasma concentration of ACE [40]. The ACE I/D polymorphism of the Alu repeat in intron 16, consists of either an insertion (I) allele or a deletion (D) allele of *Alu* repeat, leading to three possible genotypes: II (homozygous insertion), ID (heterozygous - insertion of Alu repeat in one allele and deletion of the Alu repeat in the other allele) and DD (homozygous deletion) [40]. Individuals with D/D genotype were reported to have significantly higher serum ACE concentrations compared to the I/D and I/I genotypes [40]. The dominant DD allele is associated with elevated levels of serum and tissue ACE, and subsequently increase in Ang-II levels [3].

Delanghe and his colleagues showed that variation in the distribution of D/I genotype was associated with prevalence of COVID-19 in 25 different European countries, where countries with higher ACE D allele frequency had lower prevalence of COVID-19 [54]. Similarly, nations with high COVID-19 prevalence, such as China and Korea, which were both heavily affected by the virus in the outset, have low ACE D allele frequencies [43].

SARS-CoV-2 virus enters host cells via the ACE2 receptor [20], [3]. ACE2 receptor is a type I transmembrane amino-peptidase found on cell surfaces of the gut, heart, lungs, and many other organs, and also found in solution in blood plasma and urine [25], [3]. The COVID-19 viral envelope's "spike" (S) protein binds to ACE2 receptors on nasopharyngeal mucosa and alveolar pneumocytes [20], [21], [51].

ACE2 expression is regulated by serum level of Ang-II through the AT1R, as Ang-II induces ACE in a dose dependent manner and decreases ACE2 expression [50]; Since ACE induces the production of Ang-II, high levels of Ang-II could therefore reduce ACE2 expression and activity.

Therefore, in ACE-DD homozygous individuals with increased ACE activities and Ang-II level, ACE2 expression might be down-regulated [17]. Furthermore, populations with this DD polymorphism may have reduced ACE2 expression which may play a protective role against COVID-19.

ACE2 genes contain 18 exons and maps on Xp22. Genetic variations in ACE2 gene have been proposed to also affect the COVID-19 viral spike protein interaction with the ACE2 receptor [31]. Several ACE2 gene variations can affect interaction with spike protein, protein stability, or change in ligand receptor affinity [19]. Therefore it can be inferred that genetic variations in the ACE2 gene may influence the individual's susceptibility or resistance to SARS-CoV-2 [19], [23]. Search for variants that show a correlation with clinical COVID-19 severity may explain the broad range of SARS-CoV-2 infection outcome. COVID-19. CoVID-19 symptoms and outcomes might be associated with the pattern and level of human ACE2 enzyme expression in different tissues [53]. Also, circulating ACE2 might inhibit SARS-COV-2 entry into pulmonary alveoli and CoVID-19 infection severity correlating with soluble/membrane bound ratio of ACE2 [15].

ACE2 polymorphism, which reduces ACE2 receptor expression, is consequently expected to be associated with increased Ang-II levels. Population distribution of these alleles may also explain the distribution of CoVID-19. However, many other studies have not been able to find this association. For example, a study of three ACE2 polymorphisms (rs2106809, rs4646155, and rs879922) in hypertensive patients did not show any significant relationship with circulating levels of ACE2 [32]. Also, an Italian study did not find a strong association between COVID-19 disease severity and ACE2 variants [35]. Furthermore, the SARS-CoV-2 virus may have an impact on the expression of ACE2 receptors, where it is downregulated in infected cells [28]. Also, the function of the ACE2 is likely impaired during the initial SARS-CoV-2 virus binding from ACE2 steric domain hindrance or ACE2 downregulation [13].

In the light of the potential role of ACE2 and Ang-II in the SARS-CoV-2 pathogenesis the quantification of ACE2 and Ang-II should be part of the COVID-19 patients' biological assessment and monitoring [13]. In this study, we compared the ACE2 gene expression level in patients of African-American descent with COVID-19, to that of non-African-American COVID-19 patients.

Methods

Data Acquisition

Transcriptomic data of both African-American and Non-African-American populations was downloaded from NCBI. The reads were downloaded via their accession numbers from the command line using `Fasterq-dump` (<https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software>). The non-African-American population was obtained from bioproject [PRJNA743862](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA743862) ; GEO: GSE179448 while the African-American project was obtained from bioproject [PRJNA679264](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA679264); GEO: GSE161731. The African-American dataset included patients with COVID-19 as well as seasonal COVID.

Quality Control

FastQC version 0.11.3 was used to verify the quality of raw sequencing reads [46], and subsequently trimming to eliminate adapter sequences was done Cutadapt [33].

Alignment to the human reference genome

After passing the quality check, the reads were aligned to the human reference genome (hg38) using Hisat2 [26]. Samtools v1.3.1 [30] was used to convert the aligned reads' SAM file to a BAM file, which was then sorted.

Read counts

HTSeq-count [37] was used to count reads associated with each human gene.

Normalization

EdgeR [41] was used to normalize library sizes using Trimmed Mean of M-values (TMM) by determining the scaling factor to obtain effective library sizes. Counts were then adjusted by effective library sizes to create counts per million (cpm) values.

Differential gene expression analysis

EdgeR [41] was used to perform differential expression analysis according to established principles. Genes having corrected P-values of 0.05 and absolute fold increases of 1.5 were considered differentially expressed. Differentially expressed candidate genes were evaluated between African-American and non-African-American populations. Enrichr was used to perform Gene Ontology (GO) and KEGG pathway enrichment analysis of differentially expressed genes [14]. The data analysis workflow is shown in Figure 1 below

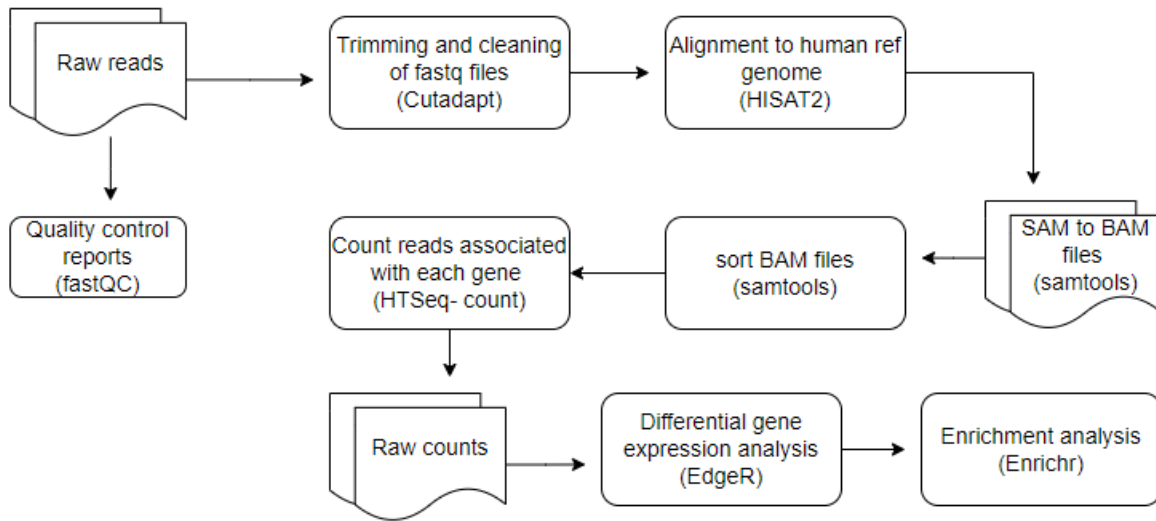


Figure 1: Materials and methods used in data analysis

Results

Quality control

The datasets included 14 and 33 transcriptomic sequences of African-American and non-African-American COVID-19 patients, respectively. The sequences were of high quality, all the sequences scored a Phred score of above 30 after trimming of the adaptor sequence.

Identification of differentially expressed genes in the African-American population

The raw gene count matrix consisting of a total of 39,590 genes generated in the gene count step of our pipeline, was used in EdgeR to analyze differential expression levels of the genes. Differential expression analysis revealed 7,718 out of 24,092 genes were significantly upregulated (log fold change > 1 and FDR < 0.05) while 16,374 genes were significantly downregulated (log fold change < -1 and FDR < 0.05) between African-American and non-African-American populations (controls) (Figure.2 and 3). We discovered that the mRNA level of ACE2 was significantly downregulated in the African-American cohort. (p value = 0.0242, p-adjusted value = 0.038).

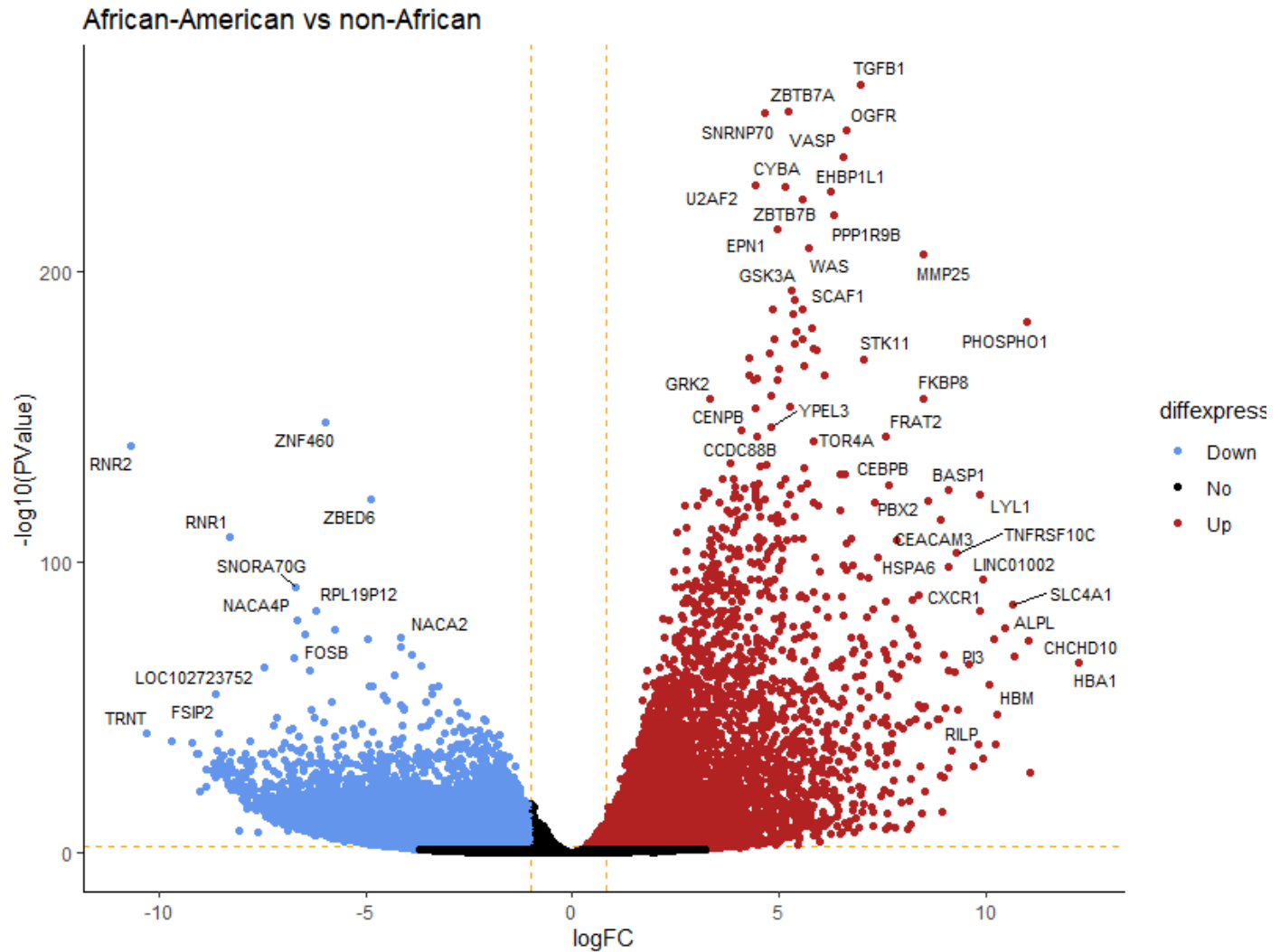


Figure 2. Volcano Plot for differentially expressed genes between African- American COVID-19 patients vs non-African-American COVID-19 patients as controls. Scattered dots represent genes; the x-axis is the log fold change for the ratio of African-American vs non-African-American COVID-19 patients, whereas the y-axis is the $-\log_{10}(P \text{ Value})$ where P is the probability that a gene has statistical significance in its differential expression. Red dots represent genes significantly upregulated, blue dots are genes significantly suppressed genes while black dots are genes with non-significant changes between the conditions

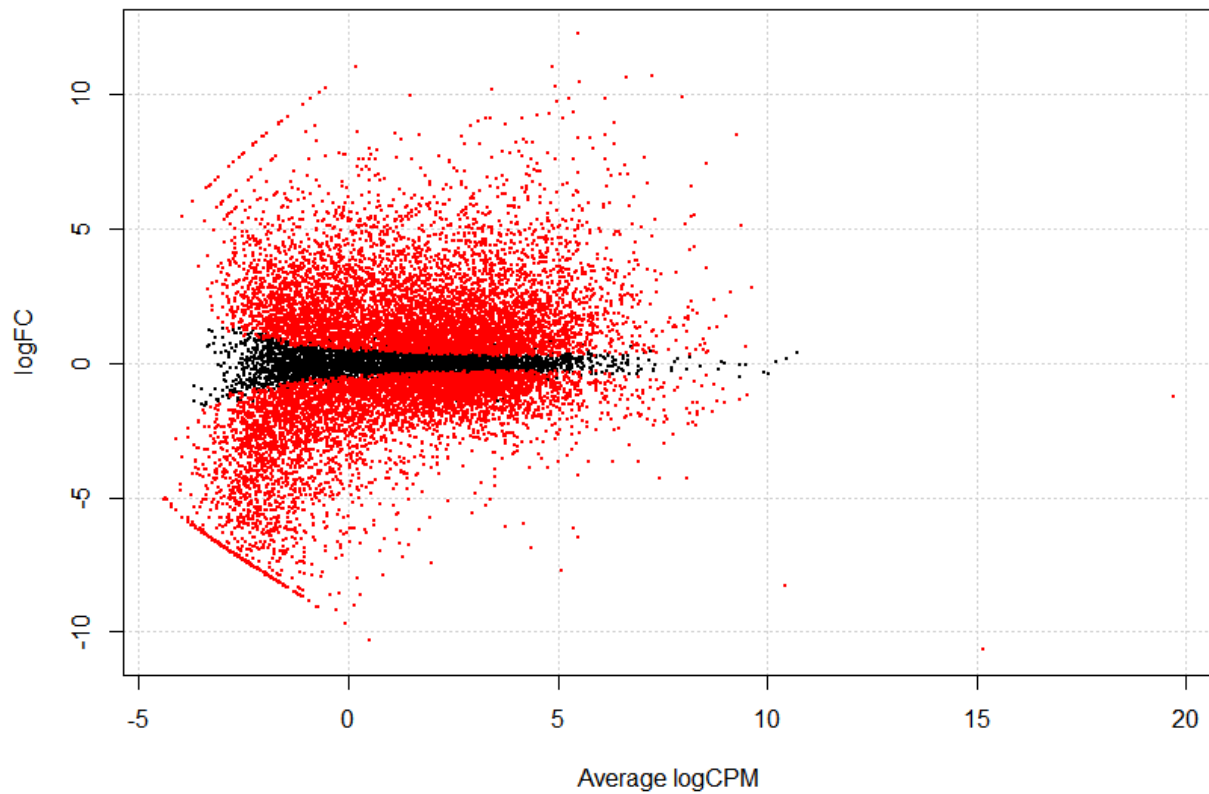


Figure 3. MA plot for transcript distribution determined by EdgeR. Red points represent differentially expressed transcripts (P-adjusted < 0.05).

The top 10 most significantly upregulated genes include TGFB1, ZBTB7A, SNRNP70, OGFR, VASP, U2AF2, CYBA, EHBP1L1, ZBTB7B, and PPP1R9B with a P-adjusted value < 0.05 (Table 1). The top 10 most significantly down regulated genes include ZNF460, RNR2, ZBED6, RNR1, SNORA70G, RPL19P12, NACA4P, RPL23P8, FOSB, NACA2 with a P-adjusted value < 0.05 (Table 2).

Table. 1 Top 10 significantly upregulated genes between African-American COVID-19 patients vs non-African-American COVID-19 patients

Gene	Annotation*	Log Fold Change	P-Value	FDR P-Value
TGFB1	Transforming Growth Factor Beta 1	6.947	4.91E-265	1.83E-260
ZBTB7A	Zinc Finger and BTB Domain Containing 7A	5.201	6.88E-256	1.28E-251
SNRNP70	Small Nuclear Ribonucleoprotein U1 Subunit 70	4.643	6.98E-255	8.66E-251
OGFR	Opioid Growth Factor Receptor	6.609	5.98E-249	5.57E-245
VASP	Vasodilator Stimulated Phosphoprotein	6.547	5.25E-240	3.91E-236
U2AF2	U2 Small Nuclear RNA Auxiliary Factor 2	4.408	1.64E-230	1.02E-226
CYBA	Cytochrome B-245 Alpha Chain	5.159	5.82E-230	3.10E-226
EHBP1L1	EH Domain Binding Protein 1 Like 1	6.234	4.15E-228	1.93E-224
ZBTB7B	Zinc Finger and BTB Domain Containing 7B	5.573	1.44E-225	5.96E-222
PPP1R9B	Protein Phosphatase 1 Regulatory Subunit 9B	6.335	8.13E-220	3.03E-216

Table. 2 Top 10 significantly downregulated genes between African-American COVID-19 patients vs non-African-American COVID-19 patient

Gene	Annotation*	Log Fold Change	P-Value	FDR P-Value
ZNF460	Zinc Finger Protein 460	-5.95926	7.16E-149	6.35E-146
RNR2	RNA, Ribosomal 45S Cluster 2	-10.6667	1.08E-140	8.35E-138
ZBED6	Zinc Finger BED-Type Containing 6	-4.86911	1.47E-122	7.02E-120
RNR1	RNA, Ribosomal 45S Cluster 1	-8.296	2.07E-109	6.75E-107
SNORA70G	Small Nucleolar RNA, H/ACA Box 70G	-6.69691	4.49E-92	8.25E-90
RPL19P12	Ribosomal Protein L19 Pseudogene 12	-6.19293	3.18E-84	4.60E-82
NACA4P	NACA family member 4, Pseudogene	-6.63674	1.10E-80	1.42E-78
RPL23P8	Ribosomal Protein L23 Pseudogene 8	-5.74711	9.70E-78	1.17E-75
FOSB	FosB Proto-Oncogene, AP-1 Transcription Factor Subunit	-6.47326	3.58E-76	4.19E-74
NACA2	Nascent Polypeptide Associated Complex Subunit Alpha 2	-4.14031	9.03E-75	1.02E-72

* - Gene name in NCBI database; FDR – False Detection Rate corrected; ** - transcriptional changes of individual genes quantified in African-American COVID-19 patients compared to non-African-American COVID-19 patients as controls. The fold-changes are computed ratios of gene expression in African-American COVID-19 patients relative to similar expressions in the respective non-African-American COVID-19 patients. Consequently, positive

and negative values indicate respective induced or suppressed genes relative to control

An unsupervised hierarchical analysis revealed that sets of genes with comparable expression characteristics clustered together, as expected. Groups of genes are systematically, either up-regulated (red) or down-regulated (green) in African-Americans versus non-African-Americans (Figure 4).

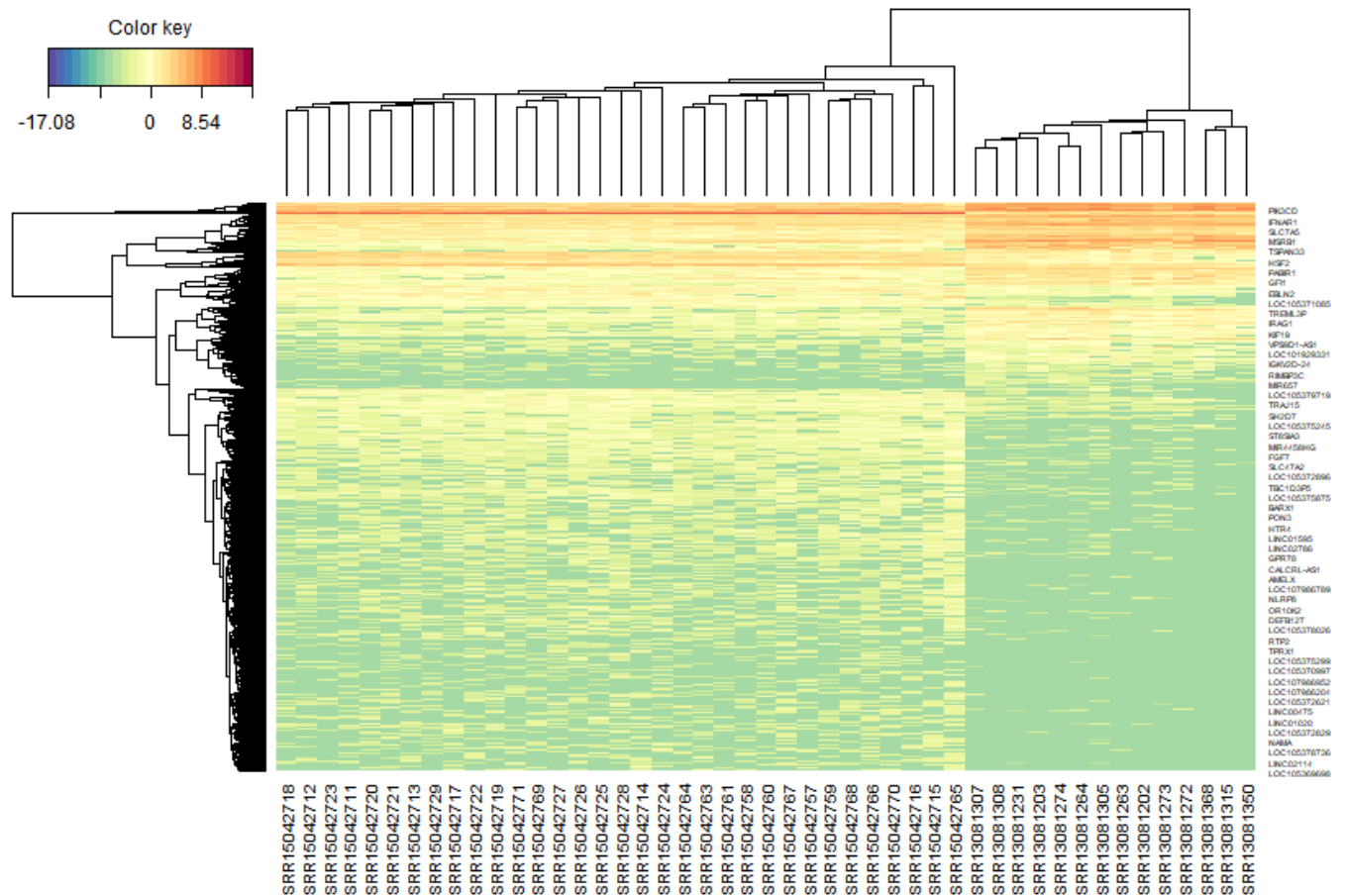


Figure 4. HeatMap for hierarchical clustering of differential gene expression between African-American COVID-19 patients vs non-African-American COVID-19 patients as controls. Groups of genes are systematically, either up-regulated (red) or down-regulated (green) in treatment versus controls.

Functional analysis of significantly upregulated genes

Enrichment analysis was performed on significantly over-expressed genes to verify the molecular functions, biological processes, cellular components, and pathways associated with the differentially expressed genes. The fifteen most significant molecular functions, biological processes, cellular components, and ten most significant pathways associated with the differentially expressed genes are represented according to Gene Ontology (GO) biological process, Gene Ontology (GO) molecular function, Gene Ontology (GO) cellular component, and KEGG pathway analysis based on p value (Figure 5A- 5D).

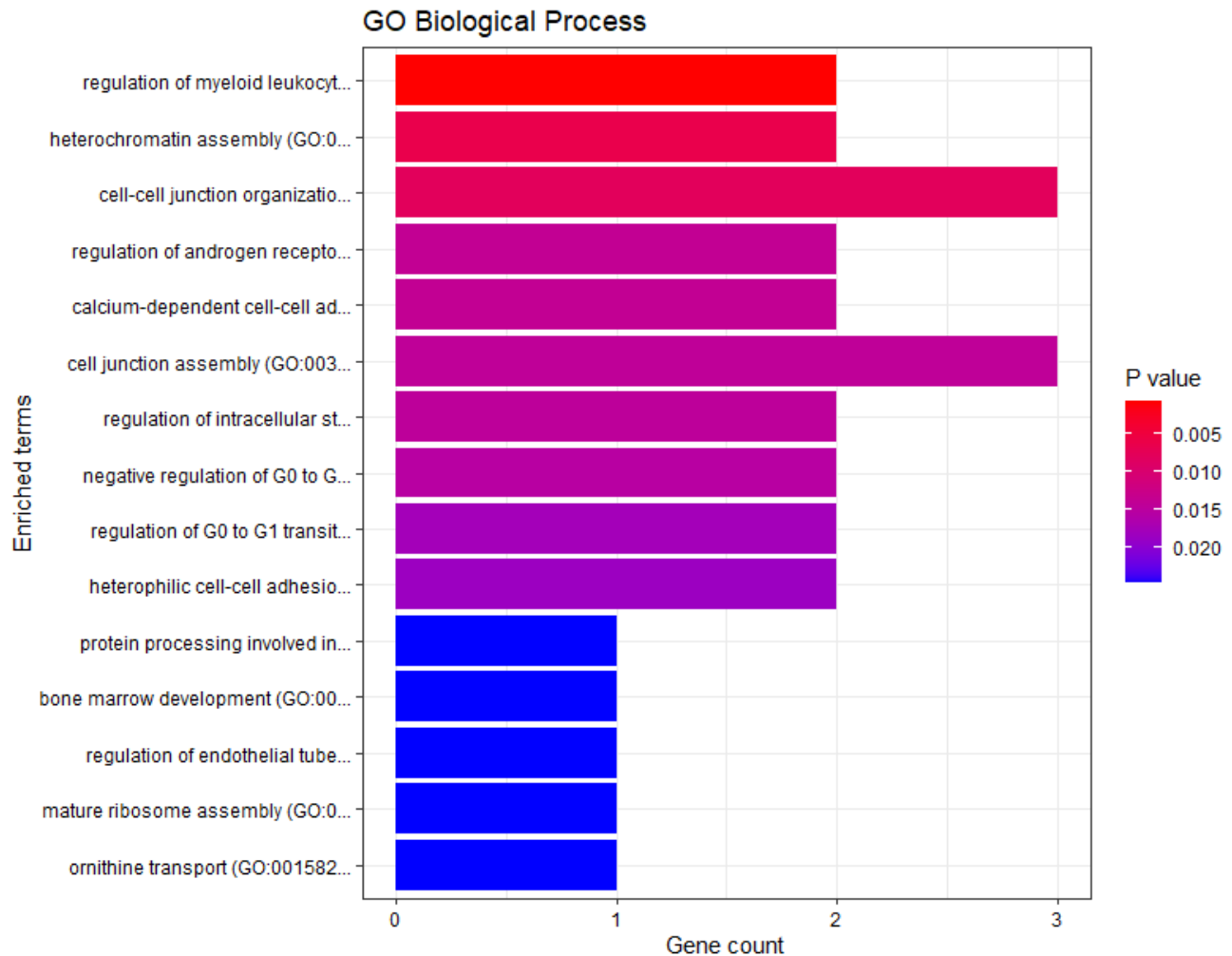


Figure 5A. GO biological process

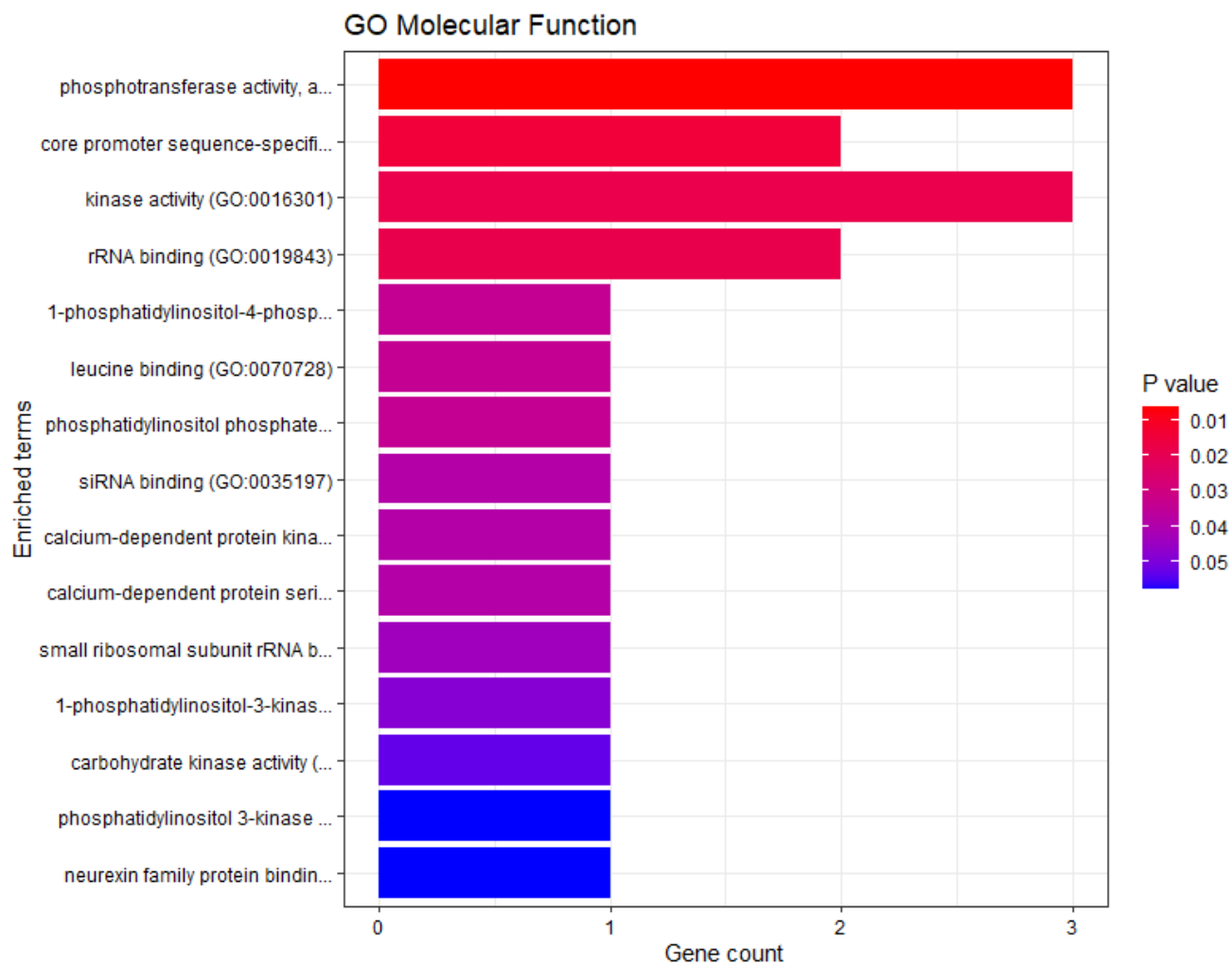


Figure 5B. GO molecular function

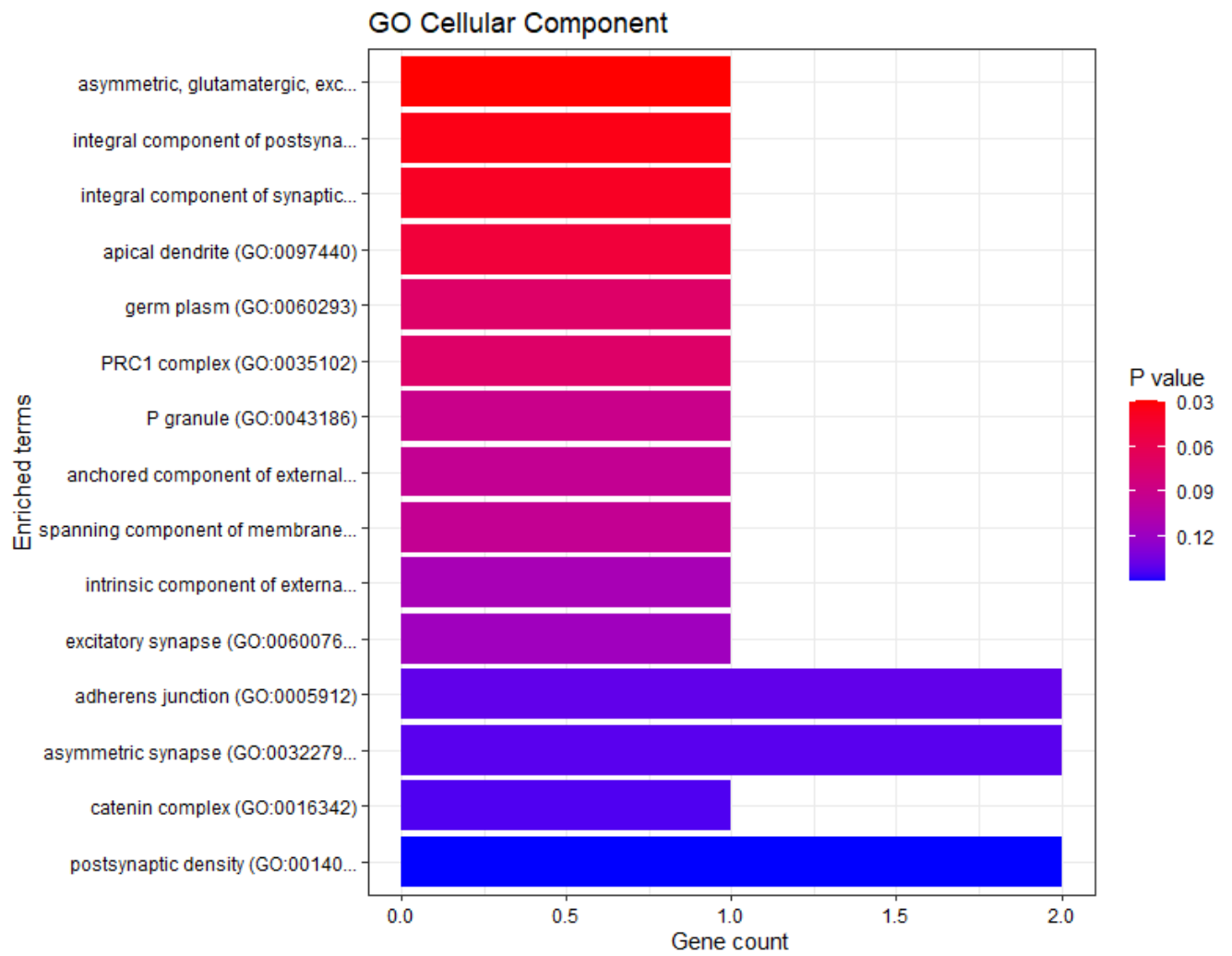


Figure 5C. GO cellular component

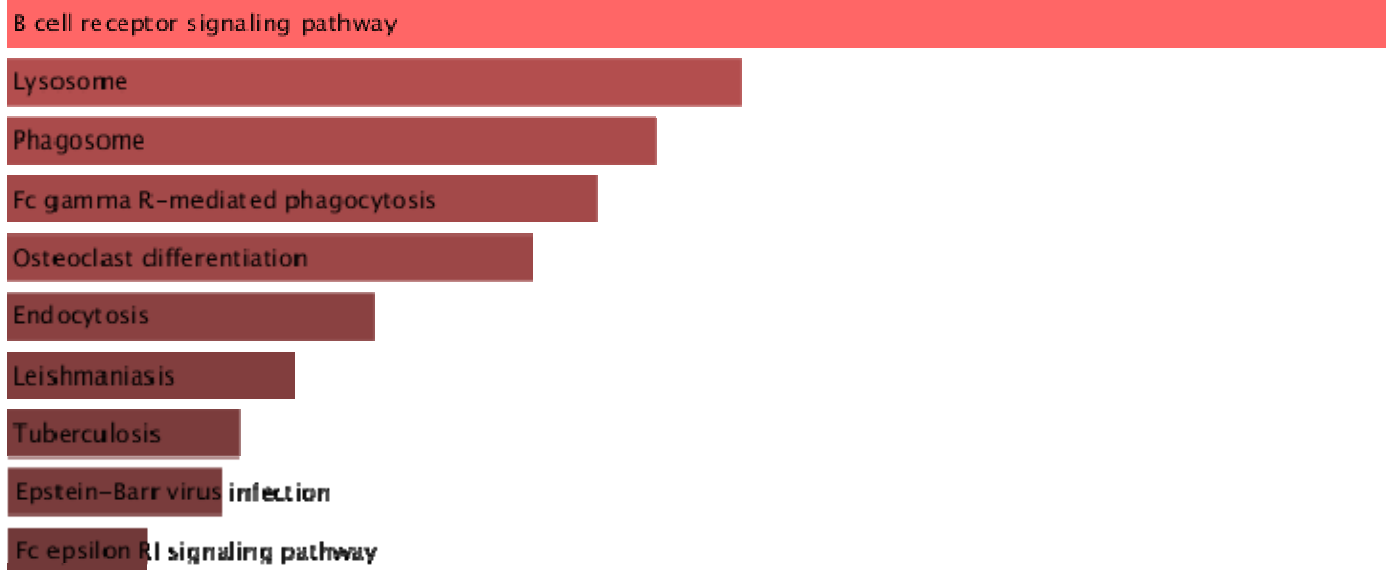


Figure 5D. KEGG pathways

*ENRICH Gene Ontology and KEGG pathways analysis of significantly over-expressed genes of the African-American population.

Significantly enriched KEGG pathways

Tuberculosis, Epstein- Barr virus infection, Influenza A, PD-L1 expression and PD-1 checkpoint pathway in cancer, B cell receptor signaling pathway, T cell receptor signaling pathways were observed to be significantly enriched in the upregulated genes, among other pathways (Figure. 5E) with multiple overlap genes (Table 3).

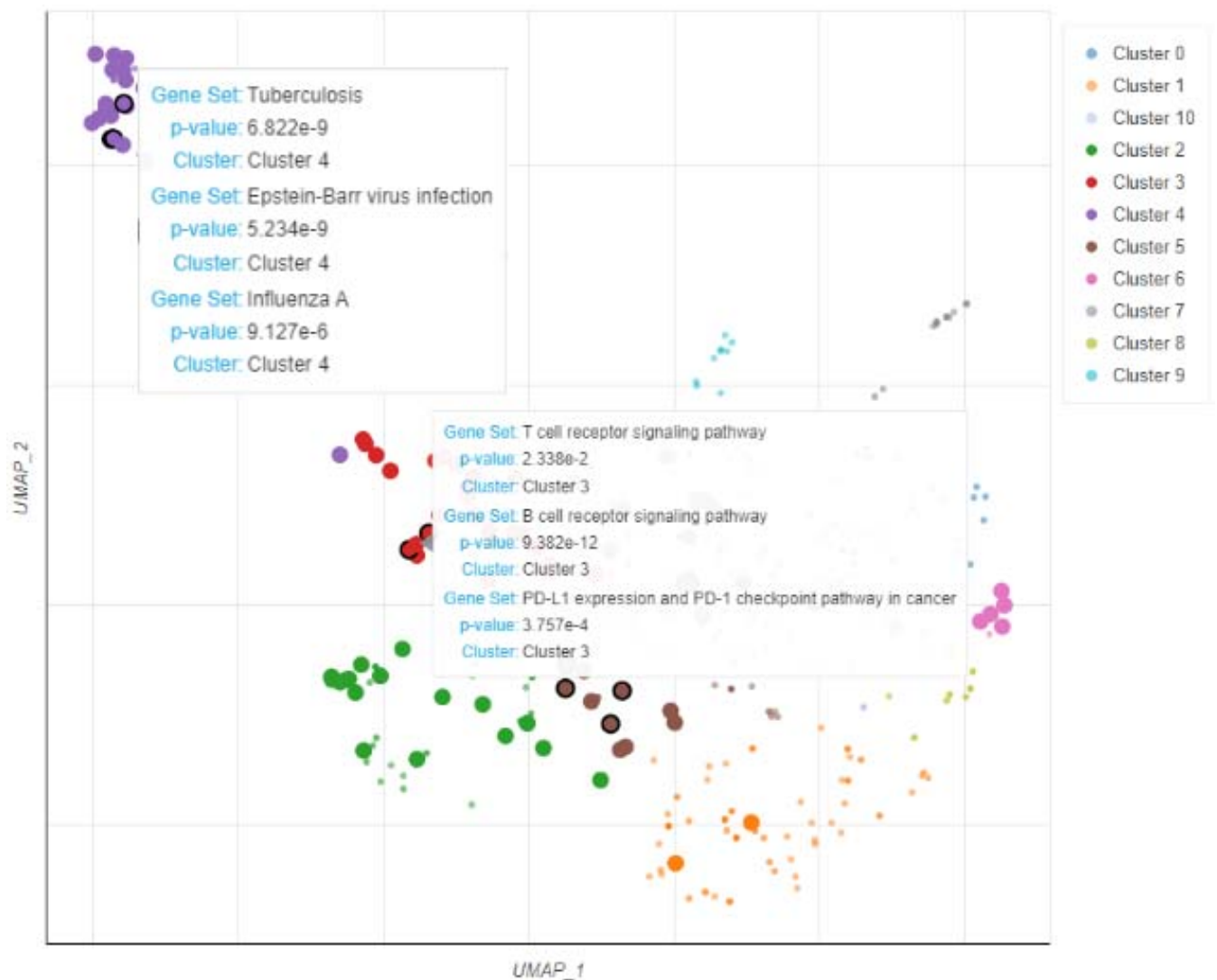


Figure 5E. Scatterplot of all terms in the KEGG_2021_Human gene set library. Each point represents a term in the library.

Table. 3 Top ten KEGG 2021 Human significant p-values and q-values, 2021

Terms	p-value	q-value	Overlap Genes
B cell receptor signaling pathway	3.67E-12	1.17E-09	[GSK3B, IFITM1, CD81, INPPL1, PIK3CD, PIK3CB, IKBKB, AKT2, AKT3, BLNK, RAC2, AKT1, RAC3, IKBKG, RAC1, HRAS, VAV3, CR2, MAP2K1, MAP2K2, SYK, PRKCB, VAV1, VAV2, BTK, RAF1, SOS2, CARD11, LILRA6, DAPP1, PIK3R2, LILRA1, LILRA2, LILRA3, LILRA4, RELA, LILRA5, RASGRP3, CD79B, CD79A, PPP3R1, CD19, INPP5D, PLCG2, MAPK1, MAPK3, LYN, CD72, LILRB1, LILRB2, LILRB3, LILRB4, LILRB5, NFKB1, GRB2, KRAS, PTPN6, FCGR2B, NFKBIE, PIK3AP1, CD22, NFKBIB]
Phagosome	9.23E-12	1.47E-09	[SCARB1, NCF1, NCF2, NCF4, MPO, ACTG1, TUBB8, TUBB6, LAMP1, LAMP2, ATP6V1E1, HLA-DOA, ATP6V1E2, HLA-DOB, ATP6V0E1, STX7, HLA-B, TAP2, CYBB, HLA-C, TAP1, CYBA, HLA-A, HLA-F, TUBA4A, HLA-E, DYNC1LI1, HGS, ATP6V0D1, ATP6V1A, RAB5B, RAB5C, CORO1A, C3, CLEC7A, ATP6V0A2, CD14, ATP6V1D, HLA-DQA2, ATP6V1C1, HLA-DQA1, ATP6V0A1, ATP6V1F, MSR1, HLA-DRB4, M6PR, MARCO, HLA-DRA, CALR, HLA-DRB3, RAB5A, TUBA8, ITGAM, ITGB5, ITGB2, TCIRG1, CTSS, FCAR, MRC2, SEC61A1, TUBA1C, FCGR3A, TUBA1B, FCGR3B, TUBA1A, CTSL, OLR1, CD36, RAC1, HLA-DPA1, ATP6V0B, ATP6AP1, TUBB, SFTPD, TUBB2B, TUBB2A, CANX, ATP6V1B2, ITGA5, TLR6, SEC22B, ATP6V0C, TLR4, HLA-DQB1, RAB7A, TLR2, VAMP3, DYNC1I2, STX12, C1R, THBS3, HLA-DMB, FCGR1A, RILP, TUBB4B, TUBB4A, FCGR2A, HLA-DPB1, FCGR2B, FCGR2C]
Lysosome	1.45E-11	1.54E-09	[SCARB2, HEXB, CLTC, CTSZ, CLTB, CLTA, CTSW, TCIRG1, LITAF, CTSS, AP4M1, LAPTM4B, LAPTM4A, CTSO, LAMP1, NAGLU, GM2A, CTSL, HYAL1, LAMP3, HYAL2, SMPD1, LAMP2, HYAL3, AP1S2, AP1S1, ACP5, CTSH, ACP2, CTSG, CTSF, CTSE, GUSB, CTSD, CTSC, CTSB, ARSA, HGSNAT, ATP6V0B, MANBA, ATP6AP1, SLC11A1, GAA, AP1B1, AP3B2,

			<p>PLA2G15, NPC2, DNASE2, ATP6V0D1, ATP6V0C, CD63, ASAH1, IDUA, GBA, ABCB9, GNS, GGA2, AP3M1, CLN3, GGA1, GGA3, CLTCL1, NEU1, PSAP, ATP6V0A2, ARSG, AP1M2, ARSB, ATP6V0A1, CTSA, ABCA2, NAPSA, SORT1, FUCA1, M6PR, AP3D1, LAPTM5, CTNS, NAGA, MCOLN1, IGF2R, GALNS, GNPTG, PPT1, MAN2B1, CD68, GLA]</p>
Endocytosis	2.66E-11	2.12E-09	<p>[WIPF1, WIPF2, ARPC5L, WASHC2C, WASHC2A, ZFYVE27, VPS35, WASHC1, AP2M1, SH3GL1, SH3GLB1, SH3GLB2, USP8, HLA-B, HLA-C, HLA-A, HLA-F, HLA-E, RNF41, RUFY1, ACAP3, ZFYVE16, ACAP1, RAB31, HGS, RAB35, CHMP4B, CHMP4A, VPS25, STAM2, VPS28, RAB5B, RAB5C, TSG101, AGAP2, AGAP1, NEDD4L, AGAP4, AGAP3, VPS26B, PLD1, PRKCZ, PLD2, SNX3, SNX1, GRK3, RAB11FIP1, GRK2, GRK5, CLTCL1, GRK6, PIP5K1B, PIP5K1C, RAB11FIP3, RAB11FIP4, RAB11FIP5, RAB4A, SMAD3, IQSEC2, SMURF2, IQSEC1, SMURF1, CAV1, IQSEC3, HSPA6, ARPC4, SNF8, ARPC5, IGF2R, EHD1, EHD4, ARPC2, ARPC3, CHMP2A, RAB5A, RNF103-CHMP3, ARF3, SH3KBP1, ARPC1B, ARPC1A, CLTC, CLTB, SNX12, CLTA, AP2A1, ARRB1, ARRB2, AP2A2, RAB22A, CAPZB, PSD4, CHMP1A, CCR5, LDLRAP1, HRAS, GIT1, RAB8A, ACTR2, VPS37C, ARAP3, ...]</p>
Osteoclast differentiation	2.77E-10	1.77E-08	<p>[SPI1, CSF1, NCF1, NCF2, NCF4, FHL2, PIK3CD, PIK3CB, SIRPB1, IKBKB, FCGR3A, FCGR3B, AKT2, AKT3, BLNK, AKT1, SIRPA, ACP5, IKBKG, RAC1, JUNB, IFNAR2, MAP2K1, SYK, IL1R1, IFNGR1, IFNGR2, CYBA, TRAF2, GAB2, TYK2, TGFBR1, OSCAR, TGFBR2, TNFRSF1A, TYROBP, BTK, LCP2, PPARG, SQSTM1, IRF9, IFNAR1, LILRA6, CSF1R, PIK3R2, LILRA1, LILRA2, LILRA3, LILRA4, RELA, LILRA5, RELB, SOCS3, PPP3R1, SOCS1, PLCG2, MAPK1, FYN, MAP2K7, FCGR1A, MAP2K6, MAPK3, JUND, TGFB1, STAT1, STAT2, LILRB1, LILRB2, LILRB3, MAPK14, LILRB4, NFKB1, LILRB5, MAPK12, FOSL2, NFKB2, MAPK13, MAPK11, FCGR2A, GRB2, TAB1, FCGR2B, MAP3K14, FCGR2C]</p>
Fc gamma R-mediated phagocytosis	1.09E-09	5.78E-08	<p>[NCF1, ARPC1B, ARPC1A, ARPC5L, INPPL1, PIK3CD, PIK3CB, CRKL, FCGR3A, FCGR3B, AKT2, AKT3, CFL1, RAC2, AKT1, RAC1, VAV3, ACTR2, MAP2K1, SYK, PRKCB, SPHK2, SPHK1, PRKCD, PLA2G4C, PLA2G4A, GAB2, VAV1, DNMT2, VAV2, HCK, RAF1, CRK, ARF6, ASAP3, WAS, PIK3R2, ASAP1,</p>

			PLA2G6, PLD1, PLD2, PAK1, INPP5D, PLCG2, MAPK1, PIP5K1B, PIP5K1C, PLCG1, FCGR1A, WASF2, MAPK3, LYN, VASP, MARCKSL1, GSN, MYO10, LIMK2, LIMK1, ARPC4, ARPC5, MARCKS, PTPRC, FCGR2A, ARPC2, ARPC3, RPS6KB2, FCGR2B]
Epstein-Barr virus infection	3.48E-09	1.59E-07	[RBPJ, ICAM1, PSMD8, CCND3, PSMD4, CCND1, AKT2, PSMD2, AKT3, PSMD3, AKT1, CIR1, HLA-DOA, B2M, HLA-DOB, IFNAR2, MAP2K3, MAP2K4, ENTPD1, USP7, HLA-B, TAP2, HLA-C, TAP1, HLA-A, HLA-F, RUNX3, HLA-E, SAP30, BTK, TP53, IFNAR1, PIK3R2, IRAK1, CD19, PLCG2, RIPK1, FADD, MAP2K7, HLA-DQA2, HLA-DQA1, MAP2K6, LYN, HLA-DRB4, GADD45B, GADD45A, EIF2AK2, ISG15, NFKB1, GADD45G, NFKB2, CXCL10, MAVS, HLA-DRA, CALR, TAB1, HLA-DRB3, NFKBIE, MAP3K14, MYD88, NFKBIB, CD40, CDKN1B, TRADD, PIK3CD, PIK3CB, CD3E, ITGAL, IKBKB, CASP9, CASP8, CASP3, BLNK, IKBKG, RAC1, JAK3, HLA-DPA1, CR2, SYK, APAF1, DDX58, TRAF2, TYK2, IRAK4, TAPBP, NCOR2, FCER2, CCNA1, IRF3, OAS1, OAS2, OAS3, IRF7, IRF9, HLA-DQB1, TLR2, RELA, RELB, HLA-DMB, SAP30L, ...]
Tuberculosis	4.42E-09	1.76E-07	[CALML4, LAMP1, AKT2, AKT3, LAMP2, AKT1, HLA-DOA, HLA-DOB, RFX5, ATP6V0D1, CLEC4E, RAF1, RAB5B, RAB5C, CORO1A, C3, PPP3R1, IRAK1, CLEC7A, IRAK2, ATP6V0A2, CD14, FADD, HLA-DQA2, HLA-DQA1, ATP6V0A1, CREBBP, HLA-DRB4, TGFB1, VDR, RFXANK, NFKB1, RFXAP, HLA-DRA, CALM3, HLA-DRB3, RAB5A, MYD88, CIITA, ITGAM, TRADD, ITGB2, LSP1, TCIRG1, CTSS, CASP9, MRC2, FCGR3A, CASP8, FCGR3B, CASP10, CASP3, ITGAX, JAK2, CTSD, CAMP, HLA-DPA1, ATP6V0B, ARHGEF12, CR1, FCER1G, APAF1, ATP6AP1, KSR1, SYK, IFNGR1, SPHK2, IFNGR2, SPHK1, IL18, IRAK4, RHOA, TIRAP, TNFRSF1A, TLR1, TLR9, TLR6, ATP6V0C, TLR4, HLA-DQB1, RAB7A, TLR2, CEBPB, SRC, NOD2, RELA, HLA-DMB, MAPK1, BID, FCGR1A, CAMK2G, MAPK3, CD74, BAD, IL10RB, STAT1, NFYC, CARD9, MAPK14, MAPK12, ...]
Salmonella infection	2.28E-08	8.07E-07	[CYFIP1, FHOD1, ARPC5L, ACTG1, TXN2, PYCARD, TUBB8, TUBB6, AKT2, AKT3, TNFSF10, AKT1, KPNA4, VPS39, KPNA1, MAP2K3, VPS18, MAP2K4, MAP2K1, MAP2K2, TUBA4A, DYNC1L1, MYL5, PFN1, RAF1, MYL9, NAIP,

ARL8A, RAB5B, RAB5C, LY96, KLC1, HSP90B1, KLC4, KLC3, KLC2, RRAS, IRAK1, VPS11, NLRP3, NCKAP1L, RIPK1, CD14, FADD, MAP2K7, VPS16, MAP2K6, GSDMD, MLKL, M6PR, TNFRSF10B, ARPC4, ARPC5, NFKB1, PTPRC, ARPC2, ARPC3, PLEKHM1, PLEKHM2, DYNLRB1, TAB1, ACBD3, RAB5A, MYD88, TUBA8, ARPC1B, TRADD, ARPC1A, PIK3CD, PIK3CB, PIK3CG, IKBKB, TUBA1C, ACTR1A, TUBA1B, TUBA1A, CASP8, CASP5, CASP3, CASP4, CASP1, IKBKG, RAC1, HRAS, ACTR2, RIPK3, EXOC7, TUBB, IL18, RHOG, TRAF2, IRAK4, DYNLL1, DYNLL2, RHOA, TIRAP, RHOB, DNM2, TNFRSF1A, TUBB2B, ...]

Yersinia infection 1.45E-07 4.61E-06

[GSK3B, WIPF1, WIPF2, ARPC1B, ARPC1A, ARPC5L, PIK3CD, PIK3CB, ACTG1, CRKL, PYCARD, IKBKB, RPS6KA3, AKT2, AKT3, RPS6KA1, CASP1, RAC2, AKT1, RAC3, IKBKG, RAC1, MAP2K3, VAV3, MAP2K4, ACTR2, MAP2K1, MAP2K2, ARHGEF12, IL18, RHOG, TRAF2, IRAK4, TICAM1, RHOA, VAV1, VAV2, CD8B, IRF3, CD8A, ARHGEF1, LCP2, ITGA5, CRK, TLR4, SKAP2, ARF6, SRC, PXN, WAS, PIK3R2, NLRC4, MEFV, RELA, IRAK1, NLRP3, CCL2, MAPK1, PTK2B, PIP5K1B, PIP5K1C, PLCG1, MAP2K7, WASF2, MAP2K6, MAPK3, GIT2, LIMK1, ARPC4, ARPC5, BAIAP2, MAPK14, NFKB1, MAPK12, MAPK13, MAPK11, FCGR2A, ARPC2, ARPC3, GNAQ, TAB1, PKN1, MYD88]

Functional analysis of significantly downregulated genes

To validate the molecular functions, biological processes, cellular components, and pathways linked with the differentially expressed genes, enrichment analysis was done on significantly under-expressed genes. According to Gene Ontology (GO) biological process, Gene Ontology (GO) molecular function, Gene Ontology (GO) cellular component, and KEGG pathway analysis based on p value, the fifteen most significant molecular functions, biological processes, cellular components, and ten most significant pathways associated with the differentially expressed genes are represented. (Figure 6A- 6D).

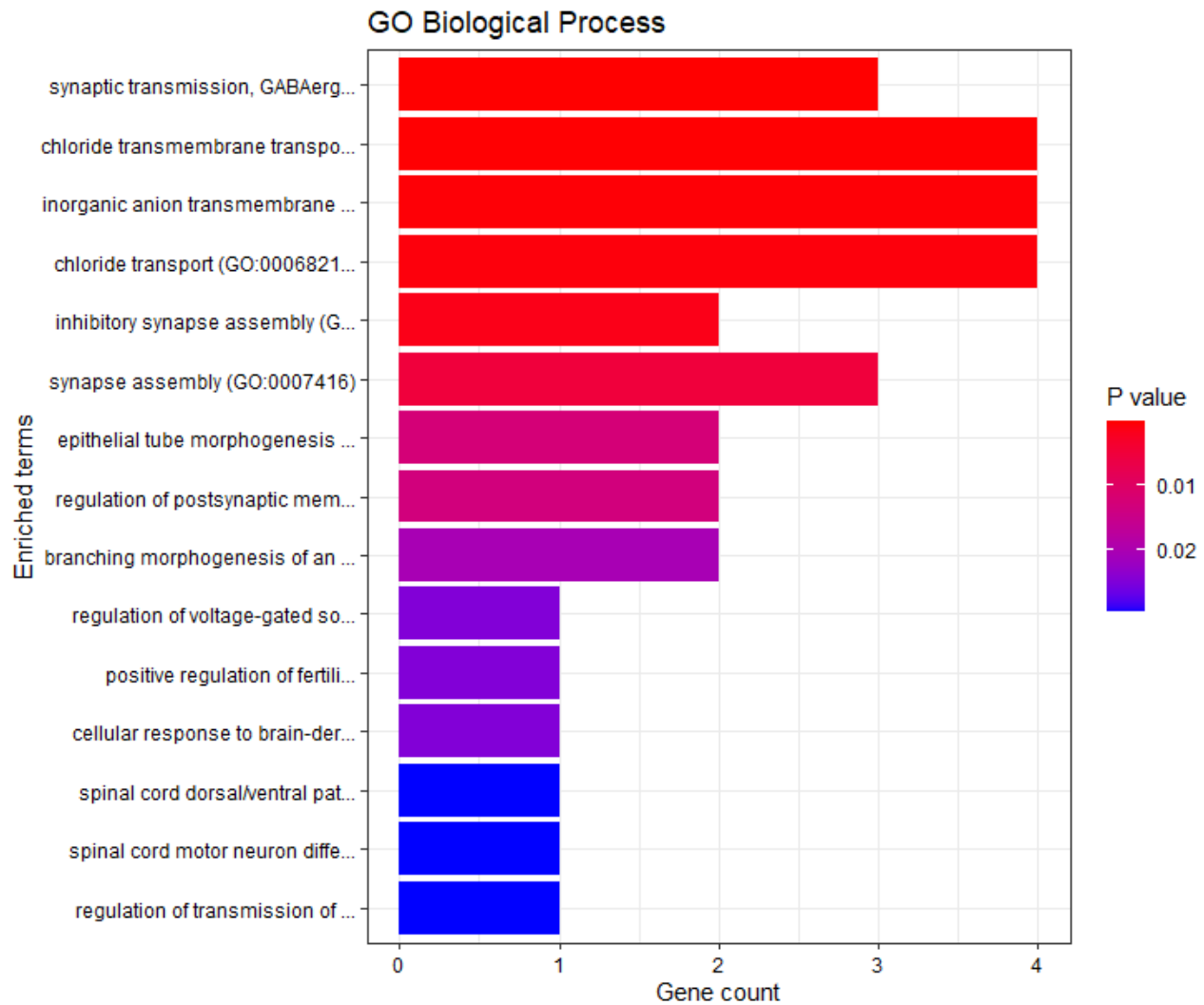


Figure 6A. GO biological process

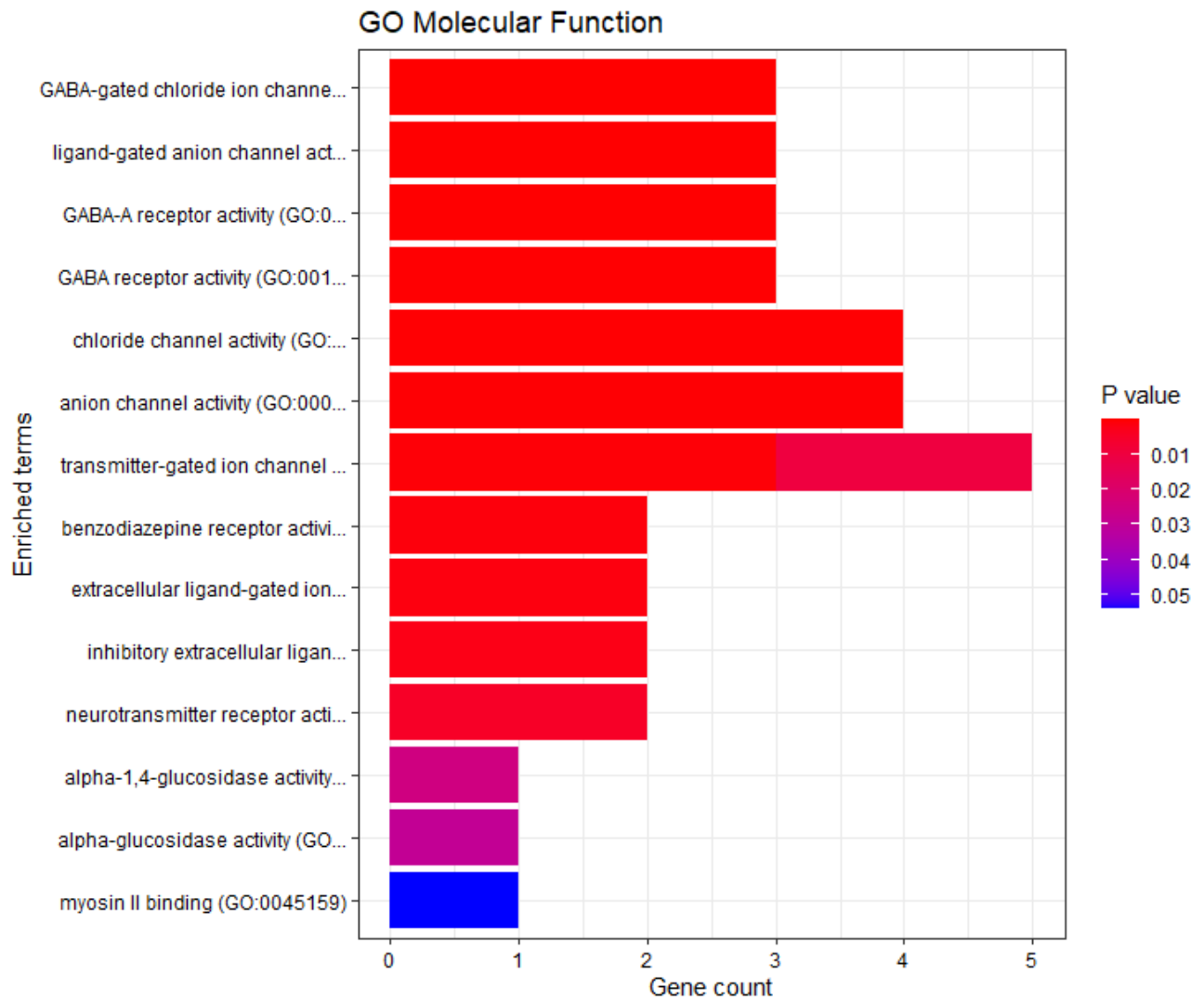


Figure 6B. GO molecular function

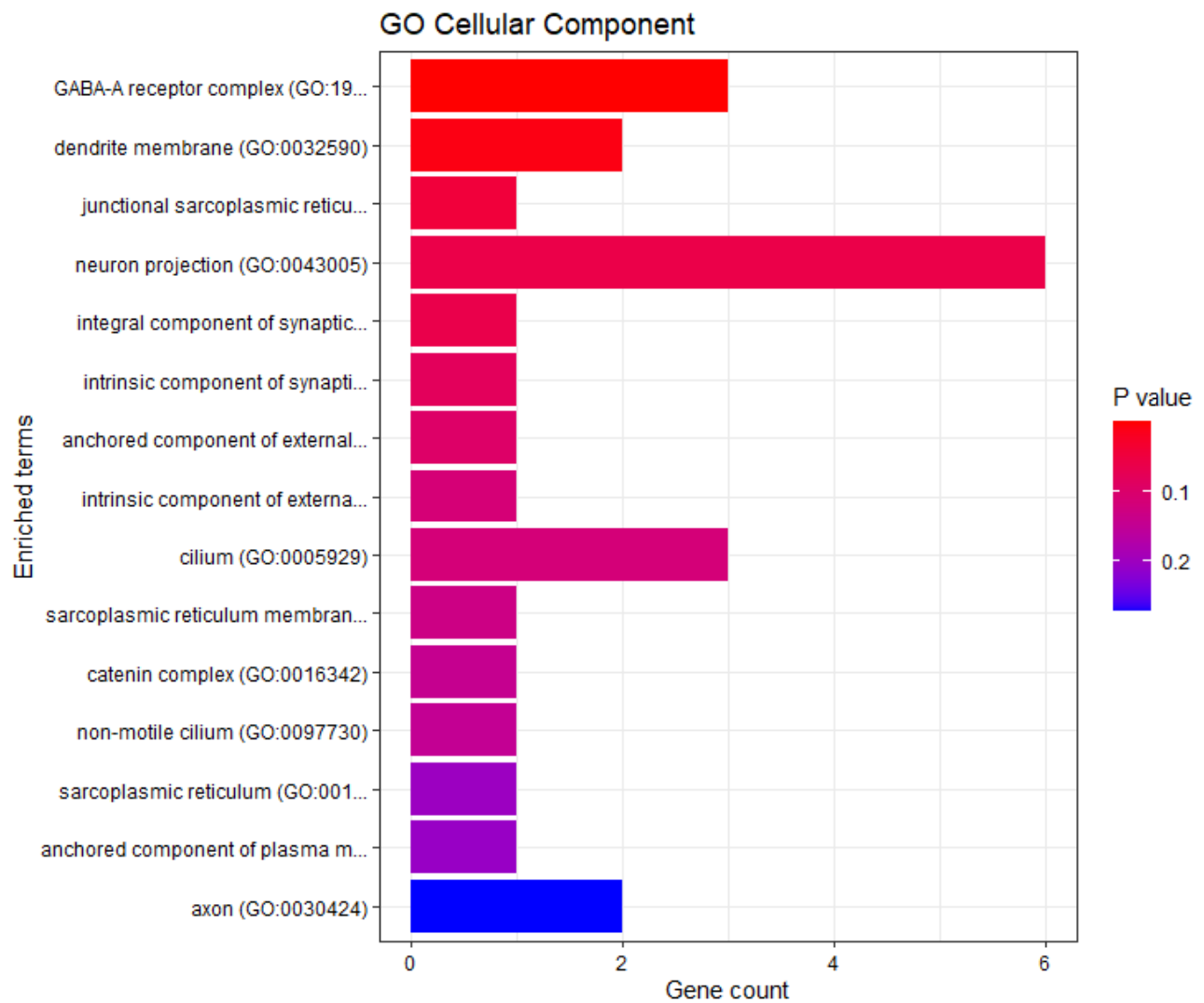


Figure 6C. GO cellular component

Nicotine addiction

Phenylalanine, tyrosine and tryptophan biosynthesis

Synthesis and degradation of ketone bodies

Caffeine metabolism

Maturity onset diabetes of the young

Neomycin, kanamycin and gentamicin biosynthesis

Taurine and hypotaurine metabolism

Vitamin B6 metabolism

Riboflavin metabolism

Fat digestion and absorption

Figure 6D. KEGG pathways

*ENRICH Gene Ontology and KEGG pathways analysis of significantly under-expressed genes of the non-African population.

Significantly enriched KEGG pathways

Nicotine addiction, caffeine metabolism, prion disease, thermogenesis, Huntington disease, drug metabolism, T cell receptor signaling among other pathways were significantly enriched in the downregulated genes (Figure 6E).

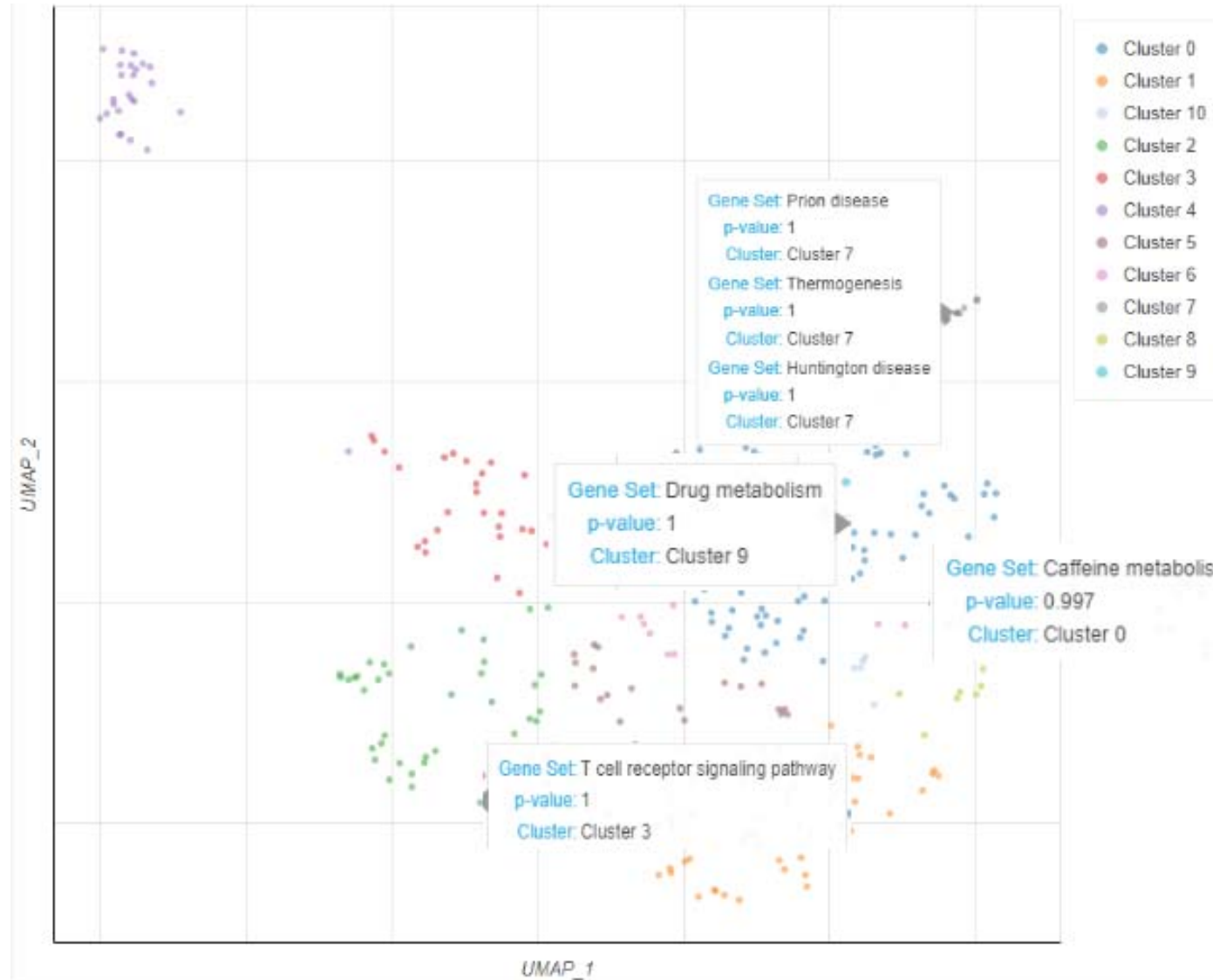


Figure 6E. Scatterplot of all terms in the KEGG_2021_Human gene set library. Each point represents a term in the library.

Discussion

The study set out to assess the difference in ACE2 expression in COVID-19 infected patients of African-American and non-African-American descent. Preliminary analysis between healthy African-American and COVID-19 infected African-American individuals did not show any difference in ACE2 expression. Likewise, a similar comparison between healthy and COVID-19 infected non-African-American individuals did not show any difference in ACE2 expression level. Our primary finding showed that the ACE2 gene was significantly downregulated in the African-American cohort compared to non-African-American cohort. The findings from this study might explain Yang and Cadejian's [15], [53] observation of reduced COVID-19 severity with reduced ACE2 expression.

There are reports that COVID-19 has a higher infection rate and severity among the African-American population [34], [52]. Besides the probable genetic cause of the disproportionate COVID-19 rates and deaths in African-Americans, other social factors such as overcrowding, working conditions, access to healthcare and background chronic health conditions may have played a greater role.

We analyzed gene expression in the two populations and discovered that African-American patients had 7,718 differentially upregulated genes and 16,374 differentially downregulated genes when compared to non-African-American patients. Many of these genes may be important genes for COVID-19 research. Our study confirms that there are significant disparities in differentially expressed genes (DEG) expression levels between African-American and non-African-American COVID-19 patients.

Residue function and enrichment analysis of downregulated genes suggests that nicotine addiction, caffeine metabolism, thermogenesis, drug metabolism, T cell receptor signaling, prion and Huntington disease pathways are actively involved in viral disease manifestation of African-American COVID-19 population as compared to a non-African-American population. Normally residues of the former two pathways prevent disease severity by blocking ACE2 from binding to COVID-19 spike protein, decreasing viral virulence or replication and inflammation [29], [42]. Therefore, the suppression of their corresponding genes in the diseased population clearly supports their significance in close protection against COVID-19 infection. The abundant presence of xanthine and caffeine/nicotine which are directly involved in the caffeine/nicotine metabolism pathway also directly reduce viral propagation in African-Americans by inhibiting red blood cells-ACE2 complex that allows viral entry [43], [57]. Autoimmune genes associated with prion and Huntington disease pathways regulate subtle peripheral and Central Nervous System (CNS) inflammation caused by pro-inflammatory cytokines such as Interleukins (IL) and Tumor Necrosis Factor (TNF) – α hence slowing down neural degeneration [22]. Hypothetically, the viral attack of the patient's nervous system by COVID-19 [55] might have significantly impacted these pathways resulting in downregulation of respective genes.

Thermogenic regulation is highly dependent on expression of ACE2 whereby its downregulation enhances microcirculation [47], thus subsequent downregulation of the thermogenic pathway in absence of any impairment. Cytochrome enzymes are active agents in toxin and drug degradation

and their suppression by infection leads to compromised drug metabolism giving rise to treatment complications especially in COVID-19 patients [11], [53]. Even though genes for these enzymes are usually upregulated in African-Americans as compared to non-African-Americans [44] discrepancies observed here might be attributed to socio-environmental factors beyond our control during sampling. Negative regulation of other pathways such as T cell receptor signaling reduces viral cell entry as well as T cell production controlling immunopathology [56]. Evidently, the respiratory, inflammatory and infectious disease pathways play a role in COVID-19 progression.

Moreover, the expression levels of TGFB1, VASP and CYBA in African-American COVID-19 patients were all elevated in our study among other genes. TGFB1 and CYBA are associated with the osteoclast differentiation pathway; TGFB1 is involved in regulatory systems that support the balance of osteoblasts and osteoclasts [45] while over-expression of CYBA is linked to decreased blood stem cells [45] Moreover, the expression of TGFB1 in Tuberculosis disease is known to be elevated as it impacts the patients' immunological response to the disease [36].

Meanwhile, the expression of VASP and MAPK1 have been reported to be associated with phagocytosis [27] and Fc Gamma R-mediated phagocytosis [24]. However, more study should be performed to confirm the function and character of VASP in with Fc Gamma R-mediated phagocytosis. Moreover, an upregulated gene linked to APOL1-associated kidney disease among African-Americans [48], [49], NLRP3 was found to be associated with salmonella infection in our study and its expression levels enhanced in our COVID-19 African- American patient transcriptomes. Our TGFB1, CYBA, and VASP results are consistent with those previously reported [24], so we speculate that these may be potential COVID-19 - related genes and more research is required.

Conclusions and next steps

The expression of ACE2 gene is downregulated in the cohort of African- American COVID-19 patients in this study. This may explain the low severity of COVID-19 in this population compared to non-African-American populations. Future exploration of the ACE2 gene and the pathway residues of respiratory, inflammatory and infectious disease pathways might offer new chemical therapeutics for this virus.

Declarations

Availability and Requirements

Project name: Expression Level Analysis of ACE2 Receptor Gene in African-American and Non-African-American COVID-19 Patients

Project home page: https://github.com/omicscodeathon/ace2_covid

Availability of data and materials

The dataset analyzed during the current study as a case study is publicly available at https://github.com/omicscodeathon/ace2_covid/blob/main/accessions. The data supporting the results reported in this manuscript is included within the article. The Project repository which also includes the entire code and other requirements can be downloaded from https://github.com/omicscodeathon/ace2_covid. The guidelines for this project and related updates, https://github.com/omicscodeathon/ace2_covid/blob/main/README.md are available at:

Abbreviations

ACE2 - angiotensin-Converting Enzyme 2

SARS-CoV-2 - severe acute respiratory syndrome coronavirus 2

SRA- sequence read archive

Ang-I - angiotensin I

TGFB1- transforming growth factor beta-1

APOL1- apolipoprotein L1

NLRP3 - NOD-, LRR- and pyrin domain-containing protein 3

STING1- Stimulator Of Interferon Response CGAMP Interactor 1

Acknowledgements

The authors thank the National Institutes of Health (NIH) Office of Data Science Strategy (ODSS) and the National Center for Biotechnology Information (NCBI) for their immense support before and during the April 2022 Omics codeathon organized in collaboration with the African Society for Bioinformatics and Computational Biology (ASBCB). The authors acknowledge the assistance of Danny Lumian for editing the manuscript.

Funding

This research was supported by the Intramural Research Program of the NIH, Office of Data Science Strategy. No grants were involved in supporting this work.

Author Information

Affiliations

Institution

Department of Bioinformatics, School of Pure and Applied Sciences, Pwani University, Kilifi, Kenya

Marion Nyaboke, Kauthar M. Omar

Biotechnology Unit, Institute of Child Health, University of Ibadan, Nigeria.

Ayorinde F. Fayehun

National School of Arts and Crafts, University of Hassan II Casablanca, Morocco.

Oumaima Dachi

Department of Biological Sciences, School of Pure and Applied Sciences, University of Embu, Kenya.

Billiah Bwana

Department of Computer Science, University of Ibadan, Ibadan, Oyo State, Nigeria

Olaitan I. Awe

African Society for Bioinformatics and Computational Biology, Cape Town, South Africa

Olaitan I. Awe

Author Name	ORCID ID
Marion N. Nyamari	0000-0003-0004-9982
Kauthar M. Omar	0000-0001-5812-3538
Ayorinde F. Fayehun	0000-0003-2302-8156
Oumaima Dachi	0000-0003-1192-2169

Billiah Bwana	0000-0001-8390-2453
Olaitan I. Awe	0000-0002-4257-3611

Contributions

MN conceived the original idea and developed the pipeline, performed the bioinformatic analysis of the data, and wrote the original draft of the manuscript. AF supervised, wrote, reviewed and edited intensively the manuscript and performed administrative roles. KO performed the bioinformatic analysis of the data, reviewed the manuscript and provided critical feedback and helped update the github repo. BK performed the bioinformatic analysis of the data, wrote, reviewed and edited the manuscript. OD performed the bioinformatic analysis of the data. OIA drafted the original abstract and provided the needed resources and guidance for the successful completion of the project. OIA reviewed the manuscript and provided critical feedback. All authors read and approved the final manuscript.

Ethics Declaration

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

No competing interests were disclosed.

References

- [1] Indari, O., Baral, B., Muduli, K., Mohanty, A. P., Swain, N., Mohakud, N. K., & Jha, H. C. (2021). Insights into Plasmodium and SARS-CoV-2 co-infection driven neurological manifestations. *Biosafety and Health*, 3(4), 230–234. <https://doi.org/10.1016/J.BSHEAL.2021.04.001>
- [2] World Health Organization. (2022). WHO coronavirus disease (COVID-19) dashboard. <https://covid19.who.int/>
- [3] Hussein, M. I. H., Albashir, A. A. D., Elawad, O., & Homeida, A. (2020). Malaria and COVID-19: unmasking their ties. *Malar J*, 19(1), 457. <https://doi.org/10.1186/s12936-020-03541-w>
- [4] Color of Coronavirus: COVID-19 deaths analyzed by race and ethnicity — APM Research Lab. (2022). Retrieved January 10, 2023, from <https://www.apmresearchlab.org/covid/deaths-by-race>
- [5] Gupta, R., Agrawal, R., Bukhari, Z., Jabbar, A., Wang, D., Diks, J., Alshal, M., Emechebe, D. Y., Brunicardi, F. C., Lazar, J. M., Chamberlain, R., Burza, A., & Haseeb, M. A. (2021). Higher comorbidities and early death in hospitalized African-American patients with Covid-19. *BMC Infectious Diseases*, 21(1), 1–11. <https://doi.org/10.1186/S12879-021-05782-9/FIGURES/3>
- [6] DiMaggio, C., Klein, M., Berry, C., & Frangos, S. (2020). Black/African American Communities are at highest risk of COVID-19: spatial modeling of New York City ZIP Code-level testing results. *Annals of Epidemiology*, 51, 7–13. <https://doi.org/10.1016/J.ANNEPIDEM.2020.08.012>
- [7] Reyes, M. V. (2020). The Disproportional Impact of COVID-19 on African Americans. *Health and Human Rights*, 22(2), 299. [/pmc/articles/PMC7762908/](https://pubmed.ncbi.nlm.nih.gov/37762908/)
- [8] worldometer. (2022). COVID-19 coronavirus pandemic. Retrieved 02/04 from <https://www.worldometers.info/coronavirus/?referer=app>
- [9] Wesonga, R.M. and Awe, O.I. (2022). An Assessment of Traditional and Genomic Screening in Newborns and their Applicability for Africa. *Informatics in Medicine Unlocked*, 32:101050. <https://doi.org/10.1016/j.imu.2022.101050>
- [10] Chikwambi, Z., Hidjo, M., Chikondowa, P., Afolabi, L., Aketch, V., Jayeoba, G., Awe, O.I. and Enoma, D.O. (2023). Multi-omics data integration approach identifies potential biomarkers for Prostate cancer. *bioRxiv*.
- [11] Deb, S., & Arrighi, S. (2021). Potential Effects of COVID 19 on Cytochrome P450 Mediated Drug Metabolism and Disposition in Infected Patients. *European Journal of Drug Metabolism and Pharmacokinetics*, 46(2), 185–203. <https://doi.org/10.1007/s13318-020->

00668-8.

[12] Zeberg, H., & Pääbo, S. (2020). The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature*, 587(7835), 610-612.

[13] Devaux, C. A., Rolain, J. M., & Raoult, D. (2020). ACE2 receptor polymorphism: Susceptibility to SARS-CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome. *J Microbiol Immunol Infect*, 53(3), 425-435. <https://doi.org/10.1016/j.jmii.2020.04.015>

[14] Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013 Apr 15;14:128. doi: 10.1186/1471-2105-14-128. PMID: 23586463; PMCID: PMC3637064.

[15] Cadegiani, F. A., Goren, A., & Wambier, C. G. (2020). Spironolactone may provide protection from SARS-CoV-2: Targeting androgens, angiotensin converting enzyme 2 (ACE2), and renin-angiotensin-aldosterone system (RAAS). *Medical Hypotheses*, 143, 110112.

[16] Donoghue, M., Hsieh, F., Baronas, E., Godbout, K., Gosselin, M., Stagliano, N., Donovan, M., Woolf, B., Robison, K., & Jeyaseelan, R. (2000). A novel angiotensin-converting enzyme–related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circulation research*, 87(5), e1-e9.

[17] Gomez, J., Albaiceta, G. M., Garcia-Clemente, M., Lopez-Larrea, C., Amado-Rodriguez, L., Lopez-Alonso, I., Hermida, T., Enriquez, A. I., Herrero, P., Melon, S., Alvarez-Arguelles, M. E., Boga, J. A., Rojo-Alba, S., Cuesta-Llavona, E., Alvarez, V., Lorca, R., & Coto, E. (2020). Angiotensin-converting enzymes (ACE, ACE2) gene variants and COVID-19 outcome. *Gene*, 762, 145102. <https://doi.org/10.1016/j.gene.2020.145102>

[18] Awe, O.I., En najih, N., Nyamari, M.N., and Mukanga, L.B. (2022). Comparative Study between Molecular and Genetic Evolutionary Analysis Tools Using African SARS-CoV-2 Variants. *Informatics in Medicine Unlocked*. <https://doi.org/10.1016/j.imu.2022.101143>

[19] Bakhshandeh, B., Sorboni, S. G., Javanmard, A. R., Mottaghi, S. S., Mehrabi, M. R., Sorouri, F., Abbasi, A., & Jahanafrooz, Z. (2021). Variants in ACE2; potential influences on virus infection and COVID-19 severity. *Infect Genet Evol*, 90, 104773. <https://doi.org/10.1016/j.meegid.2021.104773>

[20] Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N.-H., & Nitsche, A. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*, 181(2), 271-280. E278.

- [21] Hofmann, H., Geier, M., Marzi, A., Krumbiegel, M., Peipp, M., Fey, G. H., Gramberg, T., & Pöhlmann, S. (2004). Susceptibility to SARS coronavirus S protein-driven infection correlates with expression of angiotensin converting enzyme 2 and infection can be blocked by soluble receptor. *Biochemical and biophysical research communications*, 319(4), 1216-1221.
- [22] Hu, C., Chen, C., & Dong, X. (2021). Impact of COVID-19 Pandemic on Patients With Neurodegenerative Diseases. *Frontiers. Aging Neuroscience*. 13:(664965), 1–6. <https://doi.org/10.3389/fnagi.2021.664965>.
- [23] Hussain, M., Jabeen, N., Raza, F., Shabbir, S., Baig, A. A., Amanullah, A., & Aziz, B. (2020). Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. *J Med Virol*, 92(9), 1580-1586. <https://doi.org/10.1002/jmv.25832>
- [24] Awasthi, A., Rolland, D. C. M., Ayello, J., van de Ven, C., Basrur, V., Conlon, K., Fermin, D., Barth, M. J., Klein, C., Elenitoba-Johnson, K. S. J., Lim, M. S., & Cairo, M. S. (2017). A comparative global phosphoproteomics analysis of obinutuzumab (GA101) versus rituximab (RTX) against RTX sensitive and resistant Burkitt lymphoma (BL) demonstrates differential phosphorylation of signaling pathway proteins after treatment. *Oncotarget*, 8(69), 113895–113909. <https://doi.org/10.18632/ONCOTARGET.23040>
- [25] Aguiar, J. A., Tremblay, B. J., Mansfield, M. J., Woody, O., Lobb, B., Banerjee, A., Chandiramohan, A., Tiessen, N., Cao, Q., & Dvorkin-Gheva, A. (2020). Gene expression and in situ protein profiling of candidate SARS-CoV-2 receptors in human airway epithelial cells and lung tissue. *European Respiratory Journal*, 56(3).
- [26] Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 2015 12:4, 12(4), 357–360. <https://doi.org/10.1038/nmeth.3317>
- [27] Körber, S., & Faix, J. (2022). VASP boosts protrusive activity of macroendocytic cups and drives phagosome rocketing after internalization. *European Journal of Cell Biology*, 101(2). <https://doi.org/10.1016/J.EJCB.2022.151200>
- [28] Kuba, K., Imai, Y., Rao, S., Gao, H., Guo, F., Guan, B., Huan, Y., Yang, P., Zhang, Y., Deng, W., Bao, L., Zhang, B., Liu, G., Wang, Z., Chappell, M., Liu, Y., Zheng, D., Leibbrandt, A., Wada, T., . . . Penninger, J. M. (2005). A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat Med*, 11(8), 875-879. <https://doi.org/10.1038/nm1267>
- [29] Kumar, S. C., Kumar, S. A., & Wei, H. (2020). A computational insight of the improved nicotine binding with ACE2- SARS-CoV-2 complex with its clinical impact. 1–11.
- [30] Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G.,

Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/BIOINFORMATICS/BTP352>

[31] Li, W., Zhang, C., Sui, J., Kuhn, J. H., Moore, M. J., Luo, S., Wong, S. K., Huang, I. C., Xu, K., & Vasilieva, N. (2005). Receptor and viral determinants of SARS coronavirus adaptation to human ACE2. *The EMBO journal*, 24(8), 1634-1643.

[32] Liu, D., Chen, Y., Zhang, P., Zhong, J., Jin, L., Zhang, C., Lin, S., Wu, S., & Yu, H. (2016). Association between circulating levels of ACE2-Ang-(1-7)-MAS axis and ACE2 gene polymorphisms in hypertensive patients. *Medicine (Baltimore)*, 95(24), e3876. <https://doi.org/10.1097/md.0000000000003876>

[33] Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17(1), 10–12. <https://doi.org/10.14806/EJ.17.1.200>

[34] Muñoz-Price, L. S., Nattinger, A. B., Rivera, F., Hanson, R., Gmehlin, C. G., Perez, A., Singh, S., Buchan, B. W., Ledebor, N. A., & Pezzin, L. E. (2020). Racial Disparities in Incidence and Outcomes Among Patients With COVID-19. *JAMA Network Open*, 3(9), e2021892-e2021892. <https://doi.org/10.1001/jamanetworkopen.2020.21892>

[35] Novelli, A., Biancolella, M., Borgiani, P., Cocciadiferro, D., Colona, V. L., D'Apice, M. R., Rogliani, P., Zaffina, S., Leonardis, F., Campana, A., Raponi, M., Andreoni, M., Grelli, S., & Novelli, G. (2020). Analysis of ACE2 genetic variants in 131 Italian SARS-CoV-2-positive patients. *Hum Genomics*, 14(1), 29. <https://doi.org/10.1186/s40246-020-00279-z>

[36] Peresi, E., Oliveira, L. R. C., da Silva, W. L., da Costa, É. A. P. N., Araujo, J. P., Ayres, J. A., Fortes, M. R. P., Graviss, E. A., Pereira, A. C., & Calvi, S. A. (2013). Cytokine Polymorphisms, Their Influence and Levels in Brazilian Patients with Pulmonary Tuberculosis during Antituberculosis Treatment. *Tuberculosis Research and Treatment*, 2013, 1–13. <https://doi.org/10.1155/2013/285094>

[37] Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E., & Zanini, F. (2021). Analysing high-throughput sequencing data in Python with HTSeq 2.0. <http://arxiv.org/abs/2112.00939>

[38] Cambien, F. (1994). The angiotensin-converting enzyme (ACE) genetic polymorphism: its relationship with plasma ACE level and myocardial infarction. *Clinical genetics*, 46(1), 94-101.

[39] Rice, G. I., Thomas, D. A., Grant, P. J., Turner, A. J., & Hooper, N. M. (2004). Evaluation of angiotensin-converting enzyme (ACE), its homologue ACE2 and neprilysin in angiotensin peptide metabolism. *Biochemical Journal*, 383(1), 45-51.

[40] Rigat, B., Hubert, C., Alhenc-Gelas, F., Cambien, F., Corvol, P., & Soubrier, F. (1990). An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene

accounting for half the variance of serum enzyme levels. *The Journal of clinical investigation*, 86(4), 1343-1346.

[41] Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139–140. <https://doi.org/10.1093/BIOINFORMATICS/BTP616>

[42] Romero-Martínez, B. S., Montaña, L. M., Sommer, B., Lizbeth, G., Gloria, E. P., & Flores-soto, E. (2021). Possible Beneficial Actions of Caffeine in SARS-CoV-2. *International Journal. Molecular. Science*, 22, 5460, 1–12. <https://doi.org/10.3390/ijms22115460>

[43] Singh H, Choudhari R, Nema V, Khan AA. ACE2 and TMPRSS2 polymorphisms in various diseases with special reference to its impact on COVID-19 disease. *Microb Pathog*. 2021 Jan;150:104621. doi: 10.1016/j.micpath.2020.104621. Epub 2020 Dec 2. PMID: 33278516; PMCID: PMC7709597.

[44] Singh, U., Hernandez, K.M., Aronow, B.J., & Wurtele, E.S. (2021). African Americans and European Americans exhibit distinct gene expression patterns across tissues and tumors associated with immunologic functions and environmental exposures. *Scientific reports*, 11(9905), 1-13. <https://doi.org/10.1038/s41598-021-89224-1>.

[45] Chaudhari, S. D., Sharma, K. K., Marchetto, J. J., Hydren, J. R., Burton, B. M., & Moreno, A. P. (2021). Modulating OPG and TGF- β 1 mRNA expression via bioelectrical stimulation. *Bone Reports*, 15. <https://doi.org/10.1016/J.BONR.2021.101141>

[46] Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

[47] Cao, X., Shi, T., Zhang, C., Jin, W., Song, L., Liu, J., Yang, F., Rotimi, C. N., Xu, A., Hospital, B. T., Health, H., Engineering, N., Sciences, C., Health, G., Genome, N. H., Kong, H., Hospital, B. T., & Medical, C. (2021). ACE2 Pathway Regulates Thermogenesis and Energy 2 Metabolism. *bioRxiv*, 1-37. <https://doi.org/10.1101/2021.08.10.455823>.

[48] Wu, J., Ma, Z., Raman, A., Beckerman, P., Dhillon, P., Mukhi, D., Palmer, M., Chen, H. C., Cohen, C. R., Dunn, T., Reilly, J., Meyer, N., Shashaty, M., Arany, Z., Haskó, G., Laudanski, K., Hung, A., & Susztak, K. (2021). APOL1 risk variants in individuals of African genetic ancestry drive endothelial cell defects that exacerbate sepsis. *Immunity*, 54(11), 2632-2649.e2636. <https://doi.org/10.1016/j.immuni.2021.10.004>

[49] Wu, J., Raman, A., Coffey, N. J., Sheng, X., Wahba, J., Seasock, M. J., Ma, Z., Beckerman, P., Laczkó, D., Palmer, M. B., Kopp, J. B., Kuo, J. J., Pullen, S. S., Boustany-Kari, C. M., Linkermann, A., & Susztak, K. (2021). The key role of NLRP3 and STING in APOL1-associated podocytopathy. *The Journal of Clinical Investigation*, 131(20). <https://doi.org/10.1172/JCI136329>

- [50] Xia, H., Sriramula, S., Chhabra, K. H., & Lazartigues, E. (2013). Brain angiotensin-converting enzyme type 2 shedding contributes to the development of neurogenic hypertension. *Circulation research*, 113(9), 1087-1096.
- [51] Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., & Zhou, Q. (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, 367(6485), 1444-1448. <https://doi.org/10.1126/science.abb2762>
- [52] Yancy, C. W. (2020). COVID-19 and African Americans. *Jama*, 323(19), 1891-1892. <https://doi.org/10.1001/jama.2020.6548>
- [53] Yang, J., Chen, Z., & Jia, B. (2020). Higher COVID-19 risk for patients with cancer. *MedComm*, 1(3), 418.
- [54] Delanghe, J. R., Speeckaert, M. M., & De Buyzere, M. L. (2020). The host's angiotensin-converting enzyme polymorphism may explain epidemiological findings in COVID-19 infections. *Clinica chimica acta; international journal of clinical chemistry*, 505, 192.
- [55] Guerrero, J. I., Barragán, L. A., Martínez, J. D., Montoya, J. P., Peña, A., Sobrino, F. E., Tovar-spinoza, Z., & Ghotme, K. A. (2021). Central and peripheral nervous system involvement by COVID-19: a systematic review of the pathophysiology, clinical manifestations, neuropathology, cerebrospinal fluid findings. *BMC Infectious Diseases*, 21(515), 1–15. <https://doi.org/10.1186/s12879-021-06185-6>.
- [56] Diao, B., Wang, C., Tan, Y., Chen, X., Liu, Y., Ning, L., Chen, L., Li, M., Liu, Y., Wang, G., Yuan, Z., Feng, Z., Zhang, Y., Wu, Y., & Chen, Y. (2020). Reduction and Functional Exhaustion of T Cells in Patients With Coronavirus Disease 2019 (COVID-19). *Frontiers in Immunology*, 11. <https://doi.org/10.3389/fimmu.2020.00827>
- [57] Valdés, A., Moreno, L. O., Rello, S. R., Orduña, A., Bernardo, D., & Cifuentes, A. (2022). Metabolomics study of COVID-19 patients in four different clinical stages. *Scientific Reports*, 12(1), 1-11. <https://doi.org/10.1038/s41598-022-05667-0>
- [58] Nzungize, L., Kengne-Ouafo, J.A., Wesonga, M.R., Umuhoza, D., Murithi, K., Kimani, P., Awe, O.I. and Dillman, A. (2022). Transcriptional Profiles Analysis of COVID-19 and Malaria Patients Reveals Potential Biomarkers in Children. *BioRxiv*.
- [59] Mwanga, M.J., Obura, H.O., Evans, M. and Awe, O.I. (2023). Enhanced Deep Convolutional Neural Network for SARS-CoV-2 Variants Classification. *BioRxiv*.
- [60] Obura, H.O., Mlay, C.D., Moyo, L., Karumbo, B.M., Omar, K.M., Sinza, E.M., Rotich, G.J., Mudaki, W., Kamau, B.M., and Awe, O.I. (2022). Molecular Phylogenetics of HIV-1 Subtypes in African Populations: A Case Study of Sub-Saharan African Countries. *BioRxiv*.

[61] Ather, S.H., Awe, O.I., Butler, T.J., Denka, T., Semick, S.A., Tang, W. and Busby B. (2018). SeqAcademy: an educational pipeline for RNA-Seq and ChIP-Seq analysis. F1000Research.

[62] Oluwagbemi, O. and Awe, O.I. (2018). A comparative computational genomics of Ebola Virus Disease strains: In-silico Insight for Ebola control. Informatics in Medicine Unlocked.