

Genetic correlations between COVID-19 and a variety of traits and diseases

Xiao Chang^{1,#}, Yun Li^{2,#}, Kenny Nguyen¹, Huiqi Qu¹, Yichuan Liu¹, Joseph Glessner¹, Patrick M A Sleiman^{1,3,4}, Hakon Hakonarson^{1,3,4,†}

¹The Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA.

²Department of Biostatistics and Epidemiology, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, 19104, USA.

³Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Pennsylvania, 19104, USA.

⁴Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, 19104, USA.

†Corresponding Author: Hakon Hakonarson, MD, PhD Director, Center for Applied Genomics Children's Hospital of Philadelphia 3615 Civic Center Boulevard, Suite 1216 Philadelphia, PA 19104-4318, USA (Hakonarson@email.chop.edu).

#Contributed equally to this work.

Abstract

We analyzed GWAS results released by COVID-19 Host Genetics Initiative, UK biobank and GWAS Catalog to explore the genetic overlap between COVID-19 and a broad spectrum of traits and diseases. We validate previously reported medical conditions and risk factors based on epidemiological studies, including but not limited to hypertension, type 2 diabetes and obesity. We also report novel traits associated with COVID-19, which have not been previously reported from epidemiological data, such as opioid use and educational attainment. Taken together, this study extends our understanding of the genetic basis of COVID-19, and provides target traits for further epidemiological studies.

Introduction

The ongoing coronavirus disease 2019 (COVID-19) break has posed an extraordinary threat to global public health. Patients with certain underlying medical conditions such as obesity, hypertension, and diabetes are at increased risk for poor outcome in COVID-19¹. Given high genetic heritability of the aforementioned conditions, their shared genetic factors may play a crucial role in the severity of COVID-19. Indeed, a recent genome-wide association study (GWAS) of COVID-19 has reported two genomic loci associated with severe COVID-19, indicating a strong genetic influence on the severity of COVID-19². Here, we analyzed GWAS results released by COVID-19 Host Genetics Initiative³, UK biobank and GWAS Catalog to explore the genetic overlap between COVID-19 and a broad spectrum of traits and diseases.

Methods

Summary statistics of selected COVID-19 GWAS (sample size > 30,000 and percentage of Europeans > 90%) were downloaded from COVID-19 Host Genetics Initiative including A2_ALL (very severe respiratory confirmed COVID-19 against population), B2_ALL (hospitalized COVID-19 against population), B2_ALL_eur (hospitalized COVID-19 against population in Europeans), C2_ALL_eur (COVID-19 against population in Europeans). GWAS summary statistics of selected diseases/traits were downloaded from UK biobank and GWAS Catalog⁴. Genetic correlation r_g between COVID-19 and interested diseases/traits were estimated by LD score regression (LDSC) using GWAS summary statistics that overlap with HapMap3 SNPs as recommended⁵. Pre-computed linkage disequilibrium scores for HapMap3 SNPs calculated based on European-ancestry individuals from the 1000 Genomes Project were used in the analysis.

Results

We first investigated genetic correlations between COVID-19 and 1555 diseases/traits from the analysis of UK biobank data by Neale lab (<http://www.nealelab.is/uk-biobank/>). Our results are consistent with the epidemiological observation that BMI is significantly associated with severe or hospitalized COVID-19 (A2_ALL, $r_g = 0.24$, $P = 3.35 \times 10^{-6}$; B2_ALL, $r_g = 0.39$, $P = 3.33 \times 10^{-7}$). COPD (chronic obstructive pulmonary disease), heart diseases, hypertension, diabetes and smoking status exhibit substantial magnitude of genetic correlation with COVID-19, though statistical significance do not pass the strict threshold after adjustment for multiple testing (Table 1). Collectively, diseases of the circulatory system, diseases of the digestive system, and diseases of the musculoskeletal system and connective tissue are significantly associated with severe or hospitalized COVID-19 (Table 1). In agreement with this, a number of medication-taking traits linked to obesity, diabetes, hypertension and digestion display modest correlation with COVID-19 (Table 1). Tramadol, an opioid pain medication, is significantly correlated with hospitalized COVID-19 (B2_ALL, $r_g = 0.65$, $P = 1.26 \times 10^{-5}$). Interestingly, our results indicated a significant negative correlation between severe or hospitalized COVID-19 and educational attainment related traits including college or university degree (A2_ALL, $r_g = -0.24$, $P = 8.51 \times 10^{-7}$; B2_ALL, $r_g = -0.32$, $P = 1.78 \times 10^{-6}$) and fluid intelligence score (A2_ALL, $r_g = -0.25$, $P = 2.40 \times 10^{-5}$; B2_ALL, $r_g = -0.26$, $P = 7.66 \times 10^{-5}$). Hospitalized COVID-19 (B2_ALL, $r_g = 0.73$, $P = 8.60 \times 10^{-6}$) is also significantly correlated with panic attacks.

We next estimated genetic correlations between COVID-19 and 80 diseases/traits from GWAS Catalog ⁴. Consistently, hypertension, type 2 diabetes and obesity are significantly associated with severe or hospitalized COVID-19 (Table 2). Coronary artery disease, heart failure and BMI are also modestly associated with severe COVID-19. Likewise, medication-taking traits related to obesity, diabetes, hypertension and digestion are modestly associated with COVID-19 such as drugs for diabetes, and antihypertensives (Table 2). Significant correlations are also found between hospitalized COVID-19 and drugs for peptic ulcer and gastro-oesophageal reflux disease (B2_ALL, $rg = 0.34$, $P = 3.01 \times 10^{-5}$; B2_ALL_eur, $rg = 0.28$, $P = 0.0006$), diuretic use and very severe respiratory confirmed COVID-19 (A2_ALL, $rg = 0.25$, $P = 0.0002$), opioids and severe or hospitalized COVID-19 (A2_ALL, $rg = 0.30$, $P = 5.59 \times 10^{-5}$; B2_ALL, $rg = 0.44$, $P = 4.78 \times 10^{-7}$; B2_ALL_eur, $rg = 0.38$, $P = 0.0002$). In consistent with the findings from UK biobank data, a strong negative genetic correlation between educational attainment and severe or hospitalized COVID-19 (A2_ALL, $rg = -0.33$, $P = 1.90 \times 10^{-9}$; B2_ALL, $rg = -0.41$, $P = 1.82 \times 10^{-8}$; B2_ALL_eur, $rg = -0.35$, $P = 2.07 \times 10^{-6}$) is observed. Further analyses of brain function and personality traits show COVID-19 is significantly correlated with cognitive performance and verbal-numerical reasoning, but not memory performance, reaction time or neuroticism (Table 1). In addition, very severe respiratory confirmed COVID-19 (A2_ALL, $rg = -0.37$, $P = 0.0006$) is significantly correlated with systemic lupus erythematosus.

Discussion

Our genetic correlation results between COVID-19 and a variety of traits and diseases confirm medical conditions and risk factors reported from epidemiological studies such as hypertension, type 2 diabetes and obesity. However, as far as we are aware, the association between opioids and COVID-19 has not previously been reported using epidemiological data. As side effects associated with chronic opioid use at high doses may affect the immune system and increase the risk of pneumonia, there is an urgent need to evaluate the relationship between COVID-19 severity and opioid use by epidemiological studies ⁶. Also, patients using chronic opioids may be considered a vulnerable group for careful monitoring. In addition, our results suggest that immune pathways involved in systemic lupus erythematosus may also play an important role in the severity of COVID-19. Finally, the observed negative correlation between COVID-19 and educational attainment reflects an indirect link mediated by environment or human behavior. For example, patients with different educational levels may differ in diet choices (BMI) or smoking status. This study provides novel information on underlying conditions that might increase the risk of severe illness of COVID-19. Added epidemiological studies are warranted to further evaluate these findings.

Table 1. Genetic correlations between COVID-19 and a variety of diseases and other medically relevant traits

Source	Disease/Trait	Reference	A2_ALL		B2_ALL		B2_ALL_eur		C2_ALL_eur	
			rg(se)	<i>P</i>	rg(se)	<i>P</i>	rg(se)	<i>P</i>	rg(se)	<i>P</i>
UK biobank	Diseases of the circulatory system	IX_CIRCULATORY	0.34 (0.08)	2.20E-05*	0.38 (0.09)	4.33E-05	0.32 (0.10)	0.0012	0.22 (0.11)	0.042
	Diseases of the digestive system	XI_DIGESTIVE	0.25 (0.07)	0.0004	0.37 (0.08)	5.38E-06**	0.38 (0.10)	8.40E-05	0.08 (0.09)	0.342

Diseases of the musculoskeletal system and connective tissue	XIII_MUSCULOSKELET	0.33 (0.07)	1.38E-05*	0.35 (0.09)	0.0001	0.34 (0.10)	0.001	0.15 (0.10)	0.137
COPD, early/late onset	COPD_EARLYANDLATER	0.50 (0.15)	0.0011	0.55 (0.16)	0.0006	0.46 (0.17)	0.0081	0.33 (0.18)	0.069
COPD differential diagnosis	COPD_EXCL	0.16 (0.10)	0.1164	0.32 (0.11)	0.0041	0.31 (0.13)	0.0146	0.27 (0.14)	0.043
Major coronary heart disease event	I9_CHD	0.24 (0.09)	0.0088	0.26 (0.10)	0.0117	0.19 (0.11)	0.0786	0.07 (0.12)	0.534
Ischaemic heart disease, wide definition	I9_IHD	0.24 (0.08)	0.0014	0.30 (0.09)	0.0006	0.22 (0.09)	0.0155	0.14 (0.09)	0.144
Vascular/heart problems diagnosed by doctor: Heart attack	6150_1	0.33 (0.09)	0.0003	0.35 (0.11)	0.002	0.30 (0.12)	0.0126	0.19 (0.12)	0.111
Vascular/heart problems diagnosed by doctor: Angina	6150_2	0.28 (0.09)	0.0017	0.36 (0.11)	0.0014	0.26 (0.11)	0.0169	0.23 (0.11)	0.033
Vascular/heart problems diagnosed by doctor: High blood pressure	6150_4	0.17 (0.05)	0.0011	0.22 (0.07)	0.0016	0.19 (0.07)	0.0065	0.21 (0.08)	0.007
Non-cancer illness code, self-reported: stroke	20002_1081	0.53 (0.22)	0.0158	0.50 (0.22)	0.0206	0.47 (0.23)	0.0452	0.32 (0.26)	0.223
Non-cancer illness code, self-reported: diabetes	20002_1220	0.28 (0.08)	0.0003	0.36 (0.09)	7.96E-05	0.38 (0.10)	0.0003	0.24 (0.09)	0.009
Non-cancer illness code, self-reported: osteoarthritis	20002_1465	0.28 (0.10)	0.0042	0.27 (0.09)	0.0045	0.25 (0.10)	0.0164	0.02 (0.11)	0.833
Non-cancer illness code, self-reported: high cholesterol	20002_1473	0.19 (0.07)	0.0071	0.19 (0.08)	0.0172	0.19 (0.09)	0.0281	0.00 (0.08)	0.986
Smoking status: Never	20116_0	-0.06 (0.05)	0.2025	-0.22 (0.06)	0.0005	-0.22 (0.07)	0.0024	-0.06 (0.07)	0.418
Smoking status: Current	20116_2	0.20 (0.06)	0.0014	0.35 (0.08)	1.97E-05*	0.34 (0.09)	0.0002	0.13 (0.09)	0.146
BMI	21001_irt	0.24 (0.05)	3.35E-06**	0.39 (0.08)	3.33E-07***	0.32 (0.08)	3.89E-05	0.25 (0.07)	0.000
Age hay fever, rhinitis or eczema diagnosed	3761_irt	0.27 (0.08)	0.0005	0.27 (0.10)	0.0064	0.28 (0.10)	0.0058	0.29 (0.13)	0.024
Age asthma diagnosed	3786_irt	0.15 (0.09)	0.1012	0.12 (0.10)	0.2563	0.07 (0.12)	0.5406	-0.03 (0.13)	0.829
Fluid intelligence score	20016_irt	-0.25 (0.06)	2.40E-05*	-0.26 (0.07)	7.66E-05*	-0.25 (0.08)	0.001	-0.13 (0.08)	0.083
Qualifications: College or University degree	6138_1	-0.24 (0.05)	8.51E-07**	-0.32 (0.07)	1.78E-06**	-0.28 (0.07)	9.62E-05	-0.23 (0.07)	0.001
Mental health problems ever diagnosed by a professional: Panic attacks	20544_6	0.40 (0.12)	0.001	0.73 (0.16)	8.60E-06*	0.64 (0.17)	0.0001	0.38 (0.19)	0.041
Seen doctor (GP) for nerves, anxiety, tension or depression	2090	0.18 (0.06)	0.002	0.20 (0.07)	0.0022	0.18 (0.07)	0.0087	-0.03 (0.07)	0.676
Medication related adverse effects (Asthma/COPD)	PULM_MEDICATIO_COMORB	0.28 (0.10)	0.0059	0.32 (0.11)	0.0052	0.36 (0.12)	0.0022	0.15 (0.14)	0.289
Treatment/medication code: ramipril	20003_1140860806	0.31 (0.10)	0.0013	0.44 (0.12)	0.0004	0.34 (0.13)	0.0095	0.24 (0.13)	0.056
Treatment/medication code: simvastatin	20003_1140861958	0.26 (0.08)	0.001	0.26 (0.09)	0.0055	0.28 (0.10)	0.0046	0.07 (0.09)	0.432
Treatment/medication code: tramadol	20003_1140864992	0.52 (0.13)	7.61E-05	0.65 (0.15)	1.26E-05*	0.63 (0.16)	5.30E-05	0.49 (0.17)	0.003
Treatment/medication code: omeprazole	20003_1140865634	0.22 (0.09)	0.017	0.36 (0.10)	0.0004	0.30 (0.11)	0.0049	0.06 (0.13)	0.662
Treatment/medication code: aspirin	20003_1140868226	0.24 (0.09)	0.0067	0.32 (0.11)	0.0047	0.29 (0.12)	0.0162	0.17 (0.12)	0.157
Treatment/medication code: amitriptyline	20003_1140879616	0.38 (0.13)	0.0031	0.49 (0.16)	0.0016	0.44 (0.16)	0.0059	0.22 (0.16)	0.167
Treatment/medication code: metformin	20003_1140884600	0.23 (0.08)	0.0045	0.31 (0.10)	0.0015	0.33 (0.11)	0.0016	0.20 (0.10)	0.042
Treatment/medication code: co-codamol	20003_1140923346	0.40 (0.11)	0.0002	0.49 (0.13)	0.0001	0.32 (0.13)	0.0171	0.07 (0.15)	0.629
Treatment/medication code: paracetamol	20003_2038460150	0.26 (0.07)	0.0003	0.35 (0.09)	0.0002	0.27 (0.10)	0.0038	-0.02 (0.10)	0.814

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

GWAS Catalog	ADHD	30478444	-0.23 (0.07)	0.0017	-0.26 (0.09)	0.0021	-0.23 (0.09)	0.0079	-0.05 (0.10)	0.569
	Autism	30804558	-0.01 (0.09)	0.9165	0.08 (0.11)	0.4675	0.11 (0.11)	0.3164	0.08 (0.13)	0.546
	Depression	30718901	-0.15 (0.05)	0.0029	-0.16 (0.06)	0.0049	-0.13 (0.06)	0.0248	0.0219 (0.07)	0.740
	Schizophrenia	25056061	-0.07 (0.05)	0.1711	-0.06 (0.06)	0.3	-0.08 (0.06)	0.1469	-0.04 (0.07)	0.600
	Asthma	29785011	-0.16 (0.08)	0.037	-0.12 (0.09)	0.1763	-0.15 (0.10)	0.1386	-0.27 (0.12)	0.026
	Allergy	29785011	-0.20 (0.07)	0.0067	-0.26 (0.09)	0.0046	-0.19 (0.10)	0.0511	-0.32 (0.12)	0.007
	SLE	26502338	-0.37 (0.11)	0.0006*	-0.27 (0.13)	0.0489	-0.27 (0.15)	0.0736	-0.23 (0.17)	0.175
	Smoking (age of initiation)	30643251	-0.23 (0.07)	0.0008	-0.32 (0.08)	0.0001**	-0.24 (0.08)	0.0035	-0.15 (0.09)	0.099
	Smoking (initiation)	30643251	0.10 (0.05)	0.0274	0.19 (0.06)	0.0005*	0.18 (0.06)	0.0026	0.05 (0.07)	0.457
	Smoking (cessation)	30643251	0.27 (0.07)	0.0003*	0.42 (0.08)	3.60E-07***	0.39 (0.09)	2.06E-05**	0.09 (0.11)	0.428
	Smoking (cigarettes per day)	30643251	0.16 (0.06)	0.0048	0.26 (0.06)	3.88E-05**	0.17 (0.07)	0.0097	0.00 (0.08)	0.957
	Hypertension	30940143	0.21 (0.06)	0.0005*	0.27 (0.08)	0.0005*	0.18 (0.07)	0.0077	0.18 (0.08)	0.022
	Coronary artery disease	28714975	0.22 (0.07)	0.001	0.20 (0.07)	0.0077	0.20 (0.08)	0.0113	0.11 (0.09)	0.207
	Heart failure	31919418	0.20 (0.09)	0.0205	0.24 (0.09)	0.0103	0.22 (0.11)	0.0468	0.17 (0.12)	0.182
	Type 2 diabetes	25056061	0.22 (0.06)	0.0003*	0.31 (0.07)	3.75E-05**	0.29 (0.08)	0.0002*	0.20 (0.08)	0.010
	Obesity	30677029	0.35 (0.13)	0.0082	0.55 (0.15)	0.0003*	0.55 (0.17)	0.0008	0.44 (0.17)	0.009
	BMI	23563607	0.26 (0.10)	0.0079	0.41 (0.13)	0.0016	0.29 (0.12)	0.0132	0.22 (0.15)	0.147
	Drugs for peptic ulcer and gastro-oesophageal reflux disease	31015401	0.23 (0.07)	0.0007	0.34 (0.08)	3.01E-05**	0.28 (0.08)	0.0006*	0.05 (0.10)	0.649
	Drugs for diabetes	31015401	0.17 (0.07)	0.0214	0.24 (0.08)	0.004	0.26 (0.09)	0.0042	0.19 (0.10)	0.052
	Antithrombotic agents	31015401	0.24 (0.09)	0.0085	0.28 (0.12)	0.0191	0.28 (0.12)	0.0266	0.29 (0.14)	0.047
	Vasodilators for cardiac diseases	31015401	0.33 (0.12)	0.0068	0.48 (0.16)	0.0021	0.43 (0.17)	0.0095	0.51 (0.21)	0.015
	Antihypertensives	31015401	0.30 (0.12)	0.0163	0.41 (0.14)	0.0049	0.40 (0.16)	0.0098	0.22 (0.17)	0.197
	Diuretic use	31015401	0.25 (0.07)	0.0002*	0.26 (0.08)	0.0018	0.29 (0.09)	0.0011	0.25 (0.11)	0.019
	Beta blocking agents	31015401	0.20 (0.07)	0.0078	0.22 (0.09)	0.0147	0.17 (0.09)	0.0762	0.09 (0.11)	0.395
	Agents acting on the renin-angiotensin system	31015401	0.19 (0.06)	0.001	0.21 (0.08)	0.0062	0.22 (0.08)	0.0068	0.19 (0.09)	0.035
	HMG CoA reductase inhibitors	31015401	0.22 (0.07)	0.0012	0.20 (0.08)	0.0121	0.22 (0.09)	0.0112	0.08 (0.09)	0.378
	Opioids	31015401	0.30 (0.07)	5.59E-05**	0.44 (0.09)	4.78E-07***	0.38 (0.10)	0.0002*	0.26 (0.11)	0.022
	Subjective well-being	27089181	0.17 (0.09)	0.0682	0.09 (0.10)	0.3818	0.11 (0.11)	0.311	0.14 (0.14)	0.317
	Educational attainment (years of education)	27225129	-0.33 (0.06)	1.90E-09***	-0.41 (0.07)	1.82E-08***	-0.35 (0.07)	2.07E-06***	-0.24 (0.08)	0.002
	Neuroticism	27089181	0.00 (0.07)	0.9995	-0.01 (0.09)	0.9161	-0.04 (0.10)	0.6917	-0.07 (0.12)	0.565
Cognitive performance	30038396	-0.22 (0.05)	3.95E-06	-0.21 (0.06)	0.0001**	-0.20 (0.06)	0.0006*	-0.22 (0.07)	0.003	
Memory performance	27046643	0.07 (0.10)	0.5018	0.07 (0.11)	0.5433	0.15 (0.11)	0.199	0.13 (0.12)	0.267	
Reaction time	27046643	0.08 (0.08)	0.3163	0.03	0.6847	-0.03	0.7716	-0.05 (0.11)	0.645	

					(0.08)		(0.09)			
Verbal-numerical reasoning	27046643	-0.33 (0.09)	0.0002*	-0.33 (0.10)	0.0007	-0.35 (0.11)	0.0018	-0.30 (0.13)	0.016	

Genetic correlation estimates, standard errors and p-values for COVID-19 and selected diseases/traits. The “Reference” column lists the phenotype code for the GWAS from UK biobank, and PMID (PubMed ID) for the GWAS from GWAS Catalog. The *P*-values are uncorrected *P*-values. Bonferroni correction was used for correction of multiple testing by the number of tests (1555 for the analysis of data from UK biobank, and 80 for the analysis of data GWAS Catalog). *** adjusted *P*-value ≤ 0.001 , ** adjusted *P*-value ≤ 0.01 , * adjusted *P*-value ≤ 0.05 . BMI = body mass index; SLE = systemic lupus erythematosus; rg = genetic correlation estimates; se = standard error, A2_ALL = very severe respiratory confirmed COVID-19 vs population, B2_ALL = hospitalized COVID-19 vs population, B2_ALL_eur = hospitalized COVID-19 vs population (European samples only), C2_ALL_eur = COVID-19 vs population (European samples only).

References

1. Richardson S, Hirsch JS, Narasimhan M, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA*. 2020;323(20):2052-2059.
2. Severe Covid GG, Ellinghaus D, Degenhardt F, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N Engl J Med*. 2020;383(16):1522-1534.
3. Initiative C-HG. 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. *Eur J Hum Genet*. 2020;28(6):715-718.
4. Buniello A, MacArthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. 2019;47(D1):D1005-D1012.
5. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291-295.
6. Ataei M, Shirazi FM, Lamarine RJ, Nakhaee S, Mehrpour O. A double-edged sword of using opioids and COVID-19: a toxicological view. *Subst Abuse Treat Prev Policy*. 2020;15(1):91.