Phenomics
https://doi.org/10.1007/s43657-022-00066-x

ARTICLE

Human Genetic Variants Associated with COVID-19 Severity are Enriched in Immune and Epithelium Regulatory Networks

Zhanying Feng[1,2] · Xianwen Ren[3] · Zhana Duren[4] · Yong Wang[1,2,5,6]

Received: 12 January 2022 / Revised: 8 July 2022 / Accepted: 11 July 2022

Abstract

Human genetic variants can influence the severity of symptoms infected with SARS-COV-2. Several genome-wide association studies have identified human genomic risk single nucleotide polymorphisms (SNPs) associated with coronavirus disease 2019 (COVID-19) severity. However, the causal tissues or cell types underlying COVID-19 severity are uncertain. In addition, candidate genes associated with these risk SNPs were investigated based on genomic proximity instead of their functional cellular contexts. Here, we compiled regulatory networks of 77 human contexts and revealed those risk SNPs' enriched cellular contexts and associated risk SNPs with transcription factors, regulatory elements, and target genes. Twenty-one human contexts were identified and grouped into two categories: immune cells and epithelium cells. We further aggregated the regulatory networks of immune cells and epithelium cells. These two aggregated regulatory networks were investigated to reveal their association with risk SNPs' regulation. Two genomic clusters, the chemokine receptors cluster and the oligoadenylate synthetase (OAS) cluster, showed the strongest association with COVID-19 severity, and they had different regulatory programs in immune and epithelium contexts. Our findings were supported by analysis of both SNP array and whole genome sequencing-based genome wide association study (GWAS) summary statistics.

Keywords GWAS of COVID-19 severity · Regulatory network · Immune cells · Epithelium cells

Abbreviations

COVID-19	Coronavirus disease 2019
GWAS	Genome-wide association study
WGS	Whole-genome sequencing
SNP	Single nucleotide polymorphisms
RE	Regulatory elements
TF	Transcription factor
TG	Target genes
FE	Fold enrichment score
eQTL	Expression quantitative trait loci
GTEx	Genotype-tissue expression (GTEx) program
PCHi-C	Promoter-capture Hi-C

✉ Yong Wang
ywang@amss.ac.cn

1 CEMS, NCMIS, HCMS, MDIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

2 School of Mathematics, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049, China

3 School of Life Sciences, Peking University, Beijing 100871, China

4 Center for Human Genetics and Department of Genetics and Biochemistry, Clemson University, Greenwood, SC 29646, USA

5 Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China

6 Key Laboratory of Systems Biology, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Hangzhou 330106, China

Introduction

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus. Most people infected with the virus will experience mild-to-moderate respiratory illness and recover without special treatment (World Health Organization 2022). However, some people will become seriously ill and require medical attention. Scientists from all over the world have made great efforts to understand the genetic mechanism of COVID-19 severity, which may lead to efficient prevention stratagem and effective cures. Genome-wide association analysis (GWAS) has been a routine option to find genetic variants

Published online: 13 August 2022

associated with phenotypes. Recently, the application of GWAS to COVID-19 phenotypes, such as infection and severe respiratory symptoms, has helped us identify some risk loci associated with COVID-19. For example, the "The Severe COVID-19 GWAS Group" conducted a GWAS involving 1980 patients with severe COVID-19 symptoms and detected single nucleotide polymorphisms (SNP) rs11385942 at locus 3p21.31 and rs657152 at locus 9q34.2 (Severe Covid 2020). The GenOMICC (Genetics of Mortality in Critical Care) group used GWAS to study 2244 critically ill COVID-19 patients from 208 UK intensive care units and found rs10735079 on 12q24.13 in a gene cluster encoding antiviral restriction enzyme activators (2'-5'-Oligoadenylate Synthetase 1/2/3, *OAS1/2/3*), rs2109069 on 19p13.2 near the gene Tyrosine Kinase 2 (*TYK2*), rs2109069 on 19p13.3 within Dipeptidyl Peptidase 9 (*DPP9*), and rs2236757 on 21q22.1 in the interferon receptor gene Interferon Alpha And Beta Receptor Subunit 2 (*IFNAR2*) (Pairo-Castineira et al. 2021). "The COVID-19 Host Genetics Initiative" conducted several GWAS, including "very severe respiratory confirmed COVID" and "hospitalized COVID", which also found many potential loci (The 2020). These identified human risk loci allowed us to further understand the underlying mechanism of COVID-19 severity.

However, interpreting these COVID-19-associated human genetic variants remains challenging since most genetic variants are located in the non-coding regulatory regions with high linkage disequilibrium (Claussnitzer et al. 2015; Kumar et al. 2012; Smemo et al. 2014) and are functional in time- and space-specific contexts (Andersson et al. 2014). Those discoveries imply that we must elucidate these genetic variants' functions in the proper cellular contexts. Several works have been done to uncover the regulatory mechanism of genetic variants. For example, FUMA (Watanabe et al. 2017) and GREAT (McLean et al. 2010) used the routine way to link SNPs to the nearby genes. SMR (Zhu et al. 2016) and Sherlock (He et al. 2013) utilized the expression information to find causal SNP affecting gene regulation. However, these methods did not consider the effect of SNPs on either the direct cis-regulation or the whole biological networks. Recently, we developed a method, Paired Expression and Chromatin Accessibility (PECA), to infer regulatory networks with paired expression and chromatin accessibility data from diverse cellular contexts (Duren et al. 2017) and then, extended it to PECA2 to reconstruct context-specific regulatory network for a single sample (Duren et al. 2020). The reconstructed regulatory networks have been successfully applied to reveal critical regulations for time-course data (Xin et al. 2020). The availability of the public available paired expression and chromatin accessibility data, such as the Encyclopedia of DNA Elements (ENCODE) and the NIH Roadmap Epigenomics Mapping data (ROADMAP), ensured the diversity in cell types and the germ layer lineages. This allows us to use PECA2 to systematically construct a regulatory network atlas of diverse human contexts, which will serve as a valuable resource for genetic variants interpretation in multi-cellular contexts.

Here, we utilized these 77 regulatory networks to interpret COVID-19-severity-associated SNPs. We found the relevant tissues of COVID-19 severity can be categorized into two main cell types: immune cells and epithelium cells. Then, in these two cell type categories, we illustrated the detailed SNP-associated regulations: SNPs that were located in regulatory elements (REs) influenced the binding of upstream transcription factors (TFs) and the expression of downstream target genes (TGs). We found that two gene clusters (chemokine receptors and OAS cluster) showed different regulation patterns in two cell type categories. Our COVID-19-severity-associated regulatory networks will be promising to serve as a valuable perspective for studying COVID-19.

Methods

Construction of Regulatory Network from Paired Expression and Chromatin Accessibility Data

We utilized PECA2 to infer genome-wide and context-specific regulatory networks based on gene expression and chromatin accessibility data (Duren et al. 2020). Given paired RNA-seq and ATAC-seq data for a sample, PECA2 hypothesized that TF regulated the downstream TG by binding at REs. The regulatory strength of a TF on a TG was quantified by the trans-regulation score, which was calculated by integrating information from multiple REs that were bound by the TF to regulate the TG. A prior TF-TG correlation across external public data from ENCODE database was included in the trans-regulation score definition to distinguish the TFs sharing the same binding motif (i.e., TFs from the same family).

We collected paired expression and chromatin accessibility data of 76 human tissues or cell lines (Table S1) and applied PECA2 to these data to obtain regulatory networks of 76 human contexts. Furthermore, we recently constructed a high-quality regulatory network of cranial neural crest cells (Feng et al. 2021), and we included it in our analysis to form a regulatory network atlas of 77 human contexts.

Fold Enrichment Score of SNPs in the Given Region Set

Given a group of SNPs and a RE set, we defined the fold enrichment (FE) score as follows:

$$\mathrm{FE} = \frac{P_r/L_r}{P/L}. \quad (1)$$

Here, P_r was the number of SNPs in REs. L_r was the length of the REs. P was the total number of SNPs. L was the genome length. A higher FE score indicated stronger association between SNPs and RE sets.

We calculated the FE score of 542 SNPs of SNP array based GWAS for COVID-19 severity in RE sets of 77 human contexts and used a criterion of FE ≥ 3 to find 21 contexts that were associated with COVID-19 severity. The FE score was then computed with the 1322 SNPs of whole genome sequencing data based GWAS. Under the same criterion, 18 contexts were found to be relevant to the severity of COVID-19.

Extraction of SNP-Associated Regulatory Network in Given Contexts and Two Categories

Given a context, we checked every RE in its regulatory network. If there was at least one SNP that was located in this RE, we extracted this RE and its upstream TFs and downstream TGs. Then, we added SNPs to the regulatory network and linked the SNPs in RE and this RE with edges. The SNPs, REs, TFs, and TGs formed a SNP-associated regulatory network in this context.

For ten tissues of immune cells, we first aggregated the SNPs, REs, TFs, and TGs. We assigned an edge to SNP-RE if this regulatory pair existed in at least one immune cell type's SNP-associated regulatory network, assigned an edge to TF-RE if this regulatory pair existed in at least one immune cell type's SNP-associated regulatory network, and assigned an edge to RE-TG if this regulatory pair existed in at least one immune cell type's SNP-associated regulatory network. After the node union and edge assignment, we constructed the SNP-associated regulatory network of immune cells. The SNP-associated regulatory network of 11 epithelium cell types was constructed with the same procedure.

Results

Human Genetic Variants of COVID-19 Severity are Enriched in Immune and Epithelium Cells

We first collected paired expression and chromatin accessibility data from 76 human contexts and constructed a human regulatory network atlas. The samples covered all three germ layers, such as "Frontal cortex" (ectoderm), "Primary T cells" (mesoderm), and "Upper lobe of left lung" (endoderm). With each sample's paired expression and chromatin accessibility data as input, we used the PECA2 model (Duren et al. 2020) to construct a regulatory network ("Methods"). The basic unit of the regulatory network is the TF-REs-TG triplet, and each triplet denotes that a TF binds on the REs and regulates a nearby TG. We also included the regulatory network of cranial neural crest cells (Feng et al. 2021), a migratory cell population in early human craniofacial development, into our regulatory network atlas of 77 human contexts. Then, we applied this regulatory network atlas to interpret genetic variants of COVID-19 severity.

We fetched the 542 SNPs with significant associations ($p \leq 5 \times 10^{-8}$) with "very severe respiratory confirmed COVID" ("A2_ALL" GWAS study in "The COVID-19 Host Genetics Initiatives"). To evaluate the relevance between these SNPs and 77 human contexts, the fold enrichment (FE) scores of these 542 SNPs in RE sets of 77 human contexts were calculated ("Methods"). We set the FE score threshold to be 3.0 and found that 21 contexts were relevant to COVID-19 severity, such as "Primary monocytes" and "Upper lobe of left lung" (Fig. 1a). Furthermore, these 21 contexts could be classified into two categories. The first category was immune cells, including ten cell types: "Primary monocytes", "Primary B cells", "Jurkat", "GM12878", "Primary natural killer cells", "CD4 primary cells", "Fetal thymus", "Primary T cells", "CD8 primary Cells", and "Hematopoietic multipotent progenitor cell". The second category was epithelium cells and consisted of 11 cell types: "Vagina", "Omental fat pad", "Gastric", "Upper lobe of left lung", "Lower leg skin", "Esophagus squamous epithelium", "Esophagus muscularis mucosa", "Fetal spleen", "Subcutaneous adipose tissue", "Spleen", and "T47D".

We further extracted sub-network that was associated with SNPs of COVID-19 severity in every relevant context ("Methods"). For example, "Primary monocytes" was one of the immune cell types. And 12 SNPs were located in the eight REs in the regulatory network of "Primary monocytes" (Fig. 1b). These REs were predicted to regulate nine TGs, such as C–C Motif Chemokine Receptor 1 (*CCR1*) and FYVE And Coiled-Coil Domain Autophagy Adaptor 1 (*FYCO1*). *CCR1* played a key role in T-cell-mediated respiratory inflammation (Schaller et al. 2008). These REs were bound by immunity-associated TFs, such as CCAAT Enhancer Binding Protein Beta/Delta (*CEBPB*/*D*) (Chinery et al. 1997; Kinoshita et al. 1992; Pless et al. 2008; Roy et al. 2002) and Fli-1 Proto-Oncogene (*FLI1*) (Wang et al. 2020). On the other hand, "Upper lobe of left lung" was one of the epithelium cell types (Fig. 1c), and eight REs were associated with 14 SNPs of COVID-19 severity and regulated five TGs, such as *CCR1* and *OAS1*. Polymorphisms of *OAS1* have been reported to affect susceptibility to a variety of viral diseases (Burgner et al. 2006; Noguchi et al. 2013). TFs

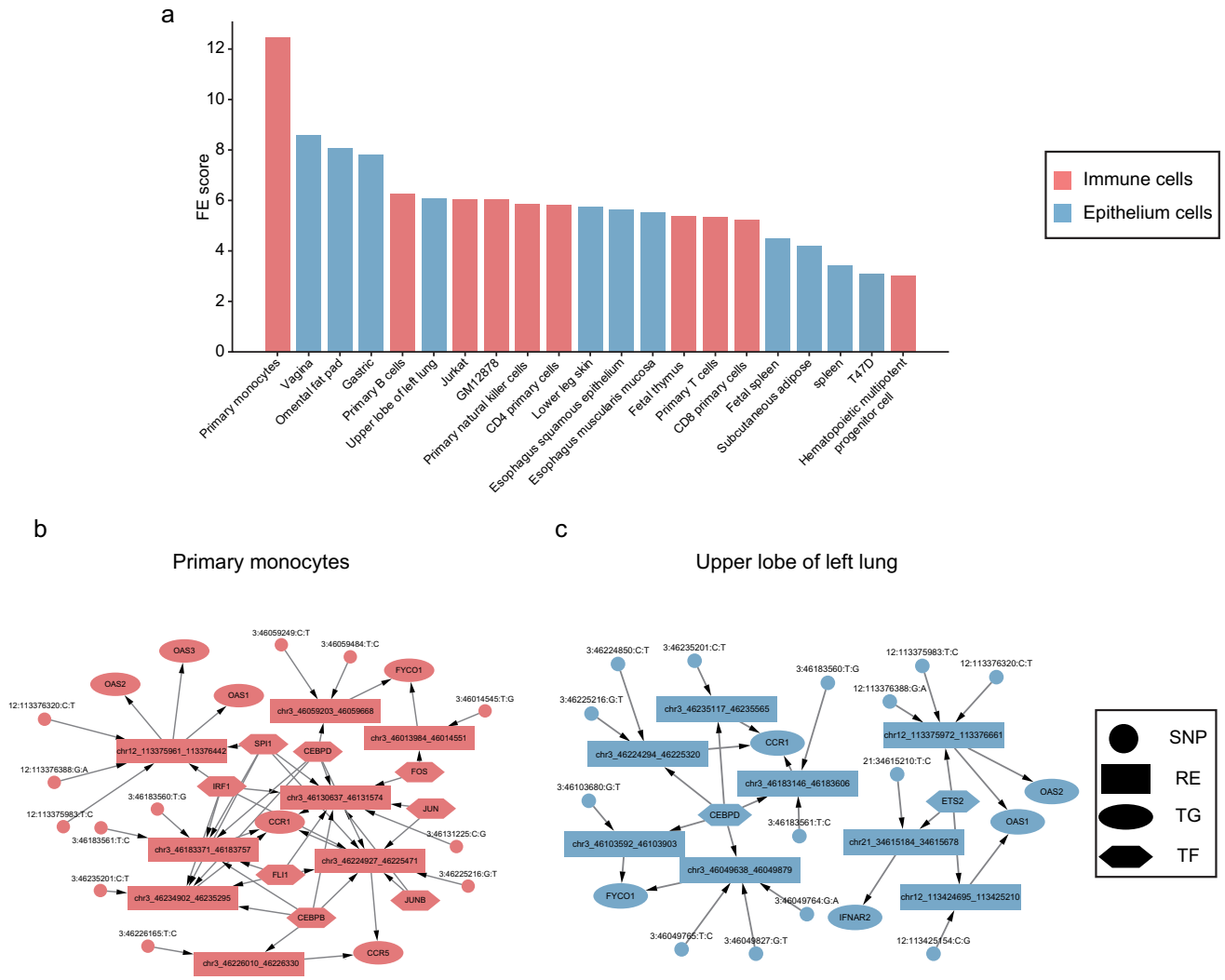

Fig. 1 Human genetic variants associated with COVID-19 severity are enriched in immune and epithelium cells. **a** 21 COVID-19 severity relevant contexts ranked by fold enrichment (FE) scores. Red: immune cells. Blue: epithelium cells. **b** COVID-19 SNP-associated regulatory network of "Primary monocytes", which is an immune cell type. **c** COVID-19 SNP-associated regulatory network of "Upper lobe of left lung", which is an epithelium cell type

regulating these REs were linked to lung epithelium, such as *CEBPD* (Cassel and Nord 2003) and ETS Proto-Oncogene 2 (*ETS2*) (Kabbout et al. 2013).

In summary, we built a regulatory network atlas of 77 human contexts and leveraged their regulatory elements to reveal relevant contexts to human genetic variants of COVID-19 severity. Two categories of cell types were found: the first was epithelium cells which might be involved with susceptibility to viral diseases; the other category was immune cells and possibly linked to severity after being infected with disease. The relevance to two categories of cell types was reproduced by a more recent GWAS with a larger cohort from "The COVID-19 Host Genetics Initiatives" (Fig. S1).

Construction of SNP-Associated Regulatory Networks of Immune and Epithelium Cells

To further understand SNPs' functions in two categories of cell types, we constructed COVID-19-severity-associated regulatory networks in immune cells and epithelium cells, respectively. Briefly, we aggregated the SNPs-associated networks of ten immune cell types into a COVID-19-severity-associated regulatory network of immune cells. Similarly, we aggregated the SNPs-associated regulatory networks of 11 epithelium cell types into a COVID-19-severity-associated regulatory network of epithelium cells ("Methods").

In the COVID-19-severity-associated regulatory network in immune cells, there were 17 TFs, 25 REs, and 15 TGs that were associated with 38 SNPs (Fig. 2a). This sub-regulatory

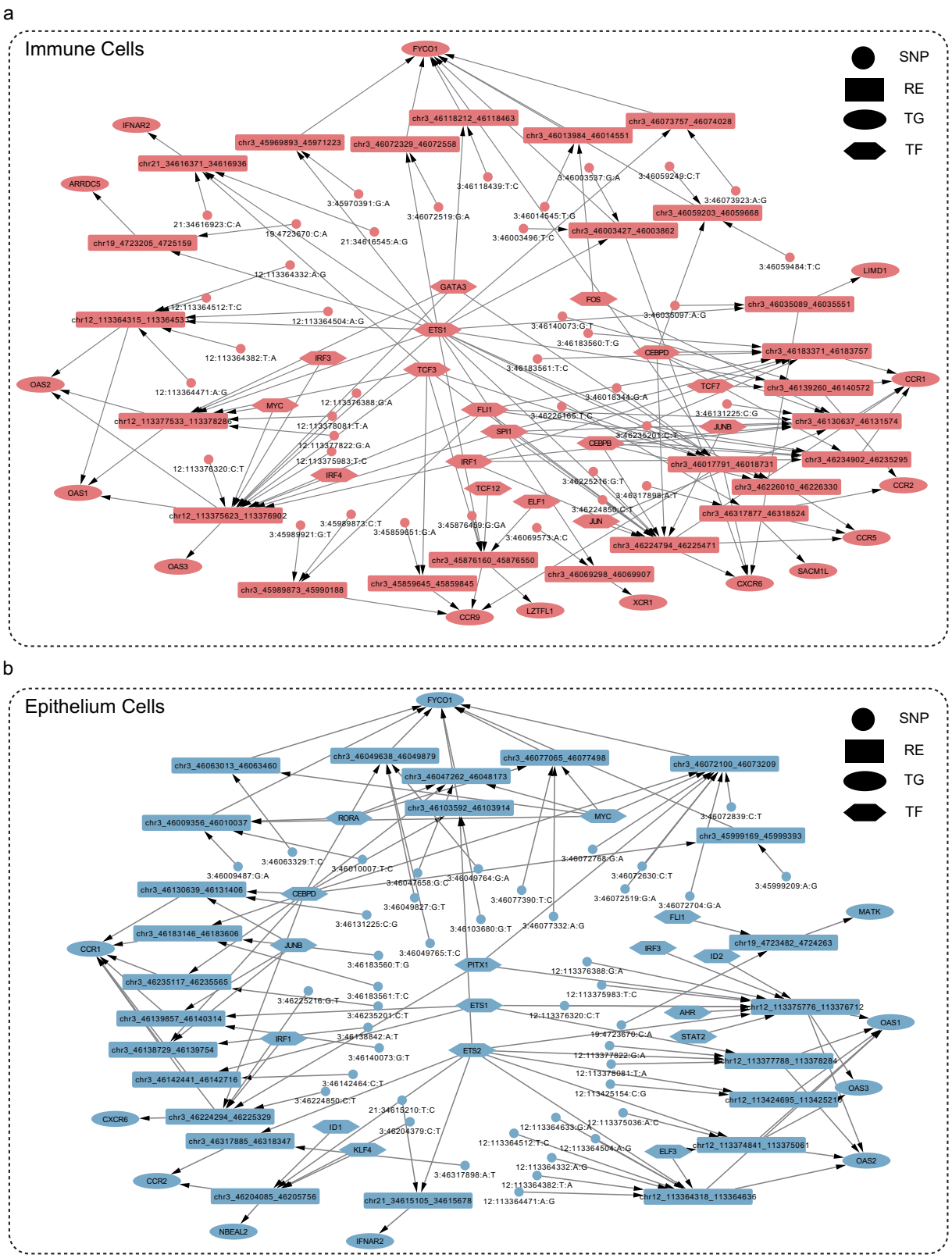

Fig. 2 COVID-19 severity-associated regulatory networks. **a** COVID-19 severity-associated regulatory network of immune cells. **b** COVID-19 severity-associated regulatory network of epithelium cells

network was highly associated with immune functions. For example, Transcription Factor 3/7/12 (*TCF3/7/12*) were important upstream TFs in the regulatory network of immune cells and played vital roles during T-cell development (Chen et al. 2019). Interferon Regulatory Factor 1/3/4 (*IRF1/3/4*) were critical TFs in the cellular differentiation of hematopoietic cells and the regulation of gene expression in response to pathogen-derived danger signals (Tamura et al. 2008). C–C Motif Chemokine Receptor 5 (*CCR5*) was one of the downstream TGs in this regulatory network, and it encoded a protein on the surface of white blood cells that are involved in the immune system (Jiao et al. 2019). We found that two TGs (*CCR5* and C-X-C Motif Chemokine Receptor 6 (*CXCR6*)) of the 15 immune TGs were also differentially expressed in blood samples of fatal COVID-19 patients (Wargodsky et al. 2022) ($p \leq 0.014$), which proved that our reconstructed regulatory networks of immune cells were significantly associated with severity of COVID-19. In the COVID-19 severity-associated regulatory network in epithelium cells, there were 16 TFs, 24 REs, and 10 TGs that were associated with 42 SNPs (Fig. 2b). This SNP-associated regulatory network was involved with the functions of epithelium cells. For example, *ETS2*, a core TF in the regulatory network of epithelium cells, could promote epithelial-to-mesenchymal transition in renal fibrosis (Kabbout et al. 2013). E74 Like ETS Transcription Factor 3 (*ELF3*) was also essential for mesenchymal-to-epithelial transition (Sengez et al. 2019). Neurobeachin Like 2 (*NBEAL2*) was a TG in the epithelium regulatory network, and the *Nbeal2*-deficient mice exhibited impaired development of functional granulation tissue due to severely reduced differentiation of myofibroblasts in the model of excisional skin wound repair (Deppermann et al. 2013). The literature showed that our reconstructed SNP-associated regulatory networks were promising to reveal COVID-19 severity-associated regulations in immune and epithelium cells.

We further overlapped our SNP-associated regulatory network with the expression quantitative trait loci (eQTL) dataset. Our SNP-associated regulatory network revealed the link between SNPs and genes by SNP's location in REs. For example, SNP "12:113364332:A:G" was located in RE "chr12:113364318–113364636" and "chr12:113364318–113364636" regulated *OAS1*, which gave the association between SNP "12:113364332:A:G" and TG *OSA1*. In this way, we obtained 91 SNP-TG associations from immune and epithelium regulatory networks. Then, we collected significant eQTL variants-gene association of 49 tissues from GTEx v8. We found that our SNP-TG associations were highly reproducible by eQTL variants-gene pairs. First, there were 78 (85.7%) SNP-TG associations that could be validated as an eQTL variant-gene association in at least one tissue. On the other hand, when we ranked the 49 tissues by their overlapping number of eQTL with our SNP-TG associations, we found that the top-ranked tissues were also immune and epithelium cell types (Fig. 3). For example, "esophagus mucosa", one of the epithelium cell types, ranked first among 49 tissues. Some other tissues were related to epithelium cells, such as "Cells Cultured fibroblasts", "Skin Not Sun Exposed Suprapubic", "Spleen", "Skin Sun Exposed Lower leg", and "Lung". And there were also immune tissues among the top-ranked tissue, such as "Artery Tibial" and "Whole Blood". To check if our SNP-TG associations were causally responsible for COVID-19 severity, we conducted the colocalization analysis with GWAS

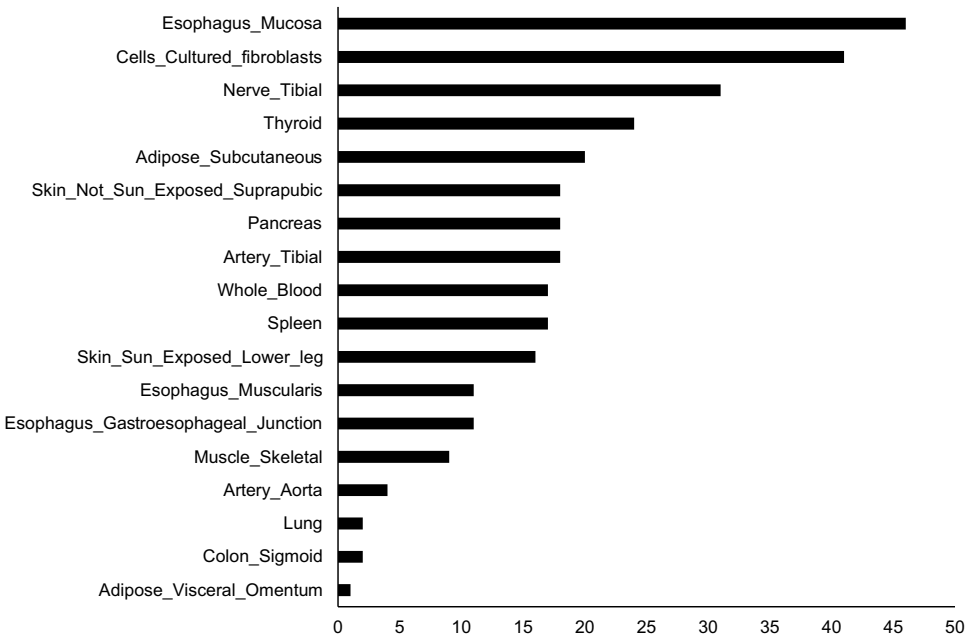

Fig. 3 Overlapping of SNP-TGs inferred from our regulatory networks with eQTL of 49 tissues in GTEx. The tissues are ranked by the number of overlapped eQTL

summary statistics of COVID-19 severity and eQTL from GTEx by SMR (Zhu et al. 2016). SMR gave 57 variant-gene pairs in 49 tissues, and we found that two of our 91 SNP-TG associations were validated by SMR colocalization analysis, which was a significantly overlapping ($p = 1 \times 10^{-7}$). These two SNP-TG associations were related to *OAS3*, revealing that "12:113376320:C:T" and "12:113375983:T:C" might causally regulate the expression of *OAS3* and influence the severity of COVID-19 symptoms. These results showed that our COVID-19-severity-associated regulatory network could be supported by the independent eQTL dataset.

In summary, we have constructed COVID-19-severity-associated regulatory network of immune cells and epithelium cells. These two regulatory networks were tightly linked to immune and epithelium functions and could be validated by eQTL of immune and epithelium tissues in GTEx.

Associating Regulatory Networks of Two Cell Type Categories Revealed Shared Regulatory Structure

After reconstructing and validating the regulatory networks of immune and epithelium cells, we next compared these two cell types to find their conservation and divergence. We first compared the TFs, TGs, REs, and SNPs in the regulatory networks of immune and epithelium cells. For TFs, seven TFs were shared by two cell types, such as ETS Proto-Oncogene 1 (*ETS1*) and *IRF1*. There were ten immune-specific TFs [such as *TCF7* and E74 Like ETS Transcription Factor 1 (*ELF1*)] and nine epithelium-specific TFs [such as *ETS2* and Kruppel Like Factor 4 (*KLF4*)]. There were eight overlapped TGs (such as *CCR1* and *FYCO1*), seven immune-specific TGs [such as *CCR5* and Leucine Zipper Transcription Factor Like 1 (*LZTFL1*)], and two epithelium-specific TGs [Megakaryocyte-Associated Tyrosine Kinase (*MATK*) and *NBEAL2*]. For SNPs and REs, about half of them in the immune regulatory network were shared by the epithelium regulatory network (Fig. 4).

We found that the regulatory network could be clustered into four clusters according to their genomic location and all these four clusters were shared between immune and epithelium cells (Fig. 2a, b). The first cluster was on chromosome 3. This cluster was involved with the regulation of C–C Motif Chemokine Receptor 1/2/5/9, (*CCR1/2/5/9*), *CXCR6*, *FYCO1*, LIM Domain Containing 1 (*LIMD1*), *LZTFL1*, SAC1 Like Phosphatidylinositide Phosphatase (*SACM1L*), X-C Motif Chemokine Receptor 1 (*XCR1*) in immune cells and *CCR1/2/5/9*, *CXCR6*, *NBEAL2* in epithelium cells. The second cluster was in chromosome 12 and involved with the regulation of *OAS1/2/3* in both immune and epithelium cells. The third cluster was in chromosome 19. In immune cells, this cluster was associated with Arrestin domain containing 5 (*ARRDC5*); and in epithelium cells, it was associated with *MATK*. The last cluster was in chromosome 21 and linked to *IFNAR2* in two cell type categories. These four genetic variants associated regulatory clusters might exert genetic influence on the infection and severity of COVID-19.

In the regulatory network of two cell types, we found two classes of upstream TFs (Fig. 2a, b). The first class of TFs regulated most of the four regulatory clusters. For example, in the regulatory network of immune cells, *ETS1*, *TCF3* and GATA Binding Protein 3 (*GATA3*) were involved with at least three regulatory clusters. And for the epithelium cells, *ETS1*, *ETS2*, and Paired Like Homeodomain 1 (*PITX1*) were classified into this class. Contrary to the broad regulation of the first class, the second class of TFs was only responsible for a small part of the regulatory network. For example, in the immune regulatory network,

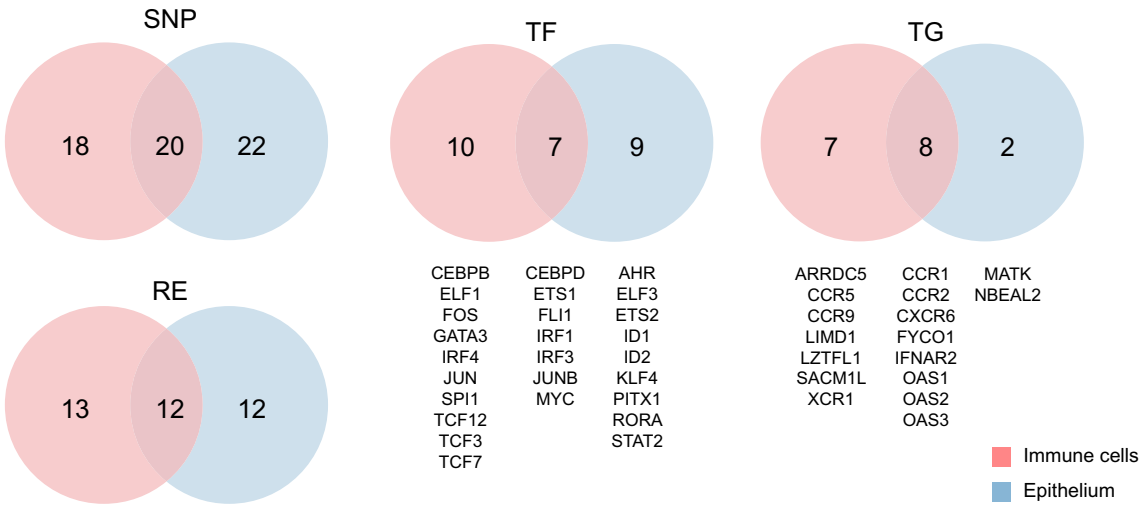

Fig. 4 Overlapping of SNP, RE, TF, TG between immune regulatory network and epithelium regulatory network

CEBPD, *CEBPB*, *TCF7*, Fos Proto-Oncogene (*FOS*), JunB Proto-Oncogene (*JUNB*), Jun Proto-Oncogene (*JUN*), and *ELF1* only regulated REs in the chromosome 3 regulatory cluster. And *IRF3*, *IRF4*, and MYC Proto-Oncogene (*MYC*) were only responsible for regulations in the chromosome 12 regulatory cluster. In the epithelium regulatory network, *CEBPD*, *JUNB*, *IRF1*, Inhibitor of DNA Binding 1 (*ID1*), RAR Related Orphan Receptor A (*RORA)*, *MYC*, and *KLF4* were only involved with the regulation of the chromosome 3 regulatory cluster. Inhibitor of DNA Binding 2 (*ID2*), *IRF3*, Aryl Hydrocarbon Receptor (*AHR*), Signal Transducer and Activator of Transcription 2 (*STAT2*), and *ELF3* were associated with the chromosome 12 regulatory cluster.

Through comparing regulatory networks of two cell types, we found these two cell types shared many TFs, TGs, REs, and SNPs. The regulatory structure (four regulatory clusters and two classes of upstream TFs) was conserved in two cell types. And there were also many cell type-specific TFs, TGs, REs, and SNPs.

Chromosome 3 and 12 Clusters Showed Distinct Regulatory Programs in Immune and Epithelium Cells

From the above analysis of regulatory structure, we noticed that two relatively denser regulatory clusters: chromosome 3 and chromosome 12 clusters, indicating that they might play an essential role in COVID-19 severity. Here, we focused on these two regulatory clusters.

For the regulatory cluster in chromosome 12, these regulatory programs were mainly involved with the *OAS* cluster (*OAS1/2/3*) in 10 immune and nine epithelium cell types. This *OAS* cluster was strongly associated with immune and epithelium functions. For example, *OAS1* polymorphisms played potential roles in respiratory infection from human bronchial epithelial cells (Noguchi et al. 2013). The *OAS2* protein was a well-known innate immune activated antiviral enzyme catalyzing the synthesis of 2′-5′-oligoadenylate for RNase L activation and inhibition of viral propagation (Mozzi et al. 2015). We found that the REs of this cluster were mainly distributed around the promoter of *OAS3* and there were many COVID-19 severity-associated SNPs within these REs (Fig. 5a), which was consistent among immune and epithelium cell types. To add additional evidence of regulation in this area, we collected previously published promoter-capture Hi-C (PCHi-C) data of primary blood cell types (Javierre et al. 2016). We used the PCHi-C interactions that have interaction score ≥ 5 in at least one analyzed cell type and found a loop between the promoters of *OAS3* and *OAS2*. This observation induced a hypothesis that the REs around the promoter of *OAS3* were cis-regulatory elements of *OAS2* in immune cells. To validate this hypothesis, we first checked the REs regulating the three OAS genes in ten immune cell types. We found that all 15 REs were predicted to regulate *OAS2*, 13 REs were predicted to regulate *OAS1*, and only seven REs were predicted to regulate *OAS3*, which supported that REs in this locus regulated *OAS2*. Then, we checked the expression of three *OAS* genes and found that *OAS2*'s expression was the highest in immune cell types (Fig. 5b). We also computed the averaged correlation between the openness score of REs and the expression of three OAS genes across 148 samples of human (Table S2). The result showed that the accessibility of these REs was more correlated to the expression of *OAS2* (Fig. 5c). In summary, the evidence of the PCHi-C loop, regulatory network, expression data, and RE-TG correlation indicated that in immune cells, the REs around *OAS3*'s promoter were more likely to regulate *OAS2*.

The regulatory program of cluster in chromosome 3 were mainly involved with chemokine receptors (*CCR1/2/5/9*, *CXCR6*, and *XCR1*) in eight immune cell types and eight epithelium cell types. In literature, there were many reports that these chemokine receptors were associated with the function of both immune and epithelium cells (Jiao et al. 2019; Schaller et al. 2008). We found five main regulatory areas associated with SNPs and regulated these chemokine receptors (we named them LOC1-LOC5, Fig. 6a). Some of these REs' regulations were shared by two cell types. For example, LOC3 was predicted to regulate *CCR1* in both immune and epithelium cells. And there was also a PCHi-C loop between LOC3 and *CCR1*. The other LOCs' regulatory programs were different. For example, LOC1 and LOC2 were immune-specific REs and regulated *CCR9* and *CXCR6*. While REs in LOC4 were shared by two cell types, they were predicted to regulate *CCR5* in immune cells but regulate *CCR1* in epithelium cells. LOC4 and *CCR5* were also contacted by a PCHi-C loop in blood cells. LOC5 regulated *CCR2/5* in immune cells but only regulated *CCR2* in the epithelium. The expression of these chemokine receptors supported the regulations above: *CCR1/2/5/9* and *CXCR6* have relatively higher expression in immune cells but only *CCR1* and *CCR2* were expressed in epithelium cells (Fig. 6b). Then, we checked the RE-TG correlation and found LOC1 showed the highest correlation with *CCR9*, LOC2 was more correlated with *CXCR6* and *CCR9*, LOC3 was most associated with *CCR1*, LOC4 showed higher correlation with *CCR1* and *CCR5*, and LOC5 was correlated with *CCR5* (Fig. 6c), which were consistent with the regulations revealed above.

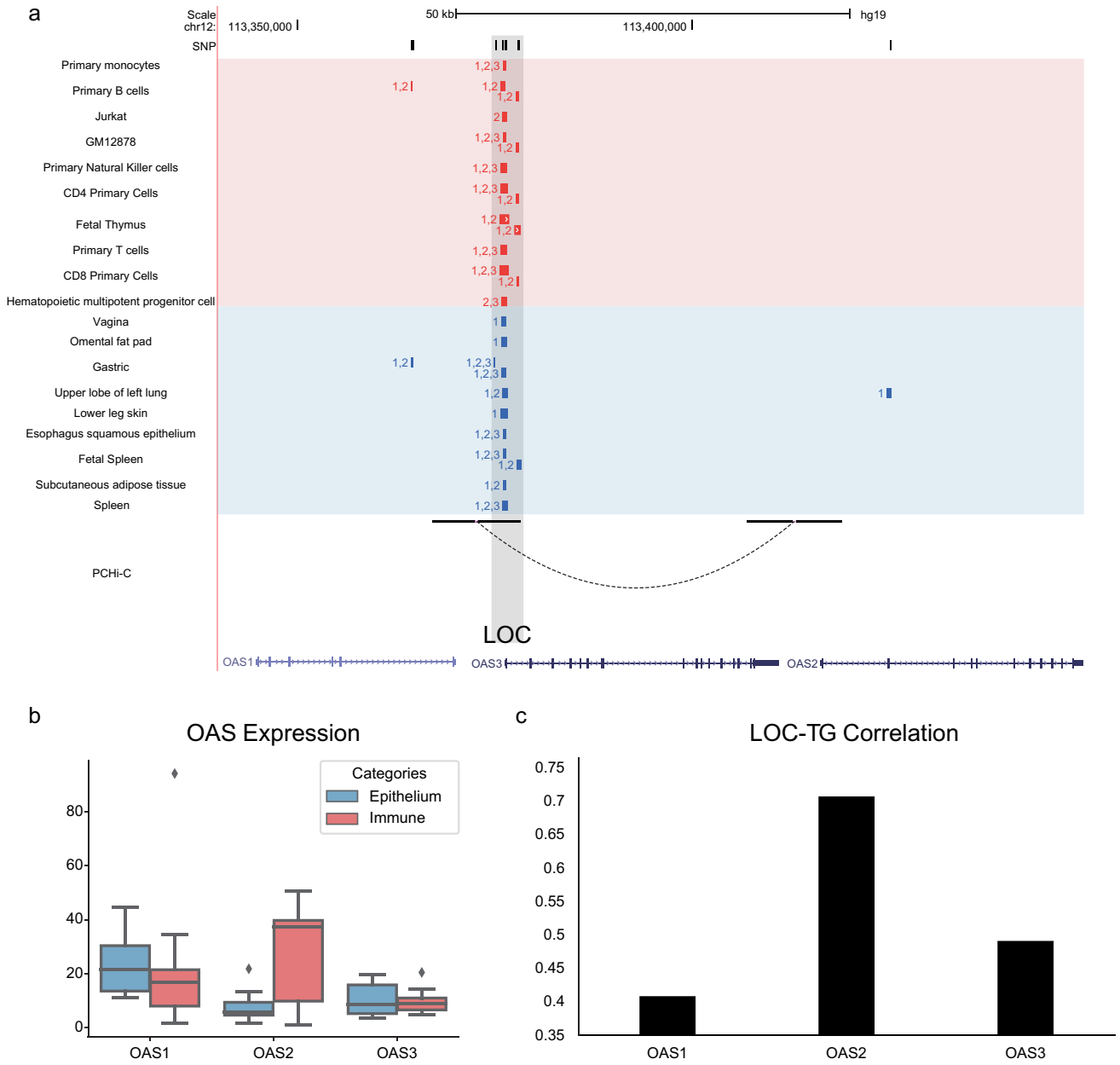

Fig. 5 Regulatory network of OAS gene cluster. **a** SNPs, REs, and PCHi-C loop associated with *OAS1/2/3*. The RE tracks consisted of 10 immune cell types and nine epithelium cell types. PCHi-C loops were from blood samples. The number labels around the RE were indicators of RE's target genes (1 means OAS1, 2 means OAS2, and 3 means OAS3). **b** Expression of *OAS1/2/3* in immune and epithelium cells. **c** Correlations between the accessibility of REs and expression of *OAS1/2/3*

Taken together, by detailed study of two main COVID-19 severity-associated regulatory clusters on chromosomes 3 and 12, we found some shared and distinct regulatory programs in immune and epithelium. These conserved and divergent regulations might give us biological insights into how virus infected human through epithelium cells and how human defensed SARS-COV-2 with our immune system.

Analyzing Whole Genome Sequencing-Based GWAS Summary Statistics Reproduced Relevant Immune and Epithelium Cells and Refined the Regulatory Networks

Microarray genotyping-based GWAS have been the most common approach for identifying disease-associated

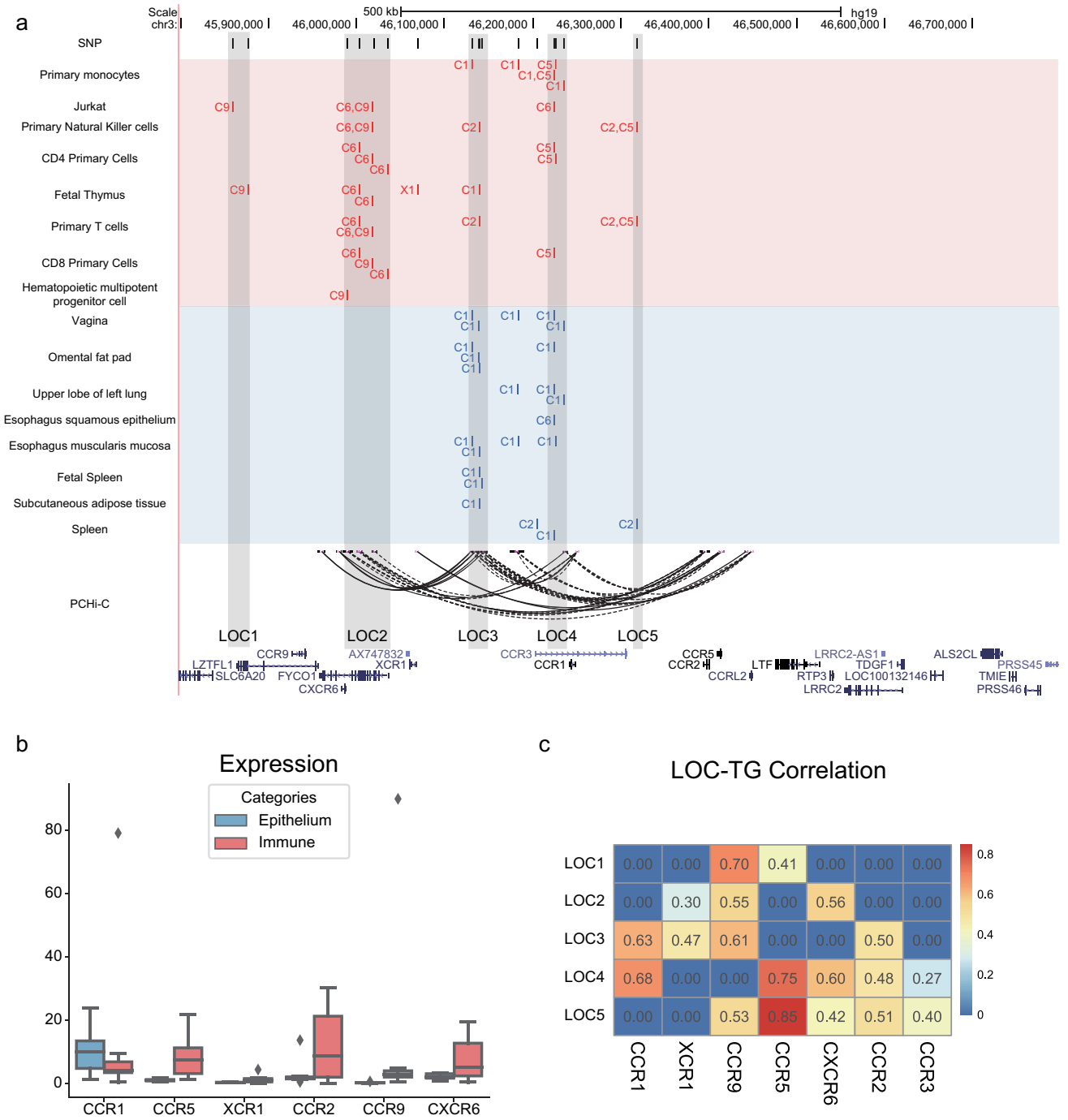

Fig. 6 Regulatory network of chemokine receptors cluster. **a** SNPs, REs, and PCHi-C loop associated with chemokine receptors. The RE tracks consisted of 11 immune cell types and nine epithelium cell types. PCHi-C loops were from blood samples. The labels around REs were indicators of TGs: C1-CCR1, C2-CCR2, C5-CCR5, C6-CXCR6, C9-CCR9, and X1-XCR1. **b** Expression of chemokine receptors in immune and epithelium cells. **c** Correlations between the accessibility of REs and expression of chemokine receptors

genetic variants across the whole genome and can interrogate over four million SNP markers per sample (Uffelmann et al. 2021). Whole-genome sequencing (WGS) provides a comprehensive base-by-base method for interrogating the 3.2 billion bases of the human genome, which can give higher-quality GWAS summary statistics. WGS-based GWAS analysis holds the promise to reveal richer and more accurate genetic variants. We noticed that Kousathanas et al. used WGS of a larger cohort to conduct GWAS analysis and found 1322 SNPs with significant association ($p \leq 5 \times 10^{-8}$)

with the severity of COVID-19 (Kousathanas et al. 2022). We then used this independent data to validate and refine our relevant contexts and regulatory networks.

We first performed the SNP fold enrichment analysis on these 1322 significant SNPs to find relevant contexts. 18 contexts obtained FE scores greater than 3.0 and were also grouped into two categories: immune cells such as "Primary monocytes", "CD4 primary cells" and epithelium cells such as "Upper lobe of left lung". These results faithfully reproduced our results in microarray genotyping-based GWAS data that the genetic variants of COVID-19 severity were enriched in immune and epithelium cells. Furthermore, this WGS-based GWAS also revealed that immune cells were more associated with the severity of COVID-19 than epithelium cells (Fig. 7a). Next, we sought to refine our regulatory networks of immune and epithelium cells with the overlapped SNPs of the WGS-based GWAS and microarray-based GWAS. We found that among the 38 SNPs in the previous reconstructed regulatory network of immune cells, 25 SNPs were reproducible in the WGS-based GWAS. These 25 SNPs formed a refined regulatory network of immune cells, including 12 TGs such as *CCR1/2/5/9*, 14 TFs such as *FLI1*, *JUNB*, *CEBPD*, and 20 REs (Fig. 7b). And for epithelium cells, we found that among the 42 SNPs in the above reconstructed regulatory network, 22 SNPs were also significant in the WGS-based GWAS. These 22 SNPs gave a refined regulatory network of epithelium cells, which include six TGs such as *CCR1/2*, *FYCO1*, 14 TFs such as *ETS1/2*, *KLF4*, and 16 REs (Fig. 7c).

In summary, we conducted regulatory network analysis with a recently published WGS-based GWAS and validated our findings that human genetic variants of COVID-19 severity were enriched in immune and epithelium cells. We also used the reproducible SNPs to obtain refined regulatory networks of immune and epithelium cells.

Discussion

The COVID-19 pandemic has influenced human life all over the world. Scientists made their efforts to gain more understanding of the pathology of COVID-19 and to find effective cures. GWAS is a powerful tool to discover suspicious loci for interesting traits. By profiling the phenotype and genotype of coronavirus infected people, the "COVID-19 host genetics initiative" has conducted GWAS analysis with increased power (Pairo-Castineira et al. 2021). Through these GWAS, many risk loci that may be associated with COVID-19 infection and severity were found. Following these discoveries, we used regulatory network atlas to interpret these risk loci and sought to gain biological insights into COVID-19.

First, we found that COVID-19 severity-associated SNPs were mainly enriched in two categories of human contexts: immune cells and epithelium cells. We used another GWAS of COVID-19 severity with a larger cohort (https://www.covid19hg.org/results/r6/) to show the association between immune cells and epithelium cells was quite reproducible (Fig. S1). We then reconstructed the COVID-19-severity-associated regulatory network in the two cell type categories. The COVID-19-severity-associated regulatory networks were validated by recently published RNA-seq data of COVID-19 patients with severe symptoms and eQTL from GTEx. And further analysis revealed the regulatory structure of COVID-19-associated regulatory networks: the four regulatory clusters and two classes of upstream TFs. The regulatory structure showed both conservation and divergence between two cell type categories. Then, we focused on two crucial regulatory clusters (OAS cluster and chemokine receptor cluster) and revealed the causal genes and differential regulations between two cell type categories. Finally, analysis based a recent WGS-based GWAS with a larger cohort validated the two COVID-19 severity's relevant cell type categories and refined the associated regulatory networks.

Several factors hindered the post-GWAS analysis of COVID-19 severity. First, one more powerful GWAS with a larger cohort is in need for such a complex trait. In this paper, the GWAS only detected 542 SNPs that were significantly associated with COVID-19's symptoms, which was not comparable to normal phenotypes, such as height or body mass index. Some efforts have been made to improve the power of GWAS. For example, the most recent GWAS conducted by "COVID-19 host genetics initiative" has included more than 18,000 very severe respiratory confirmed cases in the analysis (https://www.covid19hg.org/results/r7/). On the other hand, a more comprehensive regulatory network atlas is promising to understand the COVID-19 severity better. Although our current regulatory network atlas has covered many tissues, it is still far from complete and is limited in cell-type resolution. As the symptoms of COVID-19 are involved with many organs, a more comprehensive regulatory network atlas, even at the cell type level from single cell multi-omics data, will interpret more genetic variants associated with COVID-19.

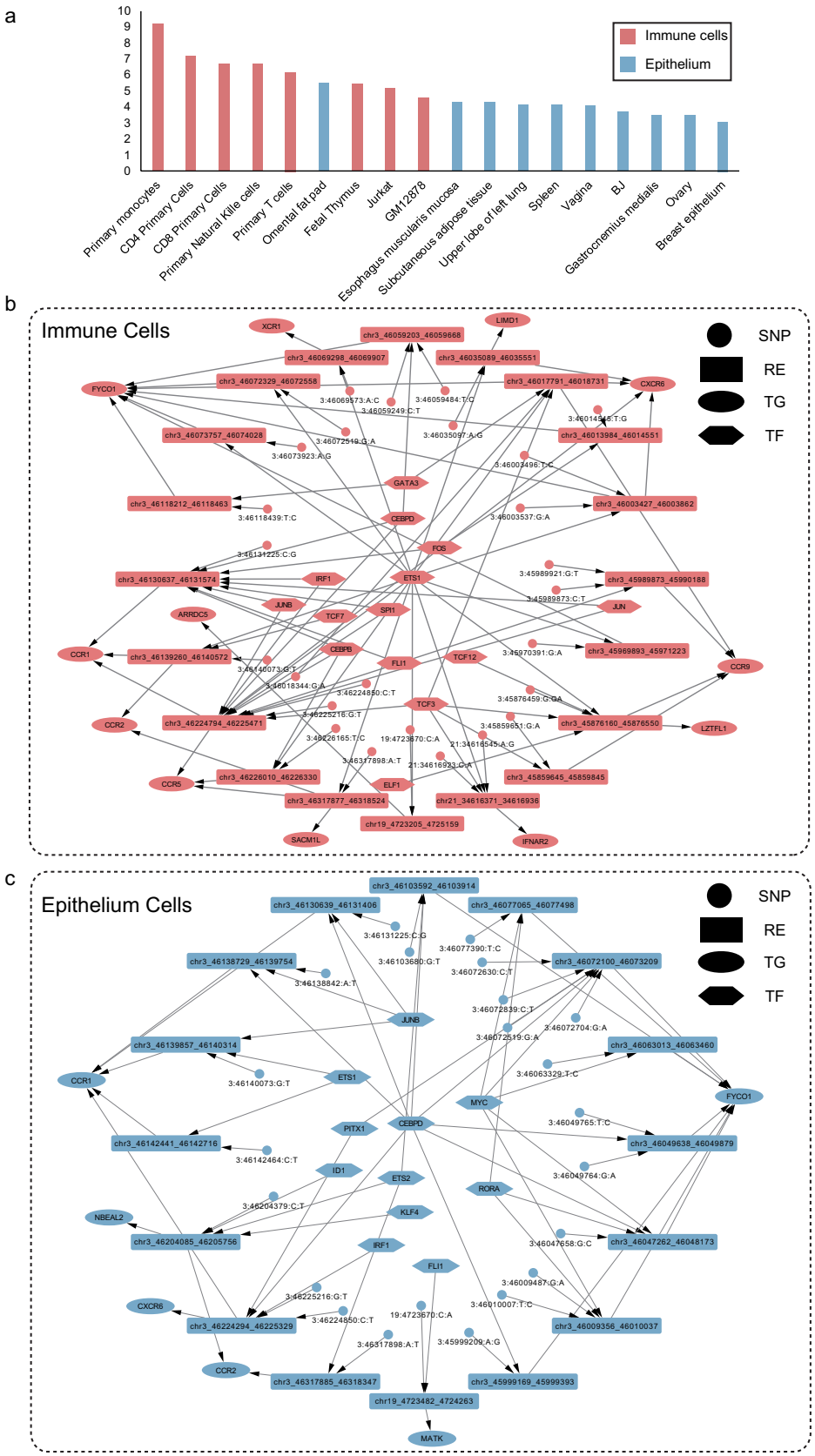

Fig. 7 Whole-genome sequencing-based GWAS reproduces and refines regulatory networks of two cell type categories. **a** Analyzing WGS-based GWAS summary statistics revealed 18 COVID-19 severity relevant contexts, which are ranked by fold enrichment (FE) scores. Red: immune cells. Blue: epithelium cells. **b** Refined COVID-19 severity-associated immune regulatory network. **c** Refined COVID-19 severity-associated epithelium regulatory network

Conclusion

We constructed regulatory networks atlas of 77 human contexts and conducted the post-GWAS analysis to interpret genetic variants of COVID-19 severity. We found SNPs of COVID-19 severity were enriched in 21 human contexts, and these 21 human contexts can be grouped into two categories: immune cells and epithelium cells. We further constructed the COVID-19 severity-associated regulatory networks in immune cells and epithelium cells. The immune-epithelium crosstalk was investigated to show their association with severity-risk SNPs' regulation. Two genomic clusters, the chemokine receptors cluster, and the OAS cluster showed the strongest association with COVID-19 severity and different regulations in immune and epithelium contexts. Our findings were supported by analysis of both SNP array and whole genome sequencing-based GWAS.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43657-022-00066-x.

Acknowledgements We acknowledge funding from the Strategic Priority Research Program of the Chinese Academy of Sciences (XDPB17), National Key Research and Development Program of China (2020YFA0712402), and the National Natural Science Foundation of China (grants 12025107, 11871463, 61621003, 11688101). The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal under dbGaP accession number phs000424.v8.p2 on 12/09/2020.

Authors' Contributions YW conceived and supervised the project. ZF designed the analytical approach and performed numerical experiments and data analysis. XR contributed to the biological interpretation of regulatory networks. ZD contributed to the construction of regulatory networks.

Data Availability The GWAS summary statistics of COVID-19 severity ("A2_ALL" study) was downloaded at https://www.covid19hg.org/results/r4/. The GWAS summary statistics from WGS data was downloaded at https://genomicc.org/data/. The constructed regulatory network atlas was freely available at https://github.com/AMSSwanglab. The eQTL dataset was downloaded at the GTEx portal https://www.gtexportal.org/home/datasets. The collected paired expression and chromatin accessibility data were summarized in Tables S1 and S2.

Code Availability Not applicable.

Declarations

Conflict of Interest The authors declare no competing interests.

Ethical Approval Not applicable.

Consent to Participate Not applicable.

Consent to Publish Not applicable.

References

Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, Ntini E, Arner E, Valen E, Li K, Schwarzfischer L, Glatz D, Raithel J, Lilje B, Rapin N, Bagger FO, Jørgensen M, Andersen PR, Bertin N, Rackham O, Burroughs AM, Baillie JK, Ishizu Y, Shimizu Y, Furuhata E, Maeda S, Negishi Y, Mungall CJ, Meehan TF, Lassmann T, Itoh M, Kawaji H, Kondo N, Kawai J, Lennartsson A, Daub CO, Heutink P, Hume DA, Jensen TH, Suzuki H, Hayashizaki Y, Müller F, Forrest ARR, Carninci P, Rehli M, Sandelin A, The FC (2014) An atlas of active enhancers across human cell types and tissues. Nature 507(7493):455–461. https://doi.org/10.1038/nature12787

Burgner D, Jamieson SE, Blackwell JM (2006) Genetic susceptibility to infectious diseases: big is beautiful, but will bigger be even better? Lancet Infect Dis 6(10):653–663. https://doi.org/10.1016/S1473-3099(06)70601-6

Cassel TN, Nord M (2003) C/EBP transcription factors in the lung epithelium. Am J Physiol Lung Cell Mol Physiol 285(4):L773–L781. https://doi.org/10.1152/ajplung.00023.2003

Chen Z, Ji Z, Ngiow SF, Manne S, Cai Z, Huang AC, Johnson J, Staupe RP, Bengsch B, Xu C, Yu S, Kurachi M, Herati RS, Vella LA, Baxter AE, Wu JE, Khan O, Beltra J-C, Giles JR, Stelekati E, McLane LM, Lau CW, Yang X, Berger SL, Vahedi G, Ji H, Wherry EJ (2019) TCF-1-centered transcriptional network drives an effector versus exhausted CD8 T cell-fate decision. Immunity 51(5):840-855.e5. https://doi.org/10.1016/j.immuni.2019.09.013

Chinery R, Brockman JA, Dransfield DT, Coffey RJ (1997) Antioxidant-induced nuclear translocation of CCAAT/enhancer-binding protein beta. A critical role for protein kinase A-mediated phosphorylation of Ser^{299}. J Biol Chem 272(48):30356–30361. https://doi.org/10.1074/jbc.272.48.30356

Claussnitzer M, Dankel SN, Kim K-H, Quon G, Meuleman W, Haugen C, Glunk V, Sousa IS, Beaudry JL, Puviindran V, Abdennur NA, Liu J, Svensson P-A, Hsu Y-H, Drucker DJ, Mellgren G, Hui C-C, Hauner H, Kellis M (2015) FTO obesity variant circuitry and adipocyte browning in humans. New Engl J Med 373(10):895–907. https://doi.org/10.1056/NEJMoa1502214

Deppermann C, Cherpokova D, Nurden P, Schulz J-N, Thielmann I, Kraft P, Vögtle T, Kleinschnitz C, Dütting S, Krohne G, Eming SA, Nurden AT, Eckes B, Stoll G, Stegner D, Nieswandt B (2013) Gray platelet syndrome and defective thrombo-inflammation in *Nbeal2*-deficient mice. J Clin Invest 123(8):3331–3342. https://doi.org/10.1172/JCI69210

Duren Z, Chen X, Jiang R, Wang Y, Wong WH (2017) Modeling gene regulation from paired expression and chromatin accessibility data. Proc Natl Acad Sci USA 114(25):E4914–E4923. https://doi.org/10.1073/pnas.1704553114

Duren Z, Chen X, Xin J, Wang Y, Wong WH (2020) Time course regulatory analysis based on paired expression and chromatin accessibility data. Genome Res 30(4):622–634. https://doi.org/10.1101/gr.257063.119

Feng Z, Duren Z, Xiong Z, Wang S, Liu F, Wong WH, Wang Y (2021) hReg-CNCC reconstructs a regulatory network in human cranial neural crest cells and annotates variants in a developmental context. Commun Biol 4(1):442. https://doi.org/10.1038/s42003-021-01970-0

He X, Fuller Chris K, Song Y, Meng Q, Zhang B, Yang X, Li H (2013) Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. Am J Hum Genet 92(5):667–680. https://doi.org/10.1016/j.ajhg.2013.03.022

Javierre BM, Burren OS, Wilder SP, Kreuzhuber R, Hill SM, Sewitz S, Cairns J, Wingett SW, Várnai C, Thiecke MJ, Burden F, Farrow S, Cutler AJ, Rehnström K, Downes K, Grassi L, Kostadima M,

Freire-Pritchett P, Wang F, Consortium B, Stunnenberg HG, Todd JA, Zerbino DR, Stegle O, Ouwehand WH, Frontini M, Wallace C, Spivakov M, Fraser P (2016) Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. Cell 167(5):1369–1384.E19. https://doi.org/10.1016/j.cell.2016.09.037

Jiao X, Nawab O, Patel T, Kossenkov AV, Halama N, Jaeger D, Pestell RG (2019) Recent advances targeting CCR5 for cancer and its role in immuno-oncology. Cancer Res 79(19):4801–4807. https://doi.org/10.1158/0008-5472.CAN-19-1167

Kabbout M, Garcia MM, Fujimoto J, Liu DD, Woods D, Chow C-W, Mendoza G, Momin AA, James BP, Solis L, Behrens C, Lee JJ, Wistuba II, Kadara H (2013) ETS2 mediated tumor suppressive function and MET oncogene inhibition in human non-small cell lung cancer. Clin Cancer Res 19(13):3383–3395. https://doi.org/10.1158/1078-0432.CCR-13-0341

Kinoshita S, Akira S, Kishimoto T (1992) A member of the C/EBP family, NF-IL6 beta, forms a heterodimer and transcriptionally synergizes with NF-IL6. Proc Natl Acad Sci USA 89(4):1473–1476. https://doi.org/10.1073/pnas.89.4.1473

Kousathanas A, Pairo-Castineira E, Rawlik K, Stuckey A, Odhams CA, Walker S, Russell CD, Malinauskas T, Wu Y, Millar J, Shen X, Elliott KS, Griffiths F, Oosthuyzen W, Morrice K, Keating S, Wang B, Rhodes D, Klaric L, Zechner M, Parkinson N, Siddiq A, Goddard P, Donovan S, Maslove D, Nichol A, Semple MG, Zainy T, Maleady-Crowe F, Todd L, Salehi S, Knight J, Elgar G, Chan G, Arumugam P, Patch C, Rendon A, Bentley D, Kingsley C, Kosmicki JA, Horowitz JE, Baras A, Abecasis GR, Ferreira MAR, Justice A, Mirshahi T, Oetjens M, Rader DJ, Ritchie MD, Verma A, Fowler TA, Shankar-Hari M, Summers C, Hinds C, Horby P, Ling L, McAuley D, Montgomery H, Openshaw PJM, Elliott P, Walsh T, Tenesa A, Fawkes A, Murphy L, Rowan K, Ponting CP, Vitart V, Wilson JF, Yang J, Bretherick AD, Scott RH, Hendry SC, Moutsianas L, Law A, Caulfield MJ, Baillie JK, GenOMICC Investigators, 23and Me Investigators, Covid-19 Human Genetics Initiative et al (2022) Whole genome sequencing reveals host factors underlying critical Covid-19. Nature 607:97–103. https://doi.org/10.1038/s41586-022-04576-6

Kumar V, Wijmenga C, Withoff S (2012) From genome-wide association studies to disease mechanisms: celiac disease as a model for autoimmune diseases. Semin Immunopathol 34(4):567–580. https://doi.org/10.1007/s00281-012-0312-1

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G (2010) GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28(5):495–501. https://doi.org/10.1038/nbt.1630

Mozzi A, Pontremoli C, Forni D, Clerici M, Pozzoli U, Bresolin N, Cagliani R, Sironi M (2015) OASes and STING: adaptive evolution in concert. Genome Biol Evol 7(4):1016–1032. https://doi.org/10.1093/gbe/evv046

Noguchi S, Hamano E, Matsushita I, Hijikata M, Ito H, Nagase T, Keicho N (2013) Differential effects of a common splice site polymorphism on the generation of OAS1 variants in human bronchial epithelial cells. Hum Immunol 74(3):395–401. https://doi.org/10.1016/j.humimm.2012.11.011

Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, Walker S, Parkinson N, Fourman MH, Russell CD, Furniss J, Richmond A, Gountouna E, Wrobel N, Harrison D, Wang B, Wu Y, Meynert A, Griffiths F, Oosthuyzen W, Kousathanas A, Moutsianas L, Yang Z, Zhai R, Zheng C, Grimes G, Beale R, Millar J, Shih B, Keating S, Zechner M, Haley C, Porteous DJ, Hayward C, Yang J, Knight J, Summers C, Shankar-Hari M, Klenerman P, Turtle L, Ho A, Moore SC, Hinds C, Horby P, Nichol A, Maslove D, Ling L, McAuley D, Montgomery H, Walsh T, Pereira AC, Renieri A, Shen X, Ponting CP, Fawkes A, Tenesa A, Caulfield M, Scott R, Rowan K, Murphy L, Openshaw PJM, Semple MG, Law A, Vitart V, Wilson JF, Baillie JK, The GenOMICC Investigators, The ISRIC4C Investigators, The COVID19 Human Genetics Initiative, 23andMe Investigators, BRACOVID Investigators, Gen-COVID Investigators et al (2021) Genetic mechanisms of critical illness in COVID-19. Nature 591(7848):92–98. https://doi.org/10.1038/s41586-020-03065-y

Pless O, Kowenz-Leutz E, Knoblich M, Lausen J, Beyermann M, Walsh MJ, Leutz A (2008) G9a-mediated lysine methylation alters the function of CCAAT/enhancer-binding protein-beta. J Biol Chem 283(39):26357–26363. https://doi.org/10.1074/jbc.M802132200

Roy SK, Hu J, Meng Q, Xia Y, Shapiro PS, Reddy SPM, Platanias LC, Lindner DJ, Johnson PF, Pritchard C, Pagés G, Pouyssegur J, Kalvakolanu DV (2002) MEKK1 plays a critical role in activating the transcription factor C/EBP-beta-dependent gene expression in response to IFN-gamma. Proc Natl Acad Sci USA 99(12):7945–7950. https://doi.org/10.1073/pnas.122075799

Schaller MA, Kallal LE, Lukacs NW (2008) A key role for CC chemokine receptor 1 in T-cell-mediated respiratory inflammation. Am J Pathol 172(2):386–394. https://doi.org/10.2353/ajpath.2008.070537

Sengez B, Aygün I, Shehwana H, Toyran N, Tercan Avci S, Konu O, Stemmler MP, Alotaibi H (2019) The transcription factor *Elf3* is essential for a successful mesenchymal to epithelial transition. Cells 8(8):858. https://doi.org/10.3390/cells8080858

Severe Covid GG (2020) Genomewide association study of severe Covid-19 with respiratory failure. N Engl J Med 383(16):1522–1534. https://doi.org/10.1056/NEJMoa2020283

Smemo S, Tena JJ, Kim K-H, Gamazon ER, Sakabe NJ, Gómez-Marín C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF, Lee JH, Puviindran V, Tam D, Shen M, Son JE, Vakili NA, Sung H-K, Naranjo S, Acemel RD, Manzanares M, Nagy A, Cox NJ, Hui C-C, Gomez-Skarmeta JL, Nóbrega MA (2014) Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature 507(7492):371–375. https://doi.org/10.1038/nature13138

Tamura T, Yanai H, Savitsky D, Taniguchi T (2008) The IRF family transcription factors in immunity and oncogenesis. Annu Rev Immunol 26:535–584. https://doi.org/10.1146/annurev.immunol.26.021607.090400

The C-HGI (2020) 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 28(6):715–718. https://doi.org/10.1038/s41431-020-0636-6

Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, Martin HC, Lappalainen T, Posthuma D (2021) Genome-wide association studies. Nat Rev Methods Primers 1(1):59. https://doi.org/10.1038/s43586-021-00056-9

Wang S, Wang Y, Yu C, Cao Y, Yu Y, Pan Y, Su D, Lu Q, Yang W, Zuo Y, Yang L (2020) Characterization of the relationship between FLI1 and immune infiltrate level in tumour immune microenvironment for breast cancer. J Cell Mol Med 24(10):5501–5514. https://doi.org/10.1111/jcmm.15205

Wargodsky R, Dela Cruz P, LaFleur J, Yamane D, Kim JS, Benjenk I, Heinz E, Irondi OO, Farrar K, Toma I, Jordan T, Goldman J, McCaffrey TA (2022) RNA Sequencing in COVID-19 patients identifies neutrophil activation biomarkers as a promising diagnostic platform for infections. PLoS ONE 17(1):e0261679. https://doi.org/10.1371/journal.pone.0261679

Watanabe K, Taskesen E, van Bochoven A, Posthuma D (2017) Functional mapping and annotation of genetic associations with FUMA. Nat Commun 8(1):1826. https://doi.org/10.1038/s41467-017-01261-5

World Health Organization (2022) Overview of Coronavirus disease (COVID-19). https://www.who.int/health-topics/coronavirus. Accessed 14 June 2022

Xin J, Zhang H, He Y, Duren Z, Bai C, Chen L, Luo X, Yan D-S, Zhang C, Zhu X, Yuan Q, Feng Z, Cui C, Qi X, Ouzhuluobu WWH, Wang Y, Su B (2020) Chromatin accessibility landscape and regulatory network of high-altitude hypoxia adaptation. Nat Commun 11(1):4928. https://doi.org/10.1038/s41467-020-18638-8

Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM, Yang J (2016) Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48(5):481–487. https://doi.org/10.1038/ng.3538