

1                    **Rare Variants in Inborn Errors of Immunity Genes**  
2                    **Associated with Covid-19 Severity**

3   **Panhong Liu<sup>1,2</sup>, Mingyan Fang<sup>2</sup>, Yuxue Luo<sup>2</sup>, Fang Zheng<sup>4</sup>, Yan Jin<sup>4</sup>, Fanjun**  
4   **Cheng<sup>5</sup>, Huanhuan Zhu<sup>2,#</sup>, Xin Jin<sup>2,3,#</sup>**

5

6   <sup>1</sup>College of Life Sciences, University of Chinese Academy of Sciences, Beijing, Beijing  
7   city, China.

8   <sup>2</sup>BGI-Shenzhen, Shenzhen, Guangdong, China.

9   <sup>3</sup>School of Medicine, South China University of Technology, Guangzhou, Guangdong,  
10   China.

11   <sup>4</sup>Department of Pediatrics, Union Hospital, Tongji Medical College, Huazhong  
12   University of Science and Technology, Wuhan, Hubei, China

13   <sup>5</sup>Department of Hematology, Union Hospital, Tongji Medical College, Huazhong  
14   University of Science and Technology, Wuhan, Hubei, China

15

16

17   #These authors are co-corresponding authors.

18   Correspondence: Xin Jin ([jinxin@genomics.cn](mailto:jinxin@genomics.cn)),

19                    Huanhuan Zhu ([zhuhuanhuan1@genomics.cn](mailto:zhuhuanhuan1@genomics.cn))

20

21

22   **Key Words:** Covid-19, inborn errors of immunity, gene-level tests, pathway-based  
23   analysis, PPI network analysis

24

25

## 26 **Abstract**

27 Covid-19 is a contagious disease caused by SARS-CoV-2, a novel severe acute  
28 respiratory syndrome coronavirus. Common variants and networks underlying host  
29 genetic mechanisms have been extensively studied to identify disease-associated  
30 genetic factors. However, there are few studies about the rare variants, typically inborn  
31 errors of immunity, in understanding the host genetics behind Covid-19 infection,  
32 especially in the Chinese population. To fill this gap, we investigate likely-deleterious  
33 missense and high-confidence predicted loss-of-function variants by (a) performing  
34 gene- and pathway-level association analyses, (b) examining known genes involved in  
35 type I interferon signaling and others previously reported in Covid-19 disease, and (c)  
36 identifying candidate genes with accumulating mutations and their potential protein-  
37 protein interactions with known genes. Based on our analyses, several putative genes  
38 and pathways are uncovered and worth further investigation, for example, genes  
39 *IL12RB1*, *TBK1*, and *TLR3*, and pathways Tuberculosis (hsa:05152), Primary  
40 Immunodeficiency (hsa:05340), and Influenza A (hsa:05164). These regions generally  
41 play an essential role in regulating antiviral innate immunity responses to foreign  
42 pathogens and in responding to many inflammatory diseases. We believe that to some  
43 extent, as an acute inflammatory disease, Covid-19 is also affected by these inborn  
44 errors of immunity. We hope that the identification of these rare genetic factors will  
45 provide new insights into the genetic architecture of Covid-19.

46

## 47 Introduction

48 Since the December of 2019, the coronavirus diseases 2019 (Covid-19) (Gorbalenya et  
49 al. 2020) caused by the SARS-CoV-2 virus (Severe Acute Respiratory Syndrome  
50 Coronavirus 2) (N. Zhu et al. 2020) has spread rapidly across the world. By January  
51 2022, the ongoing SARS-CoV-2 pandemic has caused more than three hundred and  
52 sixty million confirmed cases and more than five million deaths. Host genetic factors  
53 have been shown to play critical roles in the disease susceptibility and severity. The  
54 Covid-19 Host Genetics Initiative (Covid-19 HGI, <https://www.covid19hg.org/>) is an  
55 international initiative to share the results of host genome-wide associations study  
56 (GWAS) meta-analysis of Covid-19 disease. The most recent Covid-19 HGI release 6  
57 has reported 24 lead SNPs ( $P < 5e-8$ ) mapped to nearly 136 genes, such as *LZTFL1*,  
58 *ABO*, *OAS1*, *DPP9*, *IFNAR2* (Initiative and Ganna 2021). The estimated heritability of  
59 Covid-19 symptoms explained by these common variants was 6.5% (Pairo-Castineira  
60 et al. 2021). A twin study with participants from the TwinsUK cohort reported that 31%  
61 of phenotypic variance of predicted Covid-19 is due to host genetic factors (Williams  
62 et al. 2020). This leads to a large proportion of unexplained heritability (nearly 25%),  
63 commonly referred to as “missing heritability”. There is increasing evidence that rare  
64 variants also make a major contribution to missing heritability of many complex  
65 diseases and traits (Zuk et al. 2014; Hunt et al. 2013; Misawa et al. 2020).

66 Recently, the rare variants attracted researchers’ attention in elucidating the missing  
67 heritability of Covid-19 susceptibility and severity. For example, Zhang et al. found  
68 that the rare predicted loss-of-function (pLoF) variants in the IRF7- and TLR3-  
69 dependent type I interferon (IFN) pathway were enriched in patients who developed  
70 risky Covid-19 (Q. Zhang et al. 2020). Smieszek et al. reported that pLoF variant in  
71 gene *IFNAR2* (c.966C>A/p.Y322X) might play a role not only in clinical manifestation  
72 of Covid-19 but also in the response to vaccination (Smieszek et al. 2021). In addition,  
73 Mantovani et al. found that pLoF variants in *TLR7* (c.3094G>A/p.A1032T,  
74 c.901T>C/p.S301P) occurred in severely affected male patients and downregulated the  
75 expression of TLR7 pathway (Mantovani et al. 2021). As previously reported, the rare  
76 variants were more likely to be functional and tended to have stronger effects on  
77 complex diseases (Gorlov et al. 2011). The study of genetic effects of rare variants is  
78 necessary to elucidate the severity of Covid-19.

79 To explore the genetic contributions of rare variants in Covid-19 patients with  
80 inborn errors of immunity, we recruited and investigated nearly 500 hospitalized  
81 patients from Union hospital of Tongji Medical College of Huazhong University of  
82 Science and Technology (*abbr.* Union hospital) (H. Zhu et al. 2021). Based on patients’  
83 genomic data and clinical information, we carried out three major analyses to  
84 investigate the effects of host rare variants: (a) gene- and pathway-level tests of these  
85 rare variants between *severe* and *non-severe* patients; (b) examination of the  
86 significance of previously reported rare variants and genes in our dataset; and c) rare  
87 mutation accumulation analysis and protein-protein interaction (PPI) network analysis

88 in only *severe* patients. From these analyses, we (a) identified candidate functional  
89 pathways that are responsible for innate immune disorders and respiratory diseases,  
90 such as Tuberculosis (hsa:05152), Primary Immunodeficiency (hsa:05340), and  
91 Influenza A (hsa:05164); (b) successfully replicated two Covid-19 associated SNPs  
92 (rs780744847 and rs541048548) mapped on genes *TLR3* and *ICAM3*, respectively; and  
93 (c) suggested several candidate genes, including *IL12RB1*, *TBK1*, and *TLR3*, which  
94 might be involved in SARS-CoV-2 cell entry, host immune responses, and Covid-19  
95 disease severity.

96       Until now, literatures based on Chinese population have replicated and discovered  
97 some Covid-19-associated common variants (Wang et al. 2020; H. Zhu et al. 2021; P.  
98 Wu, Chen, et al. 2021; P. Wu, Ding, et al. 2021), but genetic background of rare variants  
99 is currently insufficiently understood in Chinese population. Our work is an effort to  
100 fill this gap. We hope that it will serve as useful scientific reference to assess the genetic  
101 mechanism of rare variants in Covid-19 and advance our understanding of disease  
102 etiology.

## 103 **Materials and Methods**

### 104 **Patient Recruitment and Quality Control**

105 All subjects in this study were collected from the Union hospital. We used PLINK 2.0  
106 (Chang et al. 2015) to infer sex of individuals from SNP genotypes and VerifyBamID  
107 (F. Zhang et al. 2020) to assess the level of DNA contamination. Individuals were  
108 excluded if their PLINK inferred sex was inconsistent with that of clinical recorded.  
109 We also removed individuals with estimated contamination rates greater than 0.05.  
110 After sample quality control, there were 451 unrelated individuals with 159 mild and  
111 moderate patients, and 292 severe and critical patients. The severity classification  
112 criteria were made by the National Health Commission of P.R. China (Z. Wu and  
113 McGoogan 2020). We reclassified the mild/moderate patients as *non-severe* patients  
114 and the severe/critical patients as *severe* patients.

### 115 **Genotype Calling**

116 The blood samples of some patients were collected at different time points during  
117 hospitalization. To increase the average depth of study, sequence fastq files of each  
118 patient were merged together to generate one GVCF file by BWA and Sentieon  
119 Genomics software (Freed et al. 2017). Joint variant calling was then performed on  
120 GVCF files of all participants using the Sentieon GVCFTyper algorithm. The resulting  
121 VCF file was used for subsequent genomic analyses. After the application of excessHet  
122 (<54.69) filter, Variant Quality Score Recalibration (VQSR) was completed by using  
123 the Genome Analysis Toolkit (GATK version 4.1.2) (DePristo et al. 2011). To improve  
124 the genotyping accuracy, we used the Beagle 4.0 software (Browning and Browning  
125 2016) to perform LD-based genotype refinement by taking genotype likelihoods as  
126 inputs. Low-quality variants with dosage imputation score  $R^2 < 0.3$  by Beagle4.0 were  
127 filtered out.

### 128 **Principal Component Analysis**

129 Principal component analysis (PCA) was performed using a subset of autosomal bi-  
130 allelic SNPs by applying PLINK 2.0 (Chang et al. 2015). Several restrictions were  
131 applied to select SNPs for PCA analysis, including keeping SNPs with minor allele  
132 frequency (MAF)  $\geq 5\%$ , Hardy–Weinberg Equilibrium  $P \geq 1e-6$ , and removing one of  
133 a pair of SNPs if the LD was greater than 0.5 (in a window of fifty SNPs with a stop of  
134 five SNPs).

### 135 **Functional Annotation**

136 We annotated rare variants (MAF < 0.5%) in our final call set by using the Ensembl  
137 Variant Effect Predictor (VEP, build 103, GRCh38) (McLaren et al. 2016) with default  
138 parameters. The databases for annotation included dbSNP (Sherry et al. 2001),  
139 gnomAD (Karczewski et al. 2020), and 1000 Genomes Project (Clarke et al. 2012). In  
140 addition, we used Combined Annotation Dependent Depletion (CADD) score to predict

141 missense variants that had potential effects on protein function. The CADD score was  
142 annotated by CADD plug-in (Kircher et al. 2014). Missense variants with CADD score >  
143 MSC (Mutation Significance Cut-off) score (95% confidence interval) (Itan et al. 2016)  
144 were predicted as likely-deleterious missense variants. We also used LOFTEE  
145 (Karczewski et al. 2020) plugin to identify high-confidence pLoF (HC-pLoF) for stop-  
146 gained, frameshift, and splice site disrupting variants. Finally, we focused on the likely-  
147 deleterious missense and HC-pLoF variants in the subsequent analyses.

## 148 **Rare Variants Analyses**

149 To investigate the cumulative effects of multiple rare variants, we performed gene-  
150 based association analysis using KGGSeq 1.0 (Li et al. 2017) with the sequence kernel  
151 association test (SKAT) (M. C. Wu et al. 2010), the Optimized SKAT (SKAT-O) (Lee,  
152 Wu, and Lin 2012), and Burden test. We further carried out pathway-based analysis by  
153 testing the Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets (Kanehisa  
154 and Goto 2000). The adjusted covariates included age, sex, and the top six principal  
155 components. The gene-based and pathway-based p-values were then multiple testing  
156 corrected by Benjamini-Hochberg method. We defined the suggestive significance  
157 threshold for gene-based association test as  $1e-6$  and for pathway-based association test  
158 as 0.05.

159 We also focused on 13 type I IFN genes (denoted as IFN-genes) that were found  
160 an enrichment in life-threatening Covid-19 study (Q. Zhang et al. 2020) and 136 genes  
161 located in 50kb of risk lead SNPs reported by the Covid-19 Host Genetics Initiative  
162 (release 6, denoted as HGI-genes) (Initiative and Ganna 2021). The mutation accuracy  
163 of variants in these 148 candidate genes (one overlap between 13 IFN-genes and 136  
164 HGI-genes) was manually checked by using Samtools 1.10 (Danecek et al. 2021).

165 Finally, we performed an analysis of rare variant accumulation in genes identified  
166 by two approaches. The first approach detected genes if there was one variant met the  
167 following two conditions: (a) the mutations occurred in only *severe* patients, and (b) the  
168 variant harbored no less than three effect allele counts. We denoted these genes as  
169 “individual variant-driven genes”. The second approach determined genes if (a) all  
170 mutations in the gene occurred in only *severe* patients, and (b) the total number of  
171 mutations in the gene is at least three. We denoted these genes as “all variant-driven  
172 genes”. We note that, genes identified by the two methods may have some overlaps.  
173 Each of the two gene sets was then used for network analysis of protein-protein  
174 interactions (PPI) with the above 148 known candidate genes. We used the STRING  
175 version 10.5 (Search Tool for the Retrieval of Interacting Genes/Proteins) (Szklarczyk  
176 et al. 2019) to build the PPI network. The minimum required interaction score to highest  
177 confidence was set to 0.900.

178

## 179 **Result**

### 180 **Participant Characteristics**

181 In this study, participants included 451 Covid-19 patients aged from 23 to 97 years old  
182 and all declared Han Chinese population. In [Table 1](#), we provided participant  
183 characteristics for *non-severe* and *severe* patients, respectively. A total of 159 (35.25%),  
184 292 (64.75%) patients were grouped as *non-severe* and *severe*, respectively. The same  
185 as previously reported (Cummings et al. 2020; Yang et al. 2020), older age (*severe*: an  
186 average of 64 y/o vs. *non-severe*: an average of 58 y/o, t-test  $p = 4.6e-05$ ) and male sex  
187 (*severe* 52.74% vs. *non-severe* 42.14%, Fisher's exact test  $p = 0.04$ ) were at a higher  
188 risk of developing severe symptoms.

### 189 **Data Quality**

190 After quality control, the dataset consisted of 22,107,585 and 680,522 variants from  
191 autosomes and X chromosome, respectively ([Figure 1](#)). Then we compared chip array  
192 sequencing results with genotype after LD-based refinement by Beagle 4.0 on 218  
193 individuals. The heterozygote concordance rate increased from an average of 94.4% to  
194 97.4%, and the improvement is more dramatic for samples with lower sequencing depth  
195 ([Figure 2A](#)). After filtering in variants by imputation score  $DR2 > 0.3$ , the final dataset  
196 for further analyses had a total of 22,532,360 variants, and the PCA on 575,888  
197 autosome SNPs detected no outlier samples ([Figure 2B](#)).

### 198 **Rare Variants Statistics**

199 After filtering by MAF, we obtained a total of 13,934,341 rare variants for VEP  
200 annotation. Among the resulting annotations, there were 88,790 missense variants and  
201 4,881 pLoF variation (including stop-gained, frameshift, and splice site disrupting  
202 variants). Damaging effects of these missense and pLoF variants were then predicted  
203 by CADD and LOFTEE plug-in, respectively. About 43.41% missense variants were  
204 predicted as likely-deleterious missense variants (38,548) and 85.68% pLoF variants  
205 were predicted as HC-pLoF variants (4,182). Thus, in total, 42,730 predicted likely  
206 damaging variants were applied for further analysis. For both likely-deleterious  
207 missense variants and HC-pLoF variants, we first tested the difference of their numbers  
208 between *non-severe* and *severe* patients and found no significant difference  
209 ([Supplementary Table S1](#)).

### 210 **Gene- and Pathway-level Analysis of Rare Variants**

211 The gene-level analysis of rare variants was performed between *severe* and *non-severe*  
212 patients via KGGSeq. We performed the gene-based tests for genes mapped by all rare  
213 variants, likely-deleterious missense variants, and HC-pLoF variants, respectively. The  
214 gene-based analyses did not identify putative genes that passed the significance  
215 threshold of  $1e-6$  ([Supplementary Figure S1A-C](#)).

216 Furthermore, we leveraged the biological knowledge that sets of genes acting



217 together in pathways. In total, we tested 307 KEGG pathways and detected  
218 Tuberculosis (hsa:05152,  $p_{\text{adjust}} = 0.036$ ) between *severe* and *non-severe* patients on  
219 likely-deleterious missense and HC-pLoF variants ([Supplementary Table S2](#)).  
220 Tuberculosis (TB) is an airborne infectious disease caused by *Mycobacterium*  
221 *tuberculosis* (Mtb). It first attacks the lungs, then other parts of the body through  
222 circulatory system. This transmission characteristics is very similar to that of Covid-19.  
223 As previously reported, Tuberculosis pathway was significant with acute respiratory  
224 distress syndrome and lung injury in mice and human (Sweeney et al. 2017). The  
225 TB/Covid-19 Global Study Group observed a phenomenon that TB and SARS-CoV-2  
226 might be co-infected, i.e., TB was often diagnosed concurrently or after Covid-19  
227 infection and the co-infection might account for increased case fatality rate (Group  
228 2021). Our finding brought up a possible explanation that patients with rare mutations  
229 enriched in Tuberculosis pathway were more likely to develop severe Covid-19  
230 symptoms. When focused on only pLoF variants enriched on KEGG, two significant  
231 pathways highlighted: Primary immunodeficiency (hsa:05340,  $p_{\text{adjust}} = 0.014$ ) and  
232 Influenza A (hsa:05164,  $p_{\text{adjust}} = 0.021$ ) ([Supplementary Table S2](#)). Primary  
233 immunodeficiencies (PID) are a group of potentially serious disorders that can cause  
234 increased susceptibility to severe infections, autoimmune diseases, and malignancy.  
235 Several studies revealed that patients with PID displayed higher morbidity and  
236 mortality from Covid-19 (Shields et al. 2021; Ho et al. 2021; Babaha and Rezaei 2020).  
237 Influenza is an infectious respiratory disease caused by influenza virus. Bibert et al.  
238 observed that gene pathways involved in the detection of Influenza A overlapped with  
239 those involved in the detection of SARS-CoV-2 virus (Bibert et al. 2021). In these two  
240 biological pathways, three functional genes, *IKBKG*, *IRF7*, and *IFNAR1*, were  
241 previously identified to have an effect on Covid-19 severity (Q. Zhang et al. 2020).

#### 242 **Tested on 148 Candidate Genes**

243 In addition to uncovering unknown possibly associated genes or pathways, we also  
244 tested 148 previously reported candidate genes, with 13 in the type I IFN pathway (Q.  
245 Zhang et al. 2020) and 136 located within 50kb of significant common variants in the  
246 Covid-19 HGI (Initiative and Ganna 2021). Specifically, we focused on likely-  
247 deleterious missense and HC-pLoF variants to aggregate potential effects of rare  
248 variants. For missense variants in both IFN- and HGI-genes, we did not detect  
249 significant difference between *severe* and *non-severe* patients. In the 13 IFN-genes, we  
250 found one HC-pLoF variant rs780744847 (c.1180C>T/p.R394\*) on *TLR3* mutated in  
251 only *severe* patients but no mutations in *non-severe* patients. It was reported that the  
252 *TLR3* deficiency may lead to increased incidences of viral infections and impair the  
253 production of type I IFN throughout SARS-CoV-2 infection (Q. Zhang et al. 2020).  
254 Moreover, mutations of inborn errors of TLR3-dependent type I IFN immunity more  
255 often occurred in highly critical patients than in mild patients and healthy controls. For  
256 the 136 HGI-genes, we found that the number of HC-pLoF mutations occurred in *severe*  
257 group was more than that of *non-severe* groups (16 in *severe* and 2 in *non-severe*



258 patients, Fisher’s exact test  $p = 0.043$ ) (Table 2). We also detected a HC-pLoF variant  
259 rs541048548 (c.1053del/p.A352fs) on *ICAM3* only mutated in *severe* patients. The  
260 gene *ICAM3* played an important role in the immunopathogenesis of SARS virus (Chan  
261 et al. 2007) and had been reported that its expression was downregulated in  
262 asymptomatic Covid-19 cases compared with symptomatic patients (Masood et al.  
263 2021).

#### 264 **Mutation Accumulation Analys**

265 In the mutation accumulation analysis, we first investigated whether there were  
266 potentially functional mutations unique to *severe* patients. We filtered in rare variants  
267 mutated in only *severe* patients and with minor allele count (MAC) greater than or equal  
268 to three. This resulted in 756 rare variants mapped to 700 genes. Among these variants,  
269 we observed a very rare mutation rs777044791 in gene *CCR3* at locus 3p21.31 (Table  
270 3). The physical distance between rs777044791 and rs11385942 is 0.43MB (GRCh38),  
271 a distance typically flanked into the same genomic region (Casto and Feldman 2011).  
272 The variant rs11385942 is a common variant located at locus 3p21.31 and was first  
273 identified to be associated with respiratory failure due to Covid-19 from GWAS  
274 analysis in the Italian and Spanish population (Ellinghaus et al. 2020). This finding was  
275 repeated in other studies based on European populations (Ellinghaus et al. 2020; Pairo-  
276 Castineira et al. 2021; Shelton et al. 2021), verifying its effects on Covid-19 disease. In  
277 Chinese population, common variant studies at this locus did not replicate significance  
278 (P. Wu, Chen, et al. 2021; H. Zhu et al. 2021; Wang et al. 2020), and no rare variant  
279 studies had been conducted. Our work filled this gap and revealed 3p21.31 as the  
280 Covid-19 risk locus.

281 Then, we performed PPI network analysis for the 700 “individual variant-driven”  
282 genes with the 148 known genes (Figure 3A). From the results, we found two candidate  
283 genes *IL12RB1* and *TRAF3IP3* that had extensive interactions with IFN- and HGI-  
284 genes. Gene *IL12RB1* (Interleukin 12 Receptor Subunit Beta 1) encodes a type I  
285 transmembrane protein that binds to interleukin-12 (IL12) and is involved in IL12  
286 transduction. Mutations in *IL12RB1* damage the development of IL17-producing T  
287 lymphocytes and increase the susceptibility to Salmonella and mycobacterial infections  
288 (van de Vosse et al. 2013). Our PPI network analysis indicated that *IL12RB1* and *TYK2*  
289 had experimentally determined interactions, which were compiled from a set of public  
290 databases and were more likely to be credible (Q. C. Zhang et al. 2013). Gene *TYK2*  
291 had been previously identified to be associated with Covid-19 critical illness (Pairo-  
292 Castineira et al. 2021), implying the potential effects of *IL12RB1* to the aggravation of  
293 Covid-19. The gene *TRAF3IP3* (TRAF3 Interacting Protein 3) encodes a protein that  
294 plays essential roles in both innate and adaptive immunity. Knockout mouse  
295 experiments of this gene observed a decrease in white blood cell count in males and an  
296 increased susceptibility to bacterial infection (GARDIN and WHITE 2011). In our  
297 results, *TRAF3IP3* was experimentally determined with protein TRAF3 encoded by  
298 gene *TRAF3*, which was included in a newly created pathway “Activation of NLRP3

299 inflammasome by SARS-CoV-2” (WP4876) (Siu et al. 2019). In response to viral  
300 infection, *TRAF3IP3* bridges *TRAF3* and *MAVS* leading to interferon production,  
301 indicating its probably strong relationship with Covid-19 disease.

302 We also performed PPI network analysis for the 778 “all variant-driven” genes with  
303 the 148 known genes (Figure 3B), from which two genes, *TBK1* and *TLR3*, were  
304 highlighted. Specifically, *TBK1* (TANK Binding Kinase 1) encodes a protein that plays  
305 important roles in antiviral innate immune response and in regulating inflammatory  
306 response to foreign agents (Fitzgerald et al. 2003; Mori et al. 2004). A previous study  
307 observed colocalization of *TBK1* with the M protein of SARS-CoV-2, which might  
308 hinder the dsRNA-induced IFN production at the step or upstream of *TBK1* (Zheng et  
309 al. 2020). The gene *TLR3* (Toll Like Receptor 3) encodes a member of TLR family that  
310 plays a primary role in recognition of pathogen and innate immunity activation. It  
311 recognizes dsRNA participated in multiple viral infections and induces type I IFNs  
312 production (Kawai and Akira 2007).

313 In summary, our mutation accumulation analyses and PPI network analyses  
314 suggested that locus 3p21.31, *IL12RB1*, *TRAF3IP3*, *TBK1*, and *TLR3* as key regions in  
315 severe Covid-19 patients compared with non-severe, implying their functions and  
316 associations with Covid-19 severity.

317

318

## 319 Discussion

320 SARS-CoV-2 is a strain of coronavirus and is highly pathogenic and transmissible.  
321 After exposure to the virus, ordinary people may do not develop noticeable symptoms  
322 or develop mild to moderate symptoms, while people with inborn errors of immunity  
323 tend to suffer severe and critical symptoms, or even death (S.-Y. Zhang et al. 2020).  
324 There is increasing evidence that the host genetic variants in genes related to  
325 immunodeficiency or inflammasomes might attribute to Covid-19 clinical  
326 manifestations (Elhabyan et al. 2020). A bunch of clinical drug treatments for Covid-  
327 19 were cultivated from this finding, including type I IFNs (e.g., IFN- $\alpha$ 1b), TNF  
328 inhibitors, anti-IFN- $\gamma$  antibodies, JAK1 inhibitors, and STAT1 inhibitors (Ku, Chen,  
329 and Lai 2021).

330 In this work, we carried out the first study of rare variants in inborn errors of  
331 immunity genes associated with Covid-19 severity in Chinese population. The  
332 identified functional candidate pathways Tuberculosis, Primary Immunodeficiency, and  
333 Influenza A were previously known to be part of antiviral immune responses and viral  
334 eradication, and we discovered their potential influences in Covid-19. We also  
335 suggested several putative genetic regions probably involved in susceptibility and  
336 severity of Covid-19, including locus 3p21.31, genes *IL12RB1*, *TRAF3IP3*, *TBK1*, and  
337 *TLR3*.

338 Despite the many compelling and significant findings of our work, there are still a  
339 few limitations to be noted. First, the sample size we used is relatively small, and the  
340 limited sample size limits the statistical power for identifying rare variants. More  
341 studies with large sample sizes are demanded to validate our results and uncover more  
342 candidate variants. Second, even though our work has suggested several candidate  
343 genes and pathways potentially related to Covid-19 severity, the true underlying genetic  
344 mechanisms of how they affect disease progression need to be explored by more  
345 persuasive experimental designs.

346 Covid-19 is assessed as a complex infectious disease and affected many risk factors.  
347 Symptoms of Covid-19 are highly variable, ranging from unnoticeable to severe and  
348 even death. The host genetic background is only partly responsible for the phenotypic  
349 heterogeneity. In recent years, multi-omics studies have proven a powerful and  
350 successful strategy to provide a broader perspective in understanding disease  
351 development and biological phenomena. Several multi-omics analyses of Covid-19  
352 have been proposed to integrate multiple “omes” data to unravel disease mechanisms  
353 at multiple omics levels (Overmyer et al. 2021; Su et al. 2020; Montaldo et al. 2021; P.  
354 Wu, Chen, et al. 2021; Stephenson et al. 2021). The integrative analyses of rare genome  
355 and other “omes” data (e.g., proteome, transcriptome, epigenome, metabolome, and  
356 microbiome) may inspire us to discover new risk factors for severe Covid-19 disease.

357

358 **Conflict of Interest**

359 The authors declare no competing interests.

360

361 **Author Contribution**

362 X.J., H.Z. conceived the study, designed the research program, and managed the project.

363 F.C., F.Z., Y.J. collected the samples.

364 Y.L. finished the laboratory processing and data acquisition.

365 P.L., H.Z. performed the statistical analyses.

366 M.F advised on statistical methods.

367 P.L., H.Z. wrote the manuscript.

368 All authors participated in revising the manuscript.

369

370 **Funding**

371 This study was supported by National Natural Science Foundation of China (No.

372 31800765, 32171441, 32000398), Natural Science Foundation of Guangdong Province,

373 China (2017A030306026), Guangdong-Hong Kong Joint Laboratory on

374 Immunological and Genetic Kidney Diseases (2019B121205005).

375

376 **Acknowledgement**

377 We thank the families who participated in the study and made this research study

378 possible.

379

380 **Data Availability**

381 The original contributions presented in the study are included in the

382 article/Supplementary Material. Further inquiries can be directed to the corresponding

383 author.

384

385

386

## 387 **References**

- 388 Babaha, Fateme, and Nima Rezaei. 2020. “Primary Immunodeficiency Diseases in  
389 COVID-19 Pandemic: A Predisposing or Protective Factor?” *The American*  
390 *Journal of the Medical Sciences*. <https://doi.org/10.1016/j.amjms.2020.07.027>.
- 391 Bibert, Stéphanie, Nicolas Guex, Joao Lourenco, Thomas Brahier, Matthaios  
392 Papadimitriou-Olivgeris, Lauro Damonti, Oriol Manuel, et al. 2021.  
393 “Transcriptomic Signature Differences Between SARS-CoV-2 and Influenza  
394 Virus Infected Patients .” *Frontiers in Immunology* .  
395 <https://www.frontiersin.org/article/10.3389/fimmu.2021.666163>.
- 396 Browning, Brian L, and Sharon R Browning. 2016. “Genotype Imputation with  
397 Millions of Reference Samples.” *American Journal of Human Genetics* 98 (1):  
398 116–26. <https://doi.org/10.1016/j.ajhg.2015.11.020>.
- 399 Casto, Amanda M, and Marcus W Feldman. 2011. “Genome-Wide Association Study  
400 SNPs in the Human Genome Diversity Project Populations: Does Selection  
401 Affect Unlinked SNPs with Shared Trait Associations?” *PLoS Genetics* 7 (1):  
402 e1001266. <https://doi.org/10.1371/journal.pgen.1001266>.
- 403 Chan, Kelvin Y K, Johannes C Y Ching, M S Xu, Annie N Y Cheung, Shea-Ping Yip,  
404 Loretta Y C Yam, Sik-To Lai, et al. 2007. “Association of ICAM3 Genetic Variant  
405 with Severe Acute Respiratory Syndrome.” *The Journal of Infectious Diseases* 196  
406 (2): 271–80. <https://doi.org/10.1086/518892>.
- 407 Chang, Christopher C, Carson C Chow, Laurent Cam Tellier, Shashaank Vattikuti,  
408 Shaun M Purcell, and James J Lee. 2015. “Second-Generation PLINK: Rising to  
409 the Challenge of Larger and Richer Datasets.” *GigaScience* 4: 7.  
410 <https://doi.org/10.1186/s13742-015-0047-8>.
- 411 Clarke, Laura, Xiangqun Zheng-Bradley, Richard Smith, Eugene Kulesha, Chunlin  
412 Xiao, Iliana Toneva, Brendan Vaughan, et al. 2012. “The 1000 Genomes Project:  
413 Data Management and Community Access.” *Nature Methods* 9 (5): 459–62.  
414 <https://doi.org/10.1038/nmeth.1974>.
- 415 Cummings, Matthew J, Matthew R Baldwin, Darryl Abrams, Samuel D Jacobson,  
416 Benjamin J Meyer, Elizabeth M Balough, Justin G Aaron, et al. 2020.  
417 “Epidemiology, Clinical Course, and Outcomes of Critically Ill Adults with  
418 COVID-19 in New York City: A Prospective Cohort Study.” *Lancet (London,*  
419 *England)* 395 (10239): 1763–70. [https://doi.org/10.1016/S0140-6736\(20\)31189-2](https://doi.org/10.1016/S0140-6736(20)31189-2).
- 420 Danecek, Petr, James K Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin  
421 O Pollard, Andrew Whitwham, et al. 2021. “Twelve Years of SAMtools and  
422 BCFtools.” *GigaScience* 10 (2): giab008.  
423 <https://doi.org/10.1093/gigascience/giab008>.

- 424 DePristo, Mark A, Eric Banks, Ryan Poplin, Kiran v Garimella, Jared R Maguire,  
425 Christopher Hartl, Anthony A Philippakis, et al. 2011. “A Framework for Variation  
426 Discovery and Genotyping Using Next-Generation DNA Sequencing Data.”  
427 *Nature Genetics* 43 (5): 491–98. <https://doi.org/10.1038/ng.806>.
- 428 Elhabyan, Abdelazeem, Saja Elyaacoub, Ehab Sanad, Abdelwahab Abukhadra, Asmaa  
429 Elhabyan, and Valentin Dinu. 2020. “The Role of Host Genetics in Susceptibility  
430 to Severe Viral Infections in Humans and Insights into Host Genetics of Severe  
431 COVID-19: A Systematic Review.” *Virus Research* 289 (November): 198163.  
432 <https://doi.org/10.1016/j.virusres.2020.198163>.
- 433 Ellinghaus, David, Frauke Degenhardt, Luis Bujanda, Maria Buti, Agustín Albillos,  
434 Pietro Invernizzi, Javier Fernández, et al. 2020. “Genomewide Association Study  
435 of Severe Covid-19 with Respiratory Failure.” *The New England Journal of  
436 Medicine*, June. <https://doi.org/10.1056/NEJMoa2020283>.
- 437 Fitzgerald, Katherine A, Sarah M McWhirter, Kerrie L Faia, Daniel C Rowe, Eicke Latz,  
438 Douglas T Golenbock, Anthony J Coyle, Sha-Mei Liao, and Tom Maniatis. 2003.  
439 “IKKepsilon and TBK1 Are Essential Components of the IRF3 Signaling  
440 Pathway.” *Nature Immunology* 4 (5): 491–96. <https://doi.org/10.1038/ni921>.
- 441 Freed, Donald, Rafael Aldana, Jessica Weber, and Jeremy Edwards. 2017. “The  
442 Sentieon Genomics Tools - A Fast and Accurate Solution to Variant Calling from  
443 next-Generation Sequence Data.” bioRxiv. <https://doi.org/10.1101/115717>.
- 444 GARDIN, A, and J WHITE. 2011. “The Sanger Mouse Genetics Programme: High  
445 Throughput Characterisation of Knockout Mice.” *Acta Ophthalmologica* 89  
446 (s248): 0. <https://doi.org/https://doi.org/10.1111/j.1755-3768.2011.4451.x>.
- 447 Gorbalenya, Alexander E, Susan C Baker, Ralph S Baric, Raoul J de Groot, Christian  
448 Drosten, Anastasia A Gulyaeva, Bart L Haagmans, et al. 2020. “The Species  
449 Severe Acute Respiratory Syndrome-Related Coronavirus: Classifying 2019-  
450 NCoV and Naming It SARS-CoV-2.” *Nature Microbiology* 5 (4): 536–44.  
451 <https://doi.org/10.1038/s41564-020-0695-z>.
- 452 Gorlov, I P, O Y Gorlova, M L Frazier, M R Spitz, and C I Amos. 2011. “Evolutionary  
453 Evidence of the Effect of Rare Variants on Disease Etiology.” *Clinical Genetics*  
454 79 (3): 199–206. <https://doi.org/https://doi.org/10.1111/j.1399-0004.2010.01535.x>.
- 455 Group, The TB/COVID-19 Global Study. 2021. “Tuberculosis and COVID-19 Co-  
456 Infection: Description of the Global Cohort.” *European Respiratory Journal*,  
457 January, 2102538. <https://doi.org/10.1183/13993003.02538-2021>.
- 458 Ho, Hsi-En, Sheryl Mathew, Michael J Peluso, and Charlotte Cunningham-Rundles.  
459 2021. “Clinical Outcomes and Features of COVID-19 in Patients with  
460 Primary Immunodeficiencies in New York City.” *The Journal of Allergy and  
461 Clinical Immunology. In Practice* 9 (1): 490-493.e2.



- 462 <https://doi.org/10.1016/j.jaip.2020.09.052>.
- 463 Hunt, Karen A, Vanisha Mistry, Nicholas A Bockett, Tariq Ahmad, Maria Ban, Jonathan  
464 N Barker, Jeffrey C Barrett, et al. 2013. “Negligible Impact of Rare Autoimmune-  
465 Locus Coding-Region Variants on Missing Heritability.” *Nature* 498 (7453): 232–  
466 35. <https://doi.org/10.1038/nature12170>.
- 467 Initiative, COVID-19 Host Genetics, and Andrea Ganna. 2021. “Mapping the Human  
468 Genetic Architecture of COVID-19: An Update.” *MedRxiv*, January,  
469 2021.11.08.21265944. <https://doi.org/10.1101/2021.11.08.21265944>.
- 470 Itan, Yuval, Lei Shang, Bertrand Boisson, Michael J Ciancanelli, Janet G Markle,  
471 Ruben Martinez-Barricarte, Eric Scott, et al. 2016. “The Mutation Significance  
472 Cutoff: Gene-Level Thresholds for Variant Predictions.” *Nature Methods* 13 (2):  
473 109–10. <https://doi.org/10.1038/nmeth.3739>.
- 474 Kanehisa, Minoru, and Susumu Goto. 2000. “KEGG: Kyoto Encyclopedia of Genes  
475 and Genomes.” *Nucleic Acids Research* 28 (1): 27–30.  
476 <https://doi.org/10.1093/nar/28.1.27>.
- 477 Karczewski, Konrad J, Laurent C Francioli, Grace Tiao, Beryl B Cummings, Jessica  
478 Alföldi, Qingbo Wang, Ryan L Collins, et al. 2020. “The Mutational Constraint  
479 Spectrum Quantified from Variation in 141,456 Humans.” *Nature* 581 (7809):  
480 434–43. <https://doi.org/10.1038/s41586-020-2308-7>.
- 481 Kawai, Taro, and Shizuo Akira. 2007. “Signaling to NF-KappaB by Toll-like Receptors.”  
482 *Trends in Molecular Medicine* 13 (11): 460–69.  
483 <https://doi.org/10.1016/j.molmed.2007.09.002>.
- 484 Kircher, Martin, Daniela M Witten, Preti Jain, Brian J O’Roak, Gregory M Cooper, and  
485 Jay Shendure. 2014. “A General Framework for Estimating the Relative  
486 Pathogenicity of Human Genetic Variants.” *Nature Genetics* 46 (3): 310–15.  
487 <https://doi.org/10.1038/ng.2892>.
- 488 Ku, Cheng-Lung, I-Ting Chen, and Ming-Zong Lai. 2021. “Infection-Induced  
489 Inflammation from Specific Inborn Errors of Immunity to COVID-19.” *The FEBS*  
490 *Journal* 288 (17): 5021–41. <https://doi.org/10.1111/febs.15961>.
- 491 Lee, Seunggeun, Michael C Wu, and Xihong Lin. 2012. “Optimal Tests for Rare Variant  
492 Effects in Sequencing Association Studies.” *Biostatistics (Oxford, England)* 13 (4):  
493 762–75. <https://doi.org/10.1093/biostatistics/kxs014>.
- 494 Li, Miaoxin, Jiang Li, Mulin Jun Li, Zhicheng Pan, Jacob Shujui Hsu, Dajiang J Liu,  
495 Xiaowei Zhan, Junwen Wang, Youqiang Song, and Pak Chung Sham. 2017.  
496 “Robust and Rapid Algorithms Facilitate Large-Scale Whole Genome  
497 Sequencing Downstream Analysis in an Integrative Framework.” *Nucleic Acids*  
498 *Research* 45 (9): e75. <https://doi.org/10.1093/nar/gkx019>.

- 499 Mantovani, Stefania, Sergio Daga, Chiara Fallerini, Margherita Baldassarri, Elisa  
500 Benetti, Nicola Picchiotti, Francesca Fava, et al. 2021. “Rare Variants in Toll-like  
501 Receptor 7 Results in Functional Impairment and Downregulation of Cytokine-  
502 Mediated Signaling in COVID-19 Patients.” *Genes and Immunity*, December, 1–  
503 6. <https://doi.org/10.1038/s41435-021-00157-1>.
- 504 Masood, Kiran Iqbal, Maliha Yameen, Javeria Ashraf, Saba Shahid, Syed Faisal  
505 Mahmood, Asghar Nasir, Nosheen Nasir, et al. 2021. “Upregulated Type I  
506 Interferon Responses in Asymptomatic COVID-19 Infection Are Associated with  
507 Improved Clinical Outcome.” *Scientific Reports* 11 (1): 22958.  
508 <https://doi.org/10.1038/s41598-021-02489-4>.
- 509 McLaren, William, Laurent Gil, Sarah E Hunt, Harpreet Singh Riat, Graham R S  
510 Ritchie, Anja Thormann, Paul Flicek, and Fiona Cunningham. 2016. “The  
511 Ensembl Variant Effect Predictor.” *Genome Biology* 17 (1): 122.  
512 <https://doi.org/10.1186/s13059-016-0974-4>.
- 513 Misawa, Kazuharu, Takanori Hasegawa, Eikan Mishima, Promsuk Jutabha, Motoshi  
514 Ouchi, Kaname Kojima, Yosuke Kawai, Masafumi Matsuo, Naohiko Anzai, and  
515 Masao Nagasaki. 2020. “Contribution of Rare Variants of the SLC22A12 Gene to  
516 the Missing Heritability of Serum Urate Levels.” *Genetics* 214 (4): 1079–90.  
517 <https://doi.org/10.1534/genetics.119.303006>.
- 518 Montaldo, Chiara, Francesco Messina, Isabella Abbate, Manuela Antonioli, Veronica  
519 Bordoni, Alessandra Aiello, Fabiola Ciccocanti, et al. 2021. “Multi-Omics  
520 Approach to COVID-19: A Domain-Based Literature Review.” *Journal of*  
521 *Translational Medicine* 19 (1): 501. <https://doi.org/10.1186/s12967-021-03168-8>.
- 522 Mori, Mitsuaki, Mitsutoshi Yoneyama, Takashi Ito, Kiyohiro Takahashi, Fuyuhiko  
523 Inagaki, and Takashi Fujita. 2004. “Identification of Ser-386 of Interferon  
524 Regulatory Factor 3 as Critical Target for Inducible Phosphorylation That  
525 Determines Activation.” *The Journal of Biological Chemistry* 279 (11): 9698–  
526 9702. <https://doi.org/10.1074/jbc.M310616200>.
- 527 Overmyer, Katherine A, Evgenia Shishkova, Ian J Miller, Joseph Balnis, Matthew N  
528 Bernstein, Trenton M Peters-Clarke, Jesse G Meyer, et al. 2021. “Large-Scale  
529 Multi-Omic Analysis of COVID-19 Severity.” *Cell Systems* 12 (1): 23-40.e7.  
530 <https://doi.org/10.1016/j.cels.2020.10.003>.
- 531 Pairo-Castineira, Erola, Sara Clohisey, Lucija Klaric, Andrew D Bretherick, Konrad  
532 Rawlik, Dorota Pasko, Susan Walker, et al. 2021. “Genetic Mechanisms of Critical  
533 Illness in COVID-19.” *Nature* 591 (7848): 92–98. <https://doi.org/10.1038/s41586-020-03065-y>.
- 535 Shelton, Janie F, Anjali J Shastri, Chelsea Ye, Catherine H Weldon, Teresa Filshtein-  
536 Sonmez, Daniella Coker, Antony Symons, Jorge Esparza-Gordillo, Stella

- 537 Aslibekyan, and Adam Auton. 2021. “Trans-Ancestry Analysis Reveals Genetic  
538 and Nongenetic Associations with COVID-19 Susceptibility and Severity.”  
539 *Nature Genetics* 53 (6): 801–8. <https://doi.org/10.1038/s41588-021-00854-7>.
- 540 Sherry, S T, M.-H. Ward, M Kholodov, J Baker, L Phan, E M Smigielski, and K Sirotkin.  
541 2001. “DbSNP: The NCBI Database of Genetic Variation.” *Nucleic Acids*  
542 *Research* 29 (1): 308–11. <https://doi.org/10.1093/nar/29.1.308>.
- 543 Shields, Adrian M, Siobhan O Burns, Sinisa Savic, and Alex G Richter. 2021. “COVID-  
544 19 in Patients with Primary and Secondary Immunodeficiency: The United  
545 Kingdom Experience.” *The Journal of Allergy and Clinical Immunology* 147 (3):  
546 870-875.e1. <https://doi.org/10.1016/j.jaci.2020.12.620>.
- 547 Siu, Kam-Leung, Kit-San Yuen, Carlos Castaño-Rodríguez, Zi-Wei Ye, Man-Lung  
548 Yeung, Sin-Yee Fung, Shuofeng Yuan, et al. 2019. “Severe Acute Respiratory  
549 Syndrome Coronavirus ORF3a Protein Activates the NLRP3 Inflammasome by  
550 Promoting TRAF3-Dependent Ubiquitination of ASC.” *FASEB Journal : Official*  
551 *Publication of the Federation of American Societies for Experimental Biology* 33  
552 (8): 8865–77. <https://doi.org/10.1096/fj.201802418R>.
- 553 Smieszek, Sandra P, Vasilios M Polymeropoulos, Changfu Xiao, Christos M  
554 Polymeropoulos, and Mihael H Polymeropoulos. 2021. “Loss-of-Function  
555 Mutations in IFNAR2 in COVID-19 Severe Infection Susceptibility.” *Journal of*  
556 *Global Antimicrobial Resistance* 26 (July): 239–40.  
557 <https://doi.org/10.1016/j.jgar.2021.06.005>.
- 558 Stephenson, Emily, Gary Reynolds, Rachel A Botting, Fernando J Calero-Nieto,  
559 Michael D Morgan, Zewen Kelvin Tuong, Karsten Bach, et al. 2021. “Single-Cell  
560 Multi-Omics Analysis of the Immune Response in COVID-19.” *Nature Medicine*  
561 27 (5): 904–16. <https://doi.org/10.1038/s41591-021-01329-2>.
- 562 Su, Yapeng, Daniel Chen, Dan Yuan, Christopher Lausted, Jongchan Choi, Chengzhen  
563 L Dai, Valentin Voillet, et al. 2020. “Multi-Omics Resolves a Sharp Disease-State  
564 Shift between Mild and Moderate COVID-19.” *Cell* 183 (6): 1479-1495.e20.  
565 <https://doi.org/10.1016/j.cell.2020.10.037>.
- 566 Sweeney, Timothy E, Shane Lofgren, Purvesh Khatri, and Angela J Rogers. 2017.  
567 “Gene Expression Analysis to Assess the Relevance of Rodent Models to Human  
568 Lung Injury.” *American Journal of Respiratory Cell and Molecular Biology* 57  
569 (2): 184–92. <https://doi.org/10.1165/rcmb.2016-0395OC>.
- 570 Szklarczyk, Damian, Annika L Gable, David Lyon, Alexander Junge, Stefan Wyder,  
571 Jaime Huerta-Cepas, Milan Simonovic, et al. 2019. “STRING V11: Protein–  
572 Protein Association Networks with Increased Coverage, Supporting Functional  
573 Discovery in Genome-Wide Experimental Datasets.” *Nucleic Acids Research* 47  
574 (D1): D607–13. <https://doi.org/10.1093/nar/gky1131>.

- 575 Vosse, Esther van de, Margje H Haverkamp, Noe Ramirez-Alejo, Mónica Martinez-  
576 Gallo, Lizbeth Blancas-Galicia, Ayşe Metin, Ben Zion Garty, et al. 2013. “IL-  
577 12Rβ1 Deficiency: Mutation Update and Description of the IL12RB1  
578 Variation Database.” *Human Mutation* 34 (10): 1329–39.  
579 <https://doi.org/10.1002/humu.22380>.
- 580 Wang, Fang, Shujia Huang, Huirong Gao, Yuwen Zhou, Changxiang Lai, Zhichao Li,  
581 Wenjie Xian, et al. 2020. “Initial Whole Genome Sequencing and Analysis of the  
582 Host Genetic Contribution to COVID-19 Severity and Susceptibility.” *MedRxiv*,  
583 January, 2020.06.09.20126607. <https://doi.org/10.1101/2020.06.09.20126607>.
- 584 Williams, Frances M K, Maxim Freydin, Massimo Mangino, Simon Couvreur, Alessia  
585 Visconti, Ruth C E Bowyer, Caroline I le Roy, et al. 2020. “Self-Reported  
586 Symptoms of Covid-19 Including Symptoms Most Predictive of SARS-CoV-2  
587 Infection, Are Heritable.” *MedRxiv*, January, 2020.04.22.20072124.  
588 <https://doi.org/10.1101/2020.04.22.20072124>.
- 589 Wu, Michael C, Peter Kraft, Michael P Epstein, Deanne M Taylor, Stephen J Chanock,  
590 David J Hunter, and Xihong Lin. 2010. “Powerful SNP-Set Analysis for Case-  
591 Control Genome-Wide Association Studies.” *American Journal of Human*  
592 *Genetics* 86 (6): 929–42. <https://doi.org/10.1016/j.ajhg.2010.05.002>.
- 593 Wu, Peng, Dongsheng Chen, Wencheng Ding, Ping Wu, Hongyan Hou, Yong Bai,  
594 Yuwen Zhou, et al. 2021. “The Trans-Omics Landscape of COVID-19.” *Nature*  
595 *Communications* 12 (1): 4543. <https://doi.org/10.1038/s41467-021-24482-1>.
- 596 Wu, Peng, Lin Ding, Xiaodong Li, Siyang Liu, Fanjun Cheng, Qing He, Mingzhong  
597 Xiao, et al. 2021. “Trans-Ethnic Genome-Wide Association Study of Severe  
598 COVID-19.” *Communications Biology* 4 (1): 1034.  
599 <https://doi.org/10.1038/s42003-021-02549-5>.
- 600 Wu, Zunyou, and Jennifer M McGoogan. 2020. “Characteristics of and Important  
601 Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China:  
602 Summary of a Report of 72 314 Cases From the Chinese Center for Disease  
603 Control and Prevention.” *JAMA* 323 (13): 1239–42.  
604 <https://doi.org/10.1001/jama.2020.2648>.
- 605 Yang, Xiaobo, Yuan Yu, Jiqian Xu, Huaqing Shu, Jia’an Xia, Hong Liu, Yongran Wu,  
606 et al. 2020. “Clinical Course and Outcomes of Critically Ill Patients with SARS-  
607 CoV-2 Pneumonia in Wuhan, China: A Single-Centered, Retrospective,  
608 Observational Study.” *The Lancet. Respiratory Medicine* 8 (5): 475–81.  
609 [https://doi.org/10.1016/S2213-2600\(20\)30079-5](https://doi.org/10.1016/S2213-2600(20)30079-5).
- 610 Zhang, Fan, Matthew Flickinger, Sarah A Gagliano Taliun, Gonçalo R Abecasis, Laura  
611 J Scott, Steven A McCarroll, Carlos N Pato, Michael Boehnke, and Hyun Min Kang.  
612 2020. “Ancestry-Agnostic Estimation of DNA Sample Contamination from

- 613 Sequence Reads.” *Genome Research* 30 (2): 185–94.  
614 <https://doi.org/10.1101/gr.246934.118>.
- 615 Zhang, Qian, Zhiyong Liu, Marcela Moncada-Velez, Jie Chen, Masato Ogishi,  
616 Benedetta Bigio, Rui Yang, et al. 2020. “Inborn Errors of Type I IFN Immunity in  
617 Patients with Life-Threatening COVID-19.” *Science* 370 (6515).  
618 <https://doi.org/10.1126/science.abd4570>.
- 619 Zhang, Qiangfeng Cliff, Donald Petrey, José Ignacio Garzón, Lei Deng, and Barry  
620 Honig. 2013. “PrePPI: A Structure-Informed Database of Protein-Protein  
621 Interactions.” *Nucleic Acids Research* 41 (Database issue): D828–33.  
622 <https://doi.org/10.1093/nar/gks1231>.
- 623 Zhang, Shen-Ying, Qian Zhang, Jean-Laurent Casanova, Helen C Su, and COVID  
624 Team. 2020. “Severe COVID-19 in the Young and Healthy: Monogenic Inborn  
625 Errors of Immunity?” *Nature Reviews. Immunology* 20 (8): 455–56.  
626 <https://doi.org/10.1038/s41577-020-0373-7>.
- 627 Zheng, Yi, Meng-Wei Zhuang, Lulu Han, Jing Zhang, Mei-Ling Nan, Peng Zhan,  
628 Dongwei Kang, Xinyong Liu, Chengjiang Gao, and Pei-Hui Wang. 2020. “Severe  
629 Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Membrane (M)  
630 Protein Inhibits Type I and III Interferon Production by Targeting RIG-I/MDA-5  
631 Signaling.” *Signal Transduction and Targeted Therapy* 5 (1): 299.  
632 <https://doi.org/10.1038/s41392-020-00438-7>.
- 633 Zhu, Huanhuan, Fang Zheng, Linxuan Li, Yan Jin, Yuxue Luo, Zhen Li, JingYu Zeng,  
634 et al. 2021. “A Chinese Host Genetic Study Discovered Type I Interferons and  
635 Causality of Cholesterol Levels and WBC Counts on COVID-19 Severity.”  
636 *MedRxiv*, January, 2021.06.04.21258335.  
637 <https://doi.org/10.1101/2021.06.04.21258335>.
- 638 Zhu, Na, Dingyu Zhang, Wenling Wang, Xingwang Li, Bo Yang, Jingdong Song, Xiang  
639 Zhao, et al. 2020. “A Novel Coronavirus from Patients with Pneumonia in China,  
640 2019.” *The New England Journal of Medicine* 382 (8): 727–33.  
641 <https://doi.org/10.1056/NEJMoa2001017>.
- 642 Zuk, Or, Stephen F Schaffner, Kaitlin Samocha, Ron Do, Eliana Hechter, Sekar  
643 Kathiresan, Mark J Daly, Benjamin M Neale, Shamil R Sunyaev, and Eric S  
644 Lander. 2014. “Searching for Missing Heritability: Designing Rare Variant  
645 Association Studies.” *Proceedings of the National Academy of Sciences* 111 (4):  
646 E455 LP-E464. <https://doi.org/10.1073/pnas.1322563111>.
- 647
- 648

649 **Figures**

650 **Figure 1. The flow diagram of rare variants analysis.** A total of 32,232,865 variants  
651 were identified from the 451 Covid-19 patients with whole genome sequencing. After  
652 filtering by VQSR and MAF, 13,934,341 rare variants were annotated by VEP, and  
653 42,730 candidate variants were included.

654  
655 **Figure 2. Quality estimate of the cohort.** (A). Heterozygote concordance rate versus  
656 sequencing depth for 218 array-genotyped individuals. (B). PCA of 159 non-severe  
657 and 292 severe patients.

658  
659 **Figure 3. The results of protein-protein interaction network analysis.** The plot of  
660 PPI network (A) between the 700 “individual variant-driven” genes with the 148  
661 known genes. (B) between the 778 “all variant-driven” genes with the 148 known  
662 genes.

663  
664 **Supplementary Figure S1. Gene-based tests.** (A). The QQ plots of gene-based  
665 association analyses between severe and non-severe patients for (A) . 42,730  
666 candidate rare variants ; (B) 38,548 rare likely-deleterious missense variants, and (C)  
667 4,182 high-confidence pLoF variants

668  
669



## Tables

**Table 1. Participant Characteristics**

	<b>Number, n (%)</b>	<b>Male gender, n (%)</b>	<b>Age, average (sd)</b>	<b>Depth, average (sd)</b>
<b>All patients</b>	451			21.67
<b>Non-Severe</b>	159 (35.25%)	67 (42.14%)	58.33 (14.62)	19.04 (8.94)
<b>Severe</b>	292 (64.75%)	154 (52.74%)	64.11 (13.31)	23.1 (10.43)

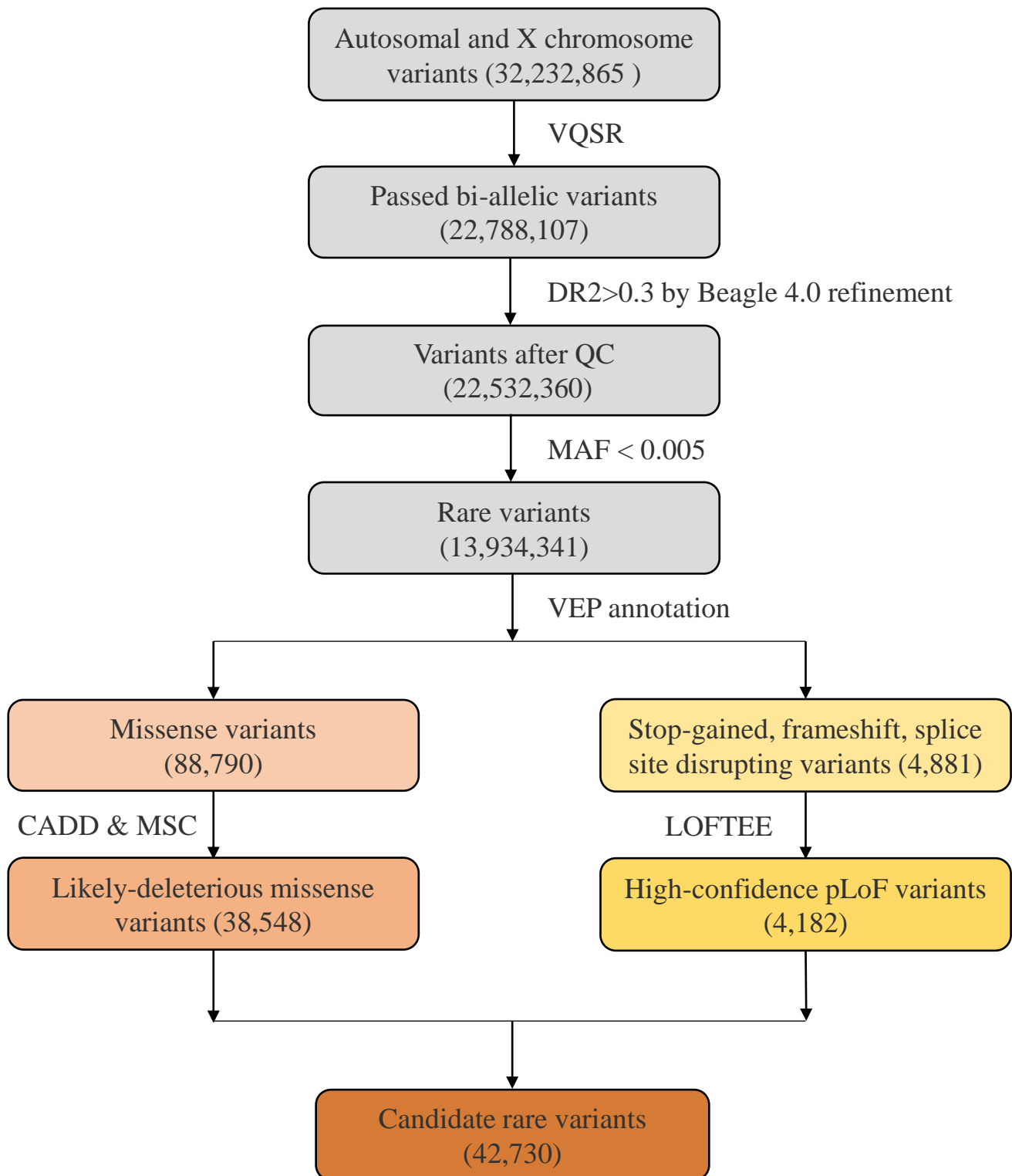
**Table 2. The pLoF variants identified in covid-19 patients in 148 candidate genes**

<b>GENE</b>	<b>SNP</b>	<b>Variant Annotation</b>	<b>HGVSc / HGVSg</b>	<b>Genotype</b>	<b>SampleID</b>	<b>Sex</b>	<b>Age Range</b>	<b>Phenotype</b>	<b>Category</b>
<i>TLR3</i>	rs780744847	stop gained	c.1180C>T / p.Arg394Ter	Het	U312	F	70-79	Severe	IFN-genes
<i>THBS3</i>	rs748584696	stop gained	c.853C>T / p.Arg285Ter	Het	U088	F	80-89	Severe	HGI-genes
<i>THBS3</i>	chr1_155206198 _A_C	splice donor variant	c.286+2T>G	Het	U359	F	70-79	Severe	HGI-genes
<i>TAC4</i>	rs372635644	splice acceptor variant	c.124-1G>A	Het	U429	F	60-69	Severe	HGI-genes
<i>TYK2</i>	rs770927552	frameshift variant	c.209_212del / p.Cys70SerfsTer21	Het	U422	F	80-89	Severe	HGI-genes
<i>C6orf15</i>	chr6_31112292_ C_T	splice acceptor variant	c.68-1G>A	Het	U107	M	60-69	Severe	HGI-genes
<i>CAT</i>	rs777641795	splice donor variant	c.1195+1G>A	Het	U012	M	40-49	Non-severe	HGI-genes
<i>CDH15</i>	chr16_89179469 _C_G	stop gained	c.96C>G / p.Tyr32*	Het	U174	F	50-59	Severe	HGI-genes
<i>CDSN</i>	chr6_31116133_ G_GA	frameshift variant	c.1481dup / p.Cys496LeufsTer20	Het	U021	F	50-59	Non-severe	HGI-genes
<i>ICAM3</i>	rs541048548	frameshift variant	c.1053delC / p.Ala352ArgfsTer11	Het	U225	M	70-79	Severe	HGI-genes
<i>ICAM3</i>	rs541048548	frameshift	c.1053delC /	Het	U047	F	70-79	Severe	HGI-genes

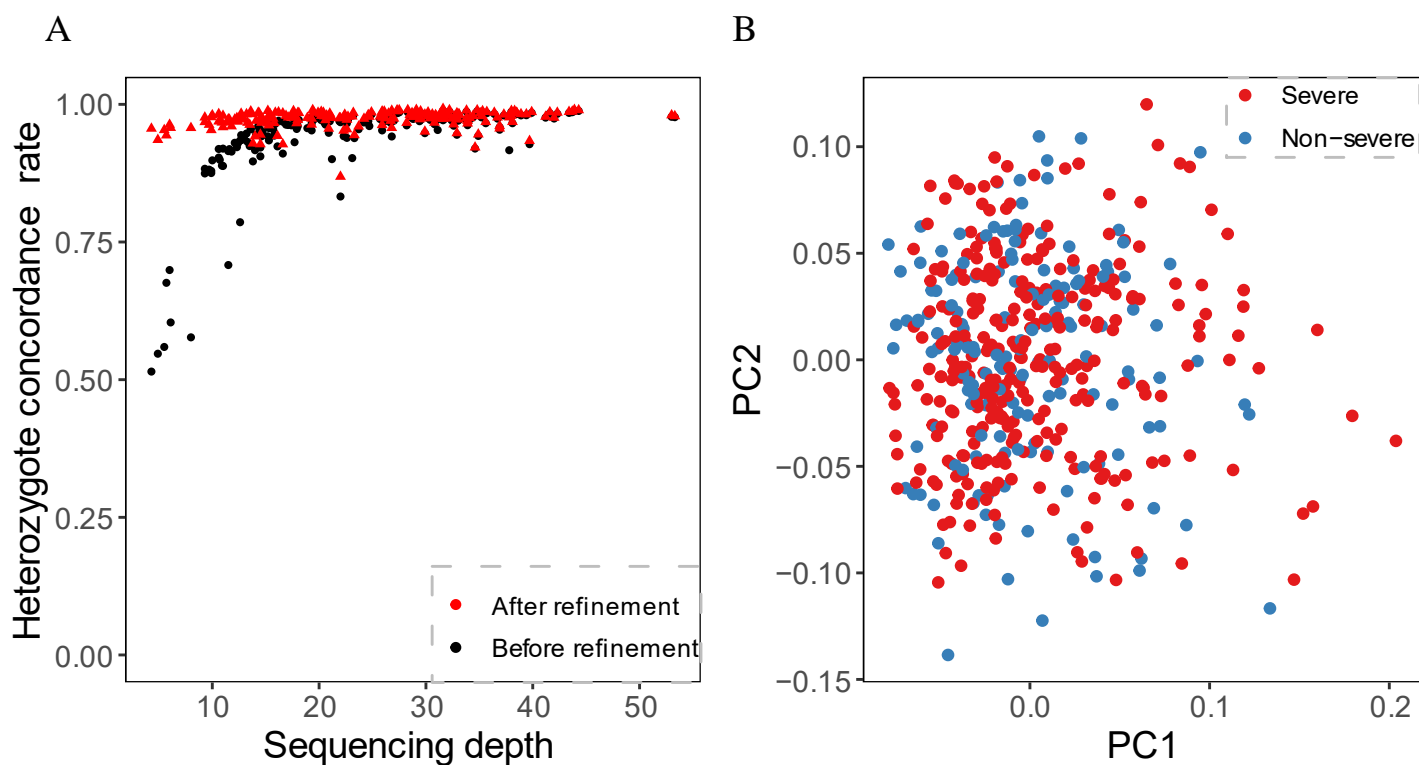
		variant	p.Ala352ArgfsTer11						
<i>MYDGF</i>	rs745851558	splice donor variant	c.174+1G>T	Het	U071	M	70-79	Severe	HGI-genes
<i>PLEKHA4</i>	chr19_48853718 _CA_C	frameshift variant	c.1289delT / p.Leu430ArgfsTer4	Het	U261	F	60-69	Severe	HGI-genes
<i>PLEKHA4</i>	chr19_48853720 _GCCGGT_G	frameshift variant	c.1283_1287delACC GG / p.Asp428AlafsTer76	Het	U261	F	60-69	Severe	HGI-genes
<i>PPP1R15 A</i>	rs768756506	frameshift variant	c.1535_1536delAT / p.Tyr512CysfsTer14	Het	U309	M	60-69	Severe	HGI-genes
<i>PSORSIC 2</i>	rs79153019	frameshift variant	c.281delC / p.Pro94LeufsTer35	Het	U075	M	60-69	Severe	HGI-genes
<i>PSORSIC 2</i>	rs79153019	frameshift variant	c.281delC / p.Pro94LeufsTer35	Het	U150	M	60-69	Severe	HGI-genes
<i>PSORSIC 2</i>	rs79153019	frameshift variant	c.281delC / p.Pro94LeufsTer35	Het	U144	M	60-69	Severe	HGI-genes
<i>TULP2</i>	chr19_48881045 _T_TC	frameshift variant	c.1528_1529insG / p.Gln510ArgfsTer17	Het	U176	F	70-79	Severe	HGI-genes

**Table 3. The comparison of allele frequency for two loci**

	<b>rs777044791</b>	<b>rs11385942</b>
CHROM	chr3	chr3
POS (hg38)	46,266,186	45,834,967
ALT	T	GA
REF	C	G
Variant Annotation	Missense Variant	Intron Variant
<b><i>Allele frequency</i></b>		
Severe (N = 292)	0.005	0
Non-severe (N = 159)	0	0
ChinaMAP	0.002	0.004
1000G_EAS	0	0.005
1000G_EUR	0	0.0805
1000G_SAS	0	0.296
1000G_AFR	0	0.053
gnomAD_EAS	0.0005	0.0006

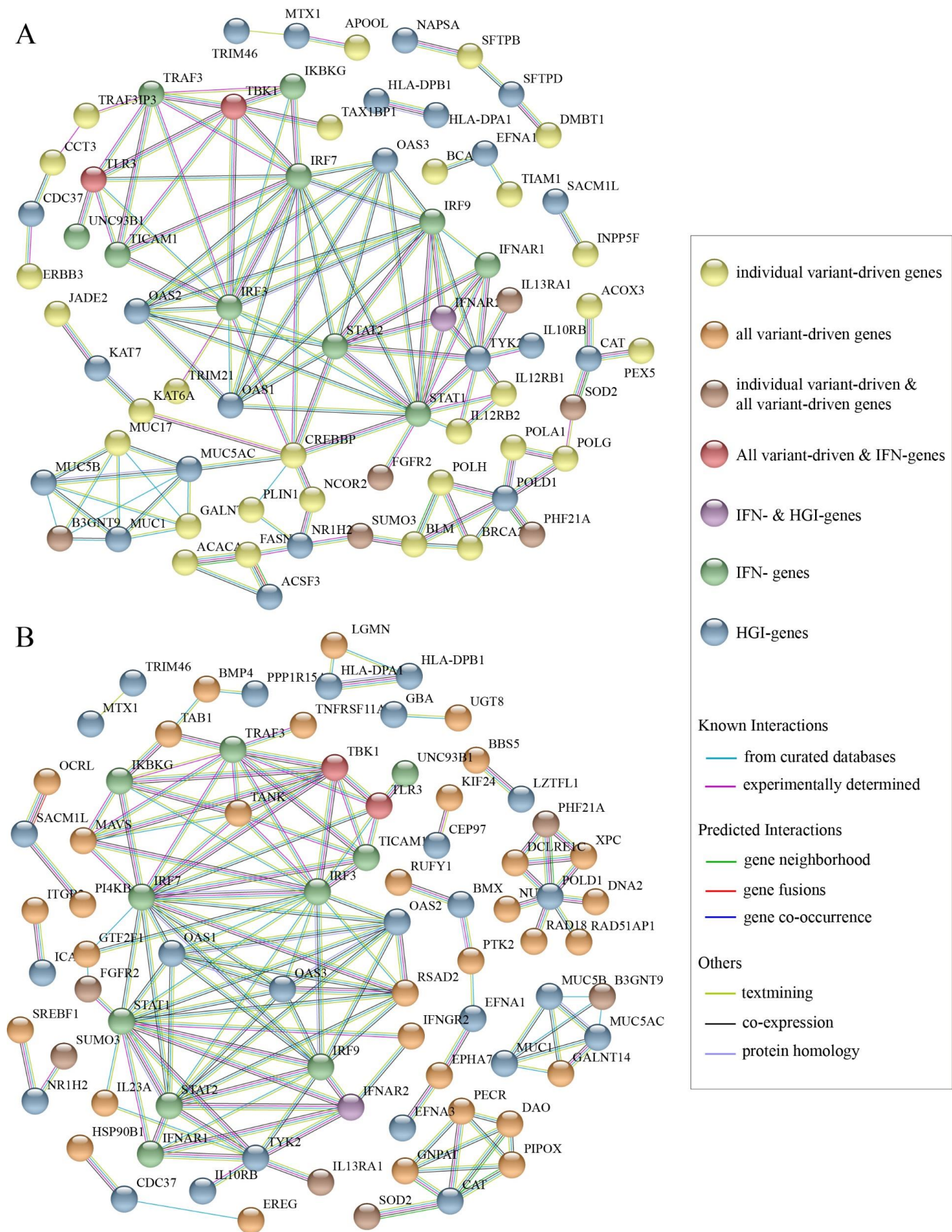


**Figure 1. The flow diagram of rare variants analysis.** A total of 32,232,865 variants were identified from the 451 Covid-19 patients with whole genome sequencing. After filtering by VQSR and MAF, 13,934,341 rare variants were annotated by VEP, and 42,730 candidate variants were included.



**Figure 2. Quality estimate of the cohort.** (A). Heterozygote concordance rate versus sequencing depth for 218 array-genotyped individuals. (B). PCA of 159 *non-severe* and 292 *severe* patients.





**Figure 3. The results of protein-protein interaction network analysis.** The plot of PPI network (A) between the 700 “individual variant-driven” genes with the 148 known genes. (B) between the 778 “all variant-driven” genes with the 148 known genes.