PLOS GENETICS

RESEARCH ARTICLE

COVID-19 and systemic lupus erythematosus genetics: A balance between autoimmune disease risk and protection against infection

Yuxuan Wang[1], Suri Guga[1], Kejia Wu[1], Zoe Khaw[1], Konstantinos Tzoumkas[1], Phil Tombleson[2], Mary E. Comeau[3], Carl D. Langefeld[3], Deborah S. Cunninghame Graham[1], David L. Morris[1]☯*, Timothy J. Vyse[1]☯

1 Department of Medical & Molecular Genetics, King's College London, London, United Kingdom, 2 NIHR GSTFT/KCL Biomedical Research Centre, London, United Kingdom, 3 Department of Biostatistics and Data Science and Center for Precision Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina, United States of America

☯ These authors contributed equally to this work.
* david.l.morris@kcl.ac.uk

Abstract

Genome wide association studies show there is a genetic component to severe COVID-19. We find evidence that the genome-wide genetic association signal with severe COVID-19 is correlated with that of systemic lupus erythematosus (SLE), having formally tested this using genetic correlation analysis by LD score regression. To identify the shared associated loci and gain insight into the shared genetic effects, using summary level data we performed meta-analyses, a local genetic correlation analysis and fine-mapping using stepwise regression and functional annotation. This identified multiple loci shared between the two traits, some of which exert opposing effects. The locus with most evidence of shared association is *TYK2*, a gene critical to the type I interferon pathway, where the local genetic correlation is negative. Another shared locus is *CLEC1A*, where the direction of effects is aligned, that encodes a lectin involved in cell signaling, and the anti-fungal immune response. Our analyses suggest that several loci with reciprocal effects between the two traits have a role in the defense response pathway, adding to the evidence that SLE risk alleles are protective against infection.

Author summary

We observed a correlation between the genetic associations with severe COVID-19 and those with systemic lupus erythematosus (SLE, Lupus), and aimed to discover which genetic loci were shared by these diseases and what biological processes were involved. This resulted in the discovery of several genetic loci, some of which had alleles that were risk for both diseases and some of which were risk for severe COVID-19 yet protective for SLE. The locus with most evidence of shared association (*TYK2*) is involved in interferon production, a process that is important in response to viral infection and known to be dysregulated in SLE patients. Other shared associated loci contained genes also involved in

---

the defense response and the immune system signaling. These results add to the growing evidence that there are alleles in the human genome that provide protection against viral infection yet are risk for autoimmune disease.

Introduction

The outbreak of COVID-19 together with modern genotyping technologies has given us the unprecedented opportunity to investigate the genetics of response to viral infection. Recent GWAS of severe COVID-19 have shown that there is a genetic component to the variability of the clinical outcome [1]. Some of the genetic loci identified unsurprisingly point to pathways involved in the host immune response. Therefore, a comparison between the genetics of severe COVID-19 and autoimmune disease (AID) may be enlightening. In this study we compare the genetics of severe COVID-19 with those of systemic lupus erythematosus (SLE). The rationale for selecting SLE is twofold: some SLE risk alleles act to augment the interferon response (e.g. *IRF5*, *IRF7*, *CXORF21-TASL*); other lupus susceptibility genes act in the intracellular viral sensing (e.g. *IFIH1*, *TLR7*, *RNASEH2C*) pathway.

Results

Genetic correlation

To investigate the shared genetics between SLE and severe COVID-19, we ran a genome-wide genetic correlation analysis between ancestry matched SLE and severe COVID-19 association data. The SLE data comprised a meta-analysis of three European GWASs [2–4] (N_{cases} = 5,734, $N_{controls}$ = 11,609, **Table A in S1 Text**) and for the COVID-19 we used the GenOMICC release 1 European data [1] (critically ill patients with COVID-19 vs. ancestry-matched control individuals from UK Biobank, N_{cases} = 1,676, $N_{controls}$ = 8,380, **Table A in S1 Text**). We found the two traits to be genetically correlated (r_g = 0.56, s.e. = 0.16, p = 3 x 10^{-04}). To identify which regions were driving this correlation we ran a local genetic correlation analysis that included Immunochip European data [5] in the SLE meta-analysis (additional N_{cases} = 3,568, $N_{controls}$ = 11,245, **Table A in S1 Text**). This identified multiple loci with both positive and negative correlation of which the *TYK2* locus was the most significantly correlated (p-value = 1 x 10^{-04}, **Table B in S1 Text**). This gene encodes a kinase that regulates transduction of IFN-I signaling. An overview of GWAS data used in the study is illustrated in **Fig 1** and Table A in **S1 Text**.

Shared genetic associations: Severe COVID-19—SLE meta-analyses

To search for shared associations between SLE and severe COVID-19, we used summary association data from the large SLE meta-analysis (three SLE European GWASs plus immunochip), and for severe COVID-19 we used summary association data from the COVID-19 Host Genetics initiative (COVID-19 hg) [6] release 6 data (GenOMICC study is a subset of these data) association results of very severe respiratory confirmed COVID-19 vs. population (A2_ALL_leave_23andme, N_{cases} = 8,779, $N_{controls}$ = 1,001,875, **Table A in S1 Text**).

We checked published associations in each trait to our summary association data for the other trait, and validated by coclisation analysis (coloc [7]), to identify potential shared risk loci (see Material and Methods). This found evidence of shared association at *TYK2*. Our colocalization analysis of all loci that had at least one SNP with p < 1 x 10^{-05} in both diseases (see Material and Methods) identified *TYK2* and *CLEC1A*, a C-type lectin that is a negative regulator of dendritic cells.

Fig 1. Overview of GWAS data used in the study.

We performed a cross-trait meta-analysis that included an analysis to highlight opposing effects (see Material and Methods, overlapped N_{SNPs} = 1,559,546). Manhattan plots from the meta-analyses can be seen in **Fig 2**. There were 15 loci that had genome-wide significant evidence of (p-values $< 5 \times 10^{-08}$, **Table 1**), the very significant p-values at the *TYK2* locus in the lower plot (**Fig 2**) highlights the negative correlation at this locus. There were six association signals in five of these loci with colocalization probabilities (PP_{H4}) greater than 0.8 and three of these, implicating *CLEC1A*, *TYK2* and *PDE4A*, had $PP_{H4} > 0.95$ (**Table 1**). The *TYK2-PDE4A* locus had opposing direction of effect across the two diseases and the other 4 loci (*CLEC1A*, *IL12B*, *PLCL1-RFTN2*, and *MIR146A*) had agreement in direction of effect. Though genome-wide significant evidence were found in the other 10 loci, there was relatively weak evidence for colocalization. Two well-known SLE associated loci, *IRF8* and *TNFSF4*, showed evidence of significant association in the opposing effect meta-analysis with some evidence for colocalisation of shared signals at both loci (*IRF8* PP_{H4} = 0.36, *TNFSF4* PP_{H4} = 0.37; **Tables C and D and Fig A in S1 Text**). LocusZoom plots for all other loci can be seen in **Figs B-L in S1 Text**. A pathway analysis showed that there was an enrichment of genes in defense response, cytokine-mediated signaling and type I interferon signaling pathway with over half the genes being included in one or more pathways (**Table E and Fig M in S1 Text**).

Tyrosine kinase 2 (TYK2)

The *TYK2* locus has previously been found to be associated with SLE [4,8–12] and severe COVID-19 [1]. There was significant negative local genetic correlation (p-value = 1×10^{-04}, ρ-HESS, overlapped N_{SNP} = 2,544) at *TYK2* between the two diseases. In a stepwise regression approach using summary meta-analysis data for both traits, we found a highly significant overlap between genetic association signals (overlapped N_{SNP} = 4,720); importantly, the SLE risk alleles were protective against severe COVID-19. The locus-wide association signals in COVID-19 and SLE are compared in **Fig 3A**. There were two independent signals that colocalized across traits (posterior probabilities of coloc = 0.991 and 0.993), referred to arbitrarily as signal-A and signal-B in **Table 2**. The top two SNPs independently associated with SLE
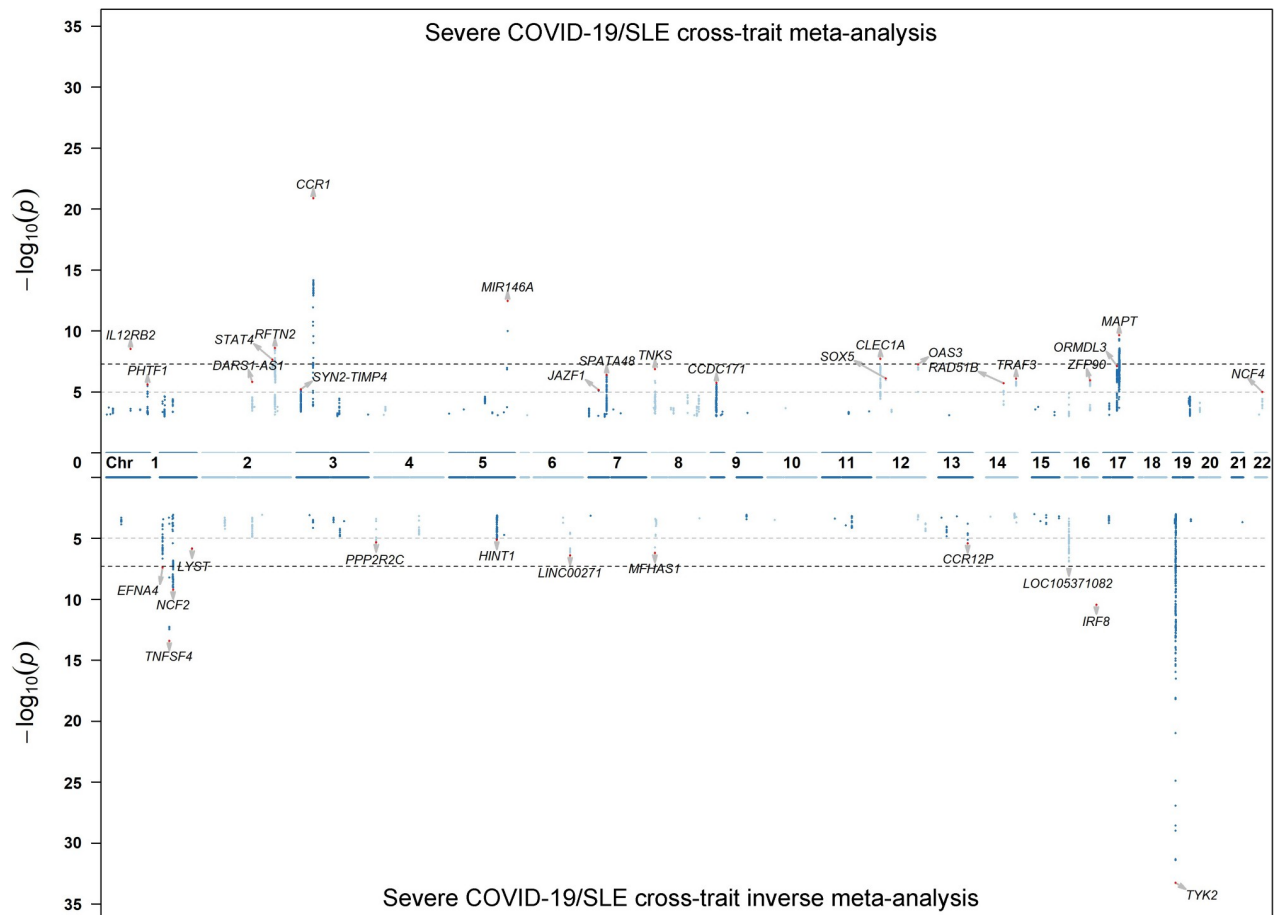
Fig 2. Cross-trait meta-analysis log10 p-values. Upper plot has results from a standard inverse variance meta-analysis. The lower plot has results form a meta-analysis when reversing the severe COVID-19 direction of effect. The MHC extended region (chr6: 24–36 Mb) was removed. Signals that the lead SNP has p-values $< 1 \times 10^{-05}$ were annotated by the closest gene names according to base pair position.

https://doi.org/10.1371/journal.pgen.1010253.g002

(rs34536443 and rs34725611) are in high LD ($r^2 = 0.88$ and 0.97 respectively), with the two SNPs we found to be independently associated with severe COVID-19 (rs74956615 and rs11085727) being reported previously in a COVID-19 GWAS [1]. For a full set of association results for these SNPs across traits see **Table F in S1 Text,** where it is shown that for all SNPs the effects have reciprocal directions of effect in SLE and COVID-19 outcome. In both traits, the relatively rare variants rs34536443/rs74956615 were associated independently from the more common variants rs34725611/rs11085727 (see conditional results bJ and pJ in **Table 2A and 2B**; $r^2 = 0.06$ and 0.09 between rs34536443 and rs34725611 and between rs74956615 and rs11085727 respectively in the EUR SLE data, $r^2 = 0.08$ and 0.07 in the 1000 genomes EUR data).

The lead SLE SNP rs34536443 for signal-A is a missense variant (**Table 2C**) and homozygosity at the SLE protective allele (C) drives a near complete loss of *TYK2* function and consequently impairs type I IFN, IL-12 and IL-23 signaling [13]. The genetic association in signal-A, which was tagged by rs34536443/rs74956615, also colocalizes with the *cis* eQTL signal for *PDE4A* in artery tibial in GTEx v8 data (**Fig 4A**, $PP_{H4} > 0.99$ for colocalisation between the two traits and with the eQTL signal) where the severe COVID-19 risk allele, that is protective for SLE, is associated with reduced expression (**Table G in S1 Text**).

Table 1. Association results for lead SNPs in the cross-trait meta-analysis/inverse meta-analysis of severe COVID-19 and SLE data. Posterior possibilities (PP_{H4}) of colocalisation between signals were estimated with region ± 1 Mb of lead SNPs, if there are multiple independent signals the highest PP_{H4} of colocalisation was shown. Predicted functional genes were inferred by lead SNP in the region and its LD with coding variants, eQTL data (GTEx v8 and eQTLGen), ENCODE ChIP-seq marker data, GeneHancer interactions data, and published functional studies.

Signals showing aligned effect in a cross-trait meta-analysis

SNP	Position	A1	meta		SLE		severe COVID-19		PP_{H4}	predicted functional genes in SLE	predicted functional genes in COVID-19
			OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P			
rs7960611	12:10230416	G	1.14 (1.09–1.19)	8.02×10^{-09}	1.11 (1.05–1.17)	2.76×10^{-04}	1.17 (1.10–1.25)	3.31×10^{-07}	95.4%	CLEC1A[1]	CLEC1A[1]
rs6869688	5:158883027	G	0.92 (0.90–0.95)	4.57×10^{-08}	0.90 (0.87–0.94)	1.79×10^{-08}	0.94 (0.91–0.98)	1.96×10^{-03}	81.5%	IL12B[2]	IL12B[2]
rs10460393	2:198548306	T	1.09 (1.06–1.13)	1.07×10^{-09}	1.11 (1.07–1.14)	2.11×10^{-08}	1.08 (1.04–1.13)	6.48×10^{-05}	80.7%	PLCL1[3], RFTN2[4]	PLCL1[3], RFTN2[4]
rs2431697	5:159879978	C	0.90 (0.87–0.92)	8.51×10^{-14}	0.84 (0.81–0.87)	2.12×10^{-20}	0.94 (0.91–0.98)	2.42×10^{-03}	80.5%	MIR146A[5]	MIR146A[5]
rs4792891	17:43973498	G	0.90 (0.88–0.93)	6.82×10^{-10}	0.92 (0.89–0.96)	1.40×10^{-05}	0.90 (0.86–0.93)	3.37×10^{-08}	56.3%	MAPT[6]	MAPT[6]
rs5022165	1:67788352	A	1.12 (1.08–1.16)	9.10×10^{-09}	1.17 (1.12–1.23)	1.82×10^{-12}	1.07 (1.02–1.13)	8.52×10^{-03}	23.0%	IL12RB2[7]	IL12RB2[7]
rs3024897	2:191896564	C	0.87 (0.83–0.91)	1.06×10^{-08}	0.82 (0.77–0.87)	1.17×10^{-10}	0.91 (0.86–0.97)	5.41×10^{-03}	14.6%	STAT1[8], STAT4[9]	STAT1[8], STAT4[9]
rs35605052	3:45916547	T	1.16 (1.12–1.21)	2.14×10^{-13}	1.08 (1.03–1.13)	6.95×10^{-04}	1.25 (1.18–1.32)	1.02×10^{-14}	6.03%	CXCR6[10]	CXCR6[10], SLC6A20, FLT1P1, FYCO1, CCR1, CCR3
rs7970893	12:113390679	T	0.92 (0.90–0.95)	2.10×10^{-08}	0.94 (0.91–0.98)	2.26×10^{-03}	0.90 (0.87–0.94)	1.28×10^{-07}	0.00%	ALDH2, SH2B3	OAS1, OAS2, OAS3

Signals showing opposing effect in a cross-trait inverse meta-analysis

SNP	Position	A1	meta		SLE		severe COVID-19		PP_{H4}	predicted functional genes in SLE	predicted functional genes in COVID-19
			OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P			
rs11085727	19:10466123	T	1.21 (1.17–1.25)	2.09×10^{-31}	0.80 (0.77–0.84)	6.92×10^{-27}	1.19 (1.14–1.23)	2.33×10^{-18}	99.3%	TYK2[11]	TYK2[11]
rs74956615	19:10427721	A	1.51 (1.39–1.63)	2.19×10^{-25}	0.58 (0.53–0.65)	2.75×10^{-24}	1.40 (1.29–1.53)	3.04×10^{-14}	99.1%	PDE4A[12]	PDE4A[12]
rs1174683	1:183650428	G	1.16 (1.11–1.22)	2.37×10^{-09}	0.82 (0.77–0.87)	3.98×10^{-10}	1.12 (1.05–1.18)	1.50×10^{-04}	60.1%	NCF2[13]	NCF2[13]
rs5778759	1:173328868	C	1.13 (1.09–1.16)	3.32×10^{-13}	0.83 (0.80–0.86)	1.98×10^{-22}	1.06 (1.01–1.11)	9.66×10^{-03}	36.8%	TNFSF4[14]	TNFSF4[14]
rs17445836	16:86017663	A	1.14 (1.09–1.19)	5.14×10^{-10}	0.83 (0.79–0.87)	3.68×10^{-16}	1.07 (1.02–1.13)	7.11×10^{-03}	35.7%	IRF8[15]	IRF8[15]
rs61811916	1:155045004	C	1.14 (1.09–1.19)	4.11×10^{-08}	1.11 (1.05–1.17)	4.59×10^{-04}	0.85 (0.80–0.91)	4.86×10^{-07}	8.01%	ADAM15[16]	ADAM15[16], GBA, MUC1, THBS3, GBAP1

(Continued)

Table 1. (Continued)

| rs76073397 | 16:11386452 | C | 1.14 (1.09–1.19) | 3.21 x 10⁻⁰⁸ | 0.90 (0.84–0.95) | 5.14 x 10⁻⁰⁴ | 1.15 (1.09–1.22) | 9.67 x 10⁻⁰⁷ | 1.18% | *RMI2*[17], *CLEC16A, DEXI* | *RMI2*[17] |

Genes associated with autoimmune or infectious diseases includes: [1]experimental autoimmune encephalomyelitis (EAE). [2]psoriasis, crohn's disease (CD), inflammatory bowel disease (IBD), ankylosing spondylitis (AS), sclerosing cholangitis (SC), ulcerative colitis (UC), psoriatic arthritis (PsA), multiple sclerosis (MS), ulcerative colitis (UC), primary biliary cirrhosis (PBC), autoimmune thyroid disease (AITD), celiac disease (CeD), type 1 diabetes (T1D), juvenile idiopathic arthritis (JIA), rheumatoid arthritis (RA). [3]UC, SLE, CD, allergic rhinitis, asthma, RA, IBD, AS, psoriasis, SC. [4]atopic asthma. [5]SLE, Sjogren's syndrome (SS), RA, MS, AITD. [6]PBC, SS. [7]SLE, PBC, systemic scleroderma (SSc), RA, CD, MS, AS, IBD, behcet's disease (BD). [8]JIA, RA. [9]SLE, RA, SSc, PBC, SS, CeD, BD, IBD, autoimmune hepatitis type-1 (AIH), T1D, MS, AITD, CD, JIA, UC, non-typhoidal Salmonella bacteremia. [10]PBC, T1D, AIH, EAE. [11]SLE, COVID-19, psoriasis, SSc, MS, RA, T1D, PBC, AS, CD, SC, UC, IBD, AITD. [12]SLE, psoriasis, MS, JIA, T1D, AS, CeD, CD, UC, AITD. [13]SLE, RA, SSc, CeD. [14]SLE, asthma, atopic asthma, SSc, allergic rhinitis, AITD. [15]SLE, PBC, SSc. [16]SLE. [17]PBC, IBD, MS, CD, T1D, psoriasis, CeD, AS, UC, SC, JIA, asthma.

We found that the genetic associations in signal-B tagged by rs34725611 and rs11085727 colocalize with a *TYK2* eQTL signal in whole blood in eQTLGen [14] (**Fig 4B**) and GTEx v8 data, and adrenal gland in GTEx v8 data: PP_H4 > 0.98 for colocalisation between the two traits and with all eQTL signals (**Fig N in S1 Text**). eQTL summary statistics can be seen in **Table G in S1 Text**, where the associated allele effects can be compared across traits and eQTL. In all
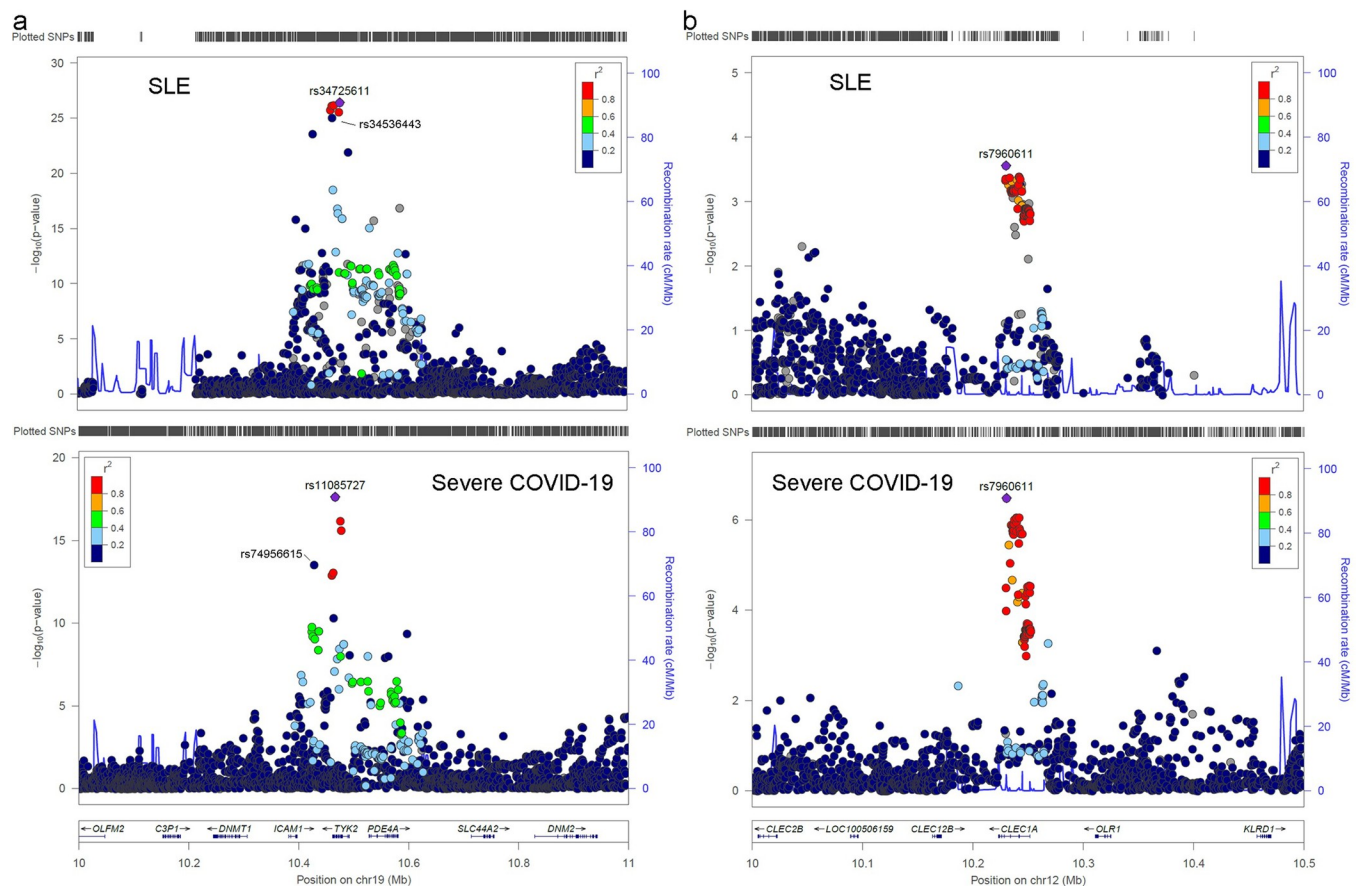

Fig 3. Locus zoom plots across a) *TYK2*, **b)** *CLEC1A* **for single marker associations with SLE, severe COVID-19.** The LD (r² in 1000 Genome project Phase 3 EUR) is identified by color.

Table 2. *TYK2* association results for a) SLE and b) severe COVID-19 data, and c) summary of functional effects of associated alleles. Independently associated SNPs in SLE and severe COVID-19 are displayed.

Table 2a. *TYK2* associations with SLE

Signal number	SNP	position	A2	A1	freq(A1)	OR	95% CI	P	bJ	bJ_se	pJ
Signal-A	rs34536443	19:10463118	G	C	0.034	0.53	0.47–0.59	9.80×10^{-26}	-0.57	0.06	9.47×10^{-20}
Signal-B	rs34725611	19:10477067	A	G	0.275	0.80	0.77–0.83	3.84×10^{-27}	-0.16	0.02	2.86×10^{-13}

Table 2b. *TYK2* associations with severe COVID-19

Signal number	SNP	position	A2	A1	freq(A1)	OR	95% CI	P	bJ	bJ_se	pJ
Signal-A	rs74956615	19:10427721	T	A	0.047	1.40	1.27–1.55	3.04×10^{-14}	0.26	0.05	4.06×10^{-08}
Signal-B	rs11085727	19:10466123	C	T	0.280	1.19	1.14–1.23	2.33×10^{-18}	0.14	0.02	2.83×10^{-12}

Table 2c. Functional effects of associated alleles in *TYK2 –PDE4A* locus.

Signal number	Ref SNP	Minor/ancestral Allele	Ancestral Allele SLE effect	Ancestral Allele COVID Effect	Function	Gene	Ancestral Allele Functional Effect
A	rs34536443	C/G; Ala1104Pro	Risk	Protective	Coding	Tyrosine Kinase 2 (*TYK2*)	Increased Gene function through increased phosphorylation [13]
A	rs34536443	C/G	Risk	Protective	Regulation	Phosphodiesterase 4A (*PDE4A*)	Increased Gene Expression: eQTL data (**Table G in S1 Text**)
B	rs2304256	A/C Phe362Val	Risk	Protective	Coding	Tyrosine Kinase 2 (*TYK2*)	Decreased Gene function though loss of exon 8 [15]
B	rs11085727	T/C	Risk	Protective	Regulation	Tyrosine Kinase 2 (*TYK2*)	Decreased Gene Expression: eQTL data (**Table G in S1 Text**)
B	rs11085727	T/C	Risk	Protective	Regulation	Serpin Family G Member 1 (*SERPING1*)	Increased Protein Expression: pQTL data [17]
B	rs11085727	T/C	Risk	Protective	Regulation	C-X-C Motif Chemokine Ligand 10 (*CXCL10*) (IP-10)	Increased Protein Expression: pQTL data [17]

* bJ, bJ_se, pJ: effect size, standard error and p-value from a joint analysis (multiple regression) of all the selected SNPs (results conditional on all other SNPs if selected from stepwise regression). † rs34536443 is in high LD with rs74956615 ($r^2 = 0.88$), rs34725611 is in high LD with rs11085727 ($r^2 = 0.97$). In table c) we refer to the common ancestral allele for effects where this is protective for severe COVID-19 and risk for SLE. Functional effects cover coding variation, *cis* acting gene transcript expression and *trans* acting protein product expression.

https://doi.org/10.1371/journal.pgen.1010253.t002

cases the protective allele for SLE, which is the risk allele for severe COVID-19, increases expression. However, signal-B is also associated with altered TYK2 function as a missense variant rs2304256 (V362F, exon 8), that is in strong LD ($r^2 = 0.98$ in SLE data) with rs11085727, acts as a splicing eQTL. The SLE protective allele promotes inclusion of exon 8 [15], which increases TYK2 function. Thus, signal-B provides conflicting results with respect to signal-A regarding the functional impact on TYK2. To understand the role of signals A and B on gene regulation, we studied the epigenetic landscape around these two association signals (**Fig O in S1 Text**). For signal A, there was evidence for localization to enhancer chromatin marks (H3K27Ac and H3K4Me1, **Fig P in S1 Text**). However, there was much less evidence for such alignment with signal-B (**Fig Q in S1 Text**). Signal-A is also observed to loop in 3D space to the promotor of *PDE4A* (**Fig R in S1 Text**). While we did observe other significant *cis* eQTLs with signal-B SNPs (**see Fig S in S1 Text**), none of them colocalized with COVID-19 or SLE signals ($PP_{H4} < 0.20$ in all cases).
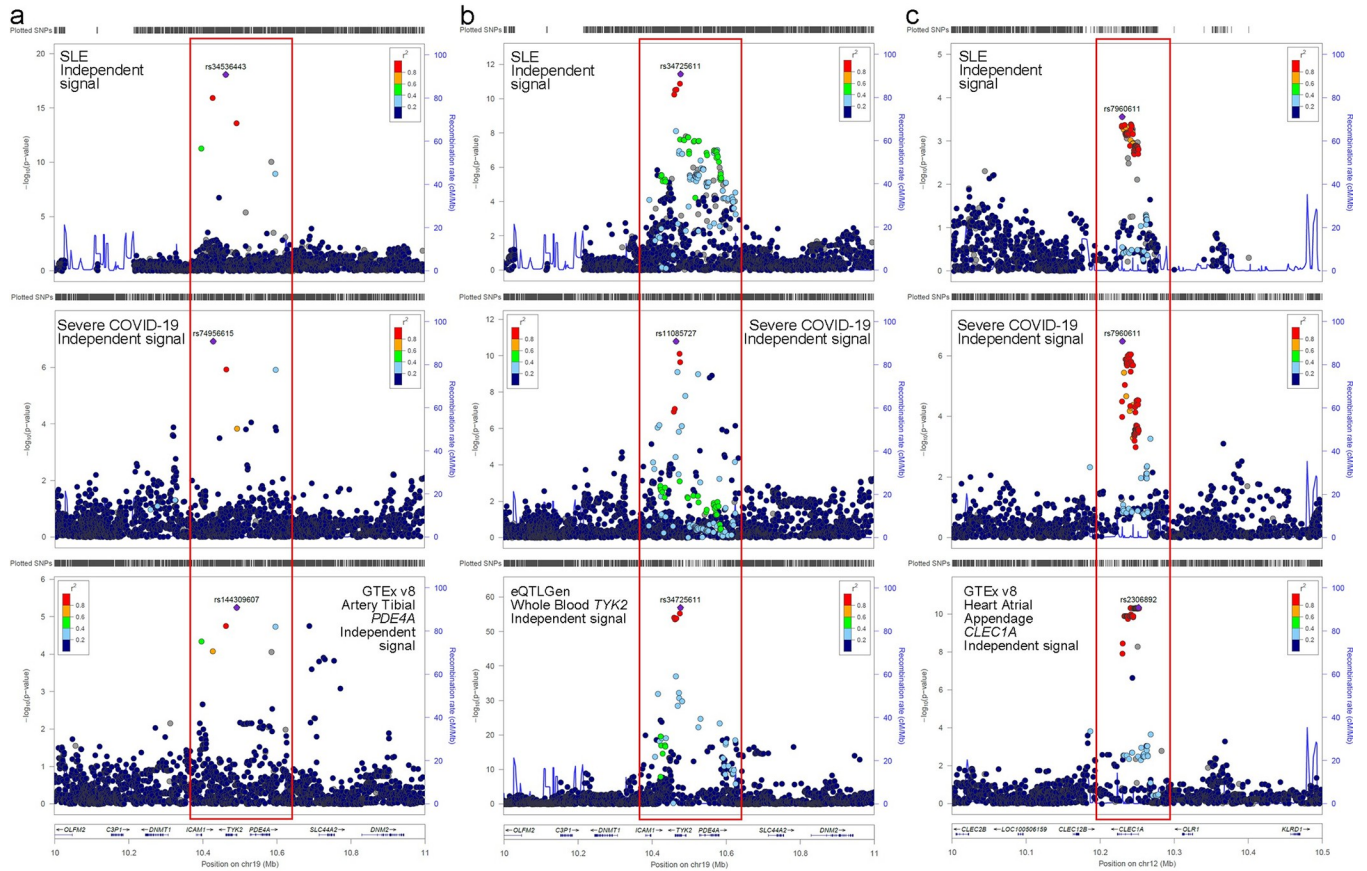
Fig 4. Locus zoom plots across loci for marginal associations with SLE, severe COVID-19 and eQTL. a) *PDE4A* locus, both diseases' signal-B colocalized with eQTL for *PDE4A*. **b) *TYK2* locus,** both diseases' signal-A colocalized with eQTL for *TYK2*. **c) *CLEC1A* locus,** both diseases colocalized with eQTL for *CLEC1A*. The LD (r^2 in 1000 Genome project Phase 3 EUR) is identified by color.

https://doi.org/10.1371/journal.pgen.1010253.g004

To explore the functional effects further, we looked for downstream effects of the shared *TYK2* associated SNPs on the expression levels of a set of 21 IFN-induced genes (dysregulated in SLE [16]) in human plasma proteome data [17]. Significant trans pQTLs (FDR < 0.01, **Table H in S1 Text**) for signal-B (rs11085727) were found for two targets: SERPING1 and CXCL10 (IP-10). Both proteins are induced by interferon and would be expected to require TYK2 activity for induction. The COVID-19 risk allele (T), that correlates with increased *TYK2* transcript expression, correlated with reduced amounts of SERPING1 and CXCL10 proteins in plasma (p = 0.0003).

C-Type lectin domain family 1 member A (CLEC1A)

The meta-analysis identified a narrow peak of association between 10.2–10.3Mb on chromosome 12 that colocalized between the two traits (**See Fig 3B**; PP_{H4} = 0.95 and **Table 3**, overlapped N_{SNP} = 3,363). Both traits' association signals colocalized with eQTLs for *CLEC1A* in multiple tissues ($PP_{H4} \geq 0.97/0.87$ for eQTL colocalisation with COVID-19/SLE) in GTEx v8 data. **Fig 4C** displays the association in both diseases and eQTL data for heart (atrial appendage), see **Fig T in S1 Text** for the other eQTL colocalization. eQTL summary statistics can be seen in **Table G in S1 Text**. The risk allele for severe COVID-19 is also risk for SLE and is associated with reduced expression of *CLEC1A*. The lead variant rs7960611 is in LD with a missense variant rs2306894 (r^2 = 0.84).

Table 3. *CLEC1A* association results for a) SLE and b) severe COVID-19 data, and c) summary of functional effects of associated alleles. Independently associated SNPs in SLE and severe COVID-19 are displayed.

Table 3a. *CLEC1A* associations with SLE											
Signal number	SNP	position	A2	A1	freq(A1)	OR	95% CI	P	bJ	bJ_se	pJ
Signal-A	rs7960611	12:10230416	A	G	0.115	1.11	1.04–1.17	2.76×10^{-04}	0.10	0.03	2.77×10^{-04}

Table 3b. *CLEC1A* associations with severe COVID-19											
Signal number	SNP	position	A2	A1	freq(A1)	OR	95% CI	P	bJ	bJ_se	pJ
Signal-A	rs7960611	12:10230416	A	G	0.124	1.17	1.11–1.24	3.31×10^{-07}	0.16	0.03	3.31×10^{-07}

Table 3c. Functional effects of associated alleles in *TYK2 –PDE4A* locus.							
Signal number	Ref SNP	Minor/ancestral Allele	Ancestral Allele SLE effect	Ancestral Allele COVID Effect	Function	Gene	Ancestral Allele Functional Effect
Signal-A	rs2306894	C/G; Gly26Ala	Protective	Protective	Coding	C-type lectin domain family 1 member A (*CLEC1A*)	Gly26Ala No publication found
Signal-A	rs7960611	G/A	Protective	Protective	Regulation	C-type lectin domain family 1 member A (*CLEC1A*)	Increased Gene Expression: eQTL data (**Table G in S1 Text**)

* bJ, bJ_se, pJ: effect size, standard error and p-value from a joint analysis (multiple regression) of all the selected SNPs (results conditional on all other SNPs if selected from stepwise regression). In table c) we refer to the common ancestral allele for effects where this is protective for severe COVID-19 and SLE. Functional effects cover coding variation and *cis* acting gene transcript. † rs2306894 is in high LD with rs7960611 ($r^2 = 0.84$).

https://doi.org/10.1371/journal.pgen.1010253.t003

Discussion

Our results indicate that there are shared genetic effects between the autoimmune disease SLE and the clinical consequences of COVID-19. The locus with the most evidence of shared effects was the Janus kinase (JAK), *TYK2*, that promotes IL-12 and IFN-I signaling. Here there are two separate genetic association signals (designated A and B) shared between severe COVID-19 and SLE. Importantly for both, the genetic factors for SLE risk mitigate the outcome following SARS-Cov2 infection. In seeking to uncover the mechanisms underlying these relationships it was apparent that the functional effects of the risk alleles are complex. Signal-A at *TYK2* is likely driven by a coding P1104A variant (rs34536443) whose COVID-19 risk allele has been shown to impair TYK2 target phosphorylation [13]. This is further supported by the therapeutic effect of a TYK2 inhibitor in psoriasis [18], and by observed risk in other infectious disease such as tuberculosis where it has been found that homozygosity for the minor allele (C) of rs34536443 is risk, in line with severe COVID-19, and strongly impairs IL-23 signaling in T cells and IFN-γ production in PBMC [19,20]. Signal-A, led by rs34536443, was also found to colocalize with an eQTL for nearby *PDE4A*, which encodes a phosphodiesterase that regulates cAMP. This enzyme has multiple potential roles, however PDE4A inhibitors have been shown to have anti-inflammatory activity and are being studied in AID and inflammatory lung diseases [21]. The severe COVID-19 risk alleles are associated with decreased expression of *PDE4A*, while they are protective for SLE. The *PDE4A* eQTL cell type is heterogeneous however and the relevance to SLE is unclear. Signal-B includes another missense variant in *TYK2*, namely rs2304256 (V362F) in exon 8, but this also acts as a splicing mutation and the missense variant is missing from the spliced transcript. The severe COVID-19 risk allele promotes inclusion of exon 8 in *TYK2* that is essential for *TYK2* binding to cognate receptors [15]. Therefore

signal-B comprises evidence for two functional effects with respect to COVID-19 risk alleles, one of which increases function of *TYK2* through altered splicing (rs2304256 (V362F)) and one that is correlated with increased expression of *TYK2* (rs11085727). It may be that the overall reduction of *TYK2* activity caused by the COVID-19 risk alleles in signal-A evokes a compensatory effect on overall gene expression, which is designed to mitigate the deleterious effect of the missense variants–an example of regulatory variants modifying the penetrance of coding variants [15,22]. This conjecture is supported by the lack of epigenetic marks in the signal-B region of *TYK2*.

The severe COVID-19 risk allele for signal-B at *TYK2* is associated with reduced SERPING1 and CXCL10 protein expression, implying that the minor allele at signal-B in the *TYK2* locus reduces some aspect of *TYK2* function. CXCL10 (IP-10) is a chemokine that acts on Th1 cells and is key regulator of the cytokine storm immune response to COVID-19 infection [23]. SERPING1, an inhibitor of complement 1 (C1-inh), is known to be reduced by infection and this reduction correlates with more severe COVID-19 [24]. Therefore genetic predisposition to low SERPING1 expression may increase risk for COVID-19 through the same dynamics as reduced levels due to infection. This and the effect of reduced levels of CXCL10 are likely just two examples of altered IFN induced activity that affects risk for disease.

We found agreement in direction of effect of association in *CLEC1A*. *CLEC1A* is interesting as C-type Lectin receptors are involved in fungal recognition and fungal immunity. Genetic variation in *CLEC1A* is a risk factor for the development of Aspergillosis in immunosuppression [25]. CLEC1A is a negative regulator of dendritic cells [26]. Therefore the SLE and severe COVID-19 risk allele, being associated with reduced expression of *CLEC1A*, would be expected to exert a pro-inflammatory effect. We also found agreement in direction of effect of associations in 3 other loci (*IL12B*, *PLCL1-RFTN2*, *MIR146A*) that showed relatively strong evidence of colocalization. The modest p-values and relatively high colocalisation possibilities support them as good candidates to follow up in larger studies. At both *IRF8* and *TNFSF4* the evidence for association in severe COVID is moderate yet the signals do show some evidence of colocalizing with opposing effects in SLE. With prominent roles in the pro-inflammatory IFN response these two loci should be a focus when larger data in severe COVID-19 are available. *IRF8* provides more evidence that the IFN pathway is important in the balance between SLE risk and infection as mutations that impair *IRF8* transcriptional activity have been found to cause immunodeficiency [27]. Interferons constitute one of the main means of host defense against viruses and hence have been well studied in the context of COVID-19 [28–30]. In SLE, evidence for interferon activity is present in about half of the patients and is often present in those with more severe disease [31–33]. Although elevated interferon has been implicated in other AID, the role is prominent in SLE. This has been exploited with therapeutic agents designed to antagonize type I interferon activity showing benefit in SLE [34]. Parallels between SLE and viral infection extend beyond interferon activation though. As stated above there are SLE risk genes that act in the intracellular viral sensing pathways. SLE is characterized by an immune response against host nucleic acids. The means by which the immune system loses tolerance to these structures appears to involve aberrant exposure of self through the pathways that are designed to sense foreign nucleic acids, as happens during viral infection [35]. Further investigation into the genetic correlation between SLE and severe COVID-19 will help explain the genetic basis of both diseases, which may be in part due to variation in response to viral infection. Risk alleles for SLE, that are also risk for severe COVID-19, may persist in the population due to protective effects against other exposures such as fungal infection. The opposing effects we find at the *TYK2* locus is compatible with the hypothesis that there are alleles in the general population that, while represent a risk for SLE, persist possibly due to an innate immune protection against pathogens [36–41] including viruses.

Material and methods

Data for genome-wide and local genetic correlation

Full summary-level GenOMICC release 1 data were downloaded from https://genomicc.org/data. These data resulted from a GWAS of 1,676 critically ill patients with COVID-19 (severe COVID-19) of European ancestry from 208 UK intensive care units (GenOMICC GWAS data release 1), and ancestry-matched control individuals (8,380 of European ancestry) selected from the large population-based cohort of UK Biobank [1]. Controls with a known positive COVID-19 test were excluded [1]. An SLE meta-analysis of three previously published European GWASs was used (the SLE main cohort [42], 4,036 cases and 6,959 controls; the Genentech cohort [2], 1,165 cases and 2,107 controls; the SLEGEN cohort [43], 533 cases and 2,543 controls), each of these data have been pre-phased (SHAPEIT [44]) and imputed (IMPUTE [45,46], 1000 Genomes phase 3 [47]) using the same pipeline as in the previous studies of these data where they were imputed to the 1000 genomes phase 1 density [4,48]. SNPTEST was run in each dataset using principal components as covariates to control for population structure as in the original studies. A standard fixed effects inverse variance approach was used for meta-analysis using our own scripts written in R, that also checked for allele matching and strand issues, and METAL [49]. The genomic inflation factor (λ_{1000}) [50] was 1.02. To evaluate genetic correlation between SLE and severe COVID-19, we used conventional cross-trait LD score regression (LDSC) [51,52] to calculate genome-wide genetic correlation (r_g). All the overlapping SNPs between the SLE meta-analysis and the COVID-19 GenOMICC European data were retained for use. The number of SNPs were reduced to common SNPs (MAF > 0.01) from the European 1000 Genomes populations [47] (N_{SNP} = 413,464 genome-wide) as these data were used as the LD reference panel in the genetic correlation analyses.

To increase power for local genetic correlation detection, while maintaining the same ancestry as required by the methodology [53], we added SLE Immunochip data [5] from a previous study (3,568 cases and 11,245 controls independent of the three European GWAS) to the SLE meta-analysis. These data were also imputed to the density of the 1000 Genomes Phase 3 data. The new meta-analysis also used a standard fixed effects inverse variance approach (MAF > 0.01 and INFO > 0.9). The genomic inflation factor (λ_{1000}) was 1.03. Local genetic correlation was performed using a recent approach that uses summary statistics (ρ-HESS) [53] to estimate local SNP-level heritability and genetic covariance (correlation).

Data for the SLE–severe-COVID-19 meta-analyses, cross disease colocalisation analyses and fine-mapping

To maximize power for genetic association [54] we obtained multi-ancestry severe COVID-19 vs. population genetic association data from round 6 of the COVID-19 Host Genetics Initiative (COVID-19 hg, https://www.covid19hg.org/) where the GenOMICC study was a subset of these data [55]. The severe COVID-19 phenotype is defined as individuals critically ill with COVID-19 based on either requiring respiratory support in hospital or who died as a consequence of the disease [55]. These data comprised 8,779 cases vs. 1,001,875 controls (A2_ALL_-leave_23andme) [55] and were obtained from a google storage bucket provided by COVID-19 hg. The association summary data was the result of a meta-analysis of 60 studies from 25 countries and was performed by the provider with fixed effects inverse variance weighting after filtering for allele frequency > 0.001 and imputation INFO > 0.6 applied to each study. The SLE meta-analysis that included Immunochip data [5] was used for the meta-analysis with severe-COVID-19 and for fine-mapping (genome-wide overlapped N_{SNPs} = 1,559,546). We also used

the African American (2,970 cases and 2,452 controls) and Hispanic (1,872 cases and 2,016 controls) samples from the Immunochip study [5] for replication (see Supplementary information).

SLE–severe-COVID-19 meta-analyses

We performed two cross-trait meta-analyses between the SLE meta-analysis (Three EUR GWAS + EUR Immunochip) summary statistics and the severe COVID-19 HGI release 6 GWAS summary statistics using R, that also checked for allele matching and strand issues, and METAL [49]. Firstly, both diseases' summary statistics were analyzed using the inverse variance approach (upper plot in **Fig 2**). A second analysis was undertaken in which the severe COVID-19 direction of effect was reversed followed by a standard inverse variance meta-analysis (lower plot in **Fig 2**). This second approach is more powerful to detect areas of the genome that have genetic association with both diseases but the direct of effect is opposing between SLE and severe COVID-19. In both meta-analyses we only retained and plotted p-values for SNPs that had p < 0.01 in both diseases and had shared direction of effects with respect to each of the two types of meta-analysis. Any SNPs that passed a significance threshold of $P < 5 \times 10^{-08}$ in meta-analysis in both traits were considered as candidates for shared association. These were followed up by fine mapping in both traits and colocalisation analysis. The MHC region was not included.

Checking published and candidate associated loci across traits for shared association loci

Loci published as associated in each trait were checked for locus-wide association with the other trait ($p < 1 \times 10^{-05}$) in our summary association data. Loci were defined as the lead published SNP +/-1mb. Candidate shared loci were visible inspected using locus-wide LocusZoom plots [56] and loci were checked for colocalisation of association between the two diseases. We also investigated any locus that had a SNP with $p < 1 \times 10^{-05}$ in both traits in our summary data for colocalization, where the loci were defined as the shared associated SNP +/-1mb. On both these analyses we only declared a locus as shared if the colocalisation probability was greater than 0.9.

eQTL data

The cis-eQTL summary statistics data was obtained from eQTLGen Consortium (https://www.eqtlgen.org/) [14], which includes eQTL data from 31,684 whole blood samples across 37 cohorts cohorts mainly of European origin, and from European specific eQTL data from GTEx v8 across 54 tissues [57].

pQTL data

Two studies' combined summary pQTL data [58,59] were downloaded from https://gwas.mrcieu.ac.uk. These two studies data were combined and analyzed previously [17]. We focused our pQTL analysis on 21 IFN induced genes previously defined [16]. Associations between the SNPs in our study and the 21 gene's expression were retrieved if included in the study, otherwise tagging SNPs were used. A Bonferroni adjustment was made for multiple testing across all SNP/gene combinations.

Fine-mapping

Our main fine-mapping analysis consisted of comparing summary association data between SLE and severe COVID-19. This consisted of an approximate stepwise regression using COJO [60] in both diseases' data to identify independent signals and colocalization analyses to

investigate whether shared association were coincidental. For supplementary information we also ran stepwise regression on the SLE individual level data, which we were unable to do on the COVID-19 data and so no comparison could be made.

COJO. To find independently associated variants using summary data, we performed an approximation of stepwise regression using GTCA 1.93 [61] (COJO [60], 'cojo-slct'). Lead SNPs from stepwise regression were taken as index SNPs and marginal signals were obtained by conditional analysis (-cojo-cond) on the set of index SNPs that were not in the signal of interest. The parameters for stepwise selection were p-value $< 1 \times 10^{-5}$, a collinearity cutoff of 0.9 and a distance of 10Mb. The SLE main cohort controls were used as the reference panel of SNPs to estimate LD.

Colocalisation. For each of the loci we found to have shared association between SLE and severe COVID-19, we used coloc [7] to perform locus-wide genetic colocalisation analysis. This returns the posterior probability that the two diseases share the same causal variant(s) in the region. We used standard coloc that assumes one casual variant and applied this to the marginal signals obtained using COJO. The SLE main cohort controls were used as the reference panel of SNPs to estimate the LD. GTEx v8 summary statistics across all tissues and eQTLGen whole blood summary statistics were used for the analysis.

For both the SLE/COVID-19 and the disease/eQTL colocalisation, signals were deemed to colocalize if: (1) when setting the prior probability of a SNP as associated with both traits = 5×10^{-05} (p_{12}, default = 1×10^{-05}), the posterior probability of colocalisation (PP_{H4}) > 0.5 and (2) when setting $p_{12} = 1 \times 10^{-5}$, the posterior probability of different causal variants (PP_{H3}) < 0.5 [62].

Haplotype analysis

Conditional haplotype-based association testing was performed on cases and controls in the European main SLE GWAS data and those with European/Hispanic/African American ancestry from the SLE Immunochip study using Plink [63]. An Independent effect for rs11085727 and rs2304256 was tested on the background of all the potential lead SNPs including rs34536443, rs74956615, rs11085727, rs2304256, rs12720356, rs280497, rs12720358 by using the PLINK '—hap-snps' command on the full set of SNPs with the '—independent-effect' option on [rs11085727, rs2304256] with '—chap'. Block estimations were performed within 200 kb. All variants with MAF < 0.001 were removed. The range of the 90% D-prime confidence interval was 0.70–0.98. The upper level for the confidence interval for historical recombination was 0.90 and strong LD pairs fraction was equal to 0.95.

Stepwise regression on individual level SLE data

In supplementary data analysis only, we analyzed the SLE GWAS individual level data for association using SNPTEST [45] fitting an additive model. This analysis was performed with three European SLE GWAS and the Immunochip data using SNPs with an imputation info score of > 0.7 and MAF > 0.01. Each dataset was included as a separate cohort in SNPTEST with covariates including principal components (PC1-3) for population structure and a discrete covariate for study. The same effect was assumed for all studies, as with a fixed effect meta-analysis. The results from the single-marker analysis using this approach were similar to those from the standard meta-analysis on the summary data (compare results for rs34536443 in **Tables F** and **I in S1 Text** for example) and the top associated SNPs was the same. We then ran forward stepwise selection by adding the top SNPs at each stage as a covariate to identify independently associated variants. This can be referred to as a one-stage approach to a meta-analysis using individual level data [64]. An alternative approach would be to use the top SNP at each step as a covariate in a regression analysis of each study separately and then meta-

analyze the results at each step (a two-stage approach). This would then rely on the meta-analysis approximation and not allow for an easy derivation of marginal signals when multiple independent associations are obtained.

Epigenetic modification and chromatin looping

Epigenetic modification and chromatin looping information were taken from resources available at the UCSC genome browser (http://genome.ucsc.edu). Enrichment of modifications to histone proteins (layered H3K27Ac, H3K4Me1, and H3K4Me3 track sets) determined by a ChIP-seq assay were from the ENCODE Consortium. Common dbSNP153 data (1000 Genomes phase 3, MAF > 0.01) was used, associated variants were highlighted. A highly filtered "double elite" subset of regulatory elements (including enhancers and promoters) and their inferred target genes in the plotting region were added as track sets, data was provided by the GeneHancer database [65].

Network and pathway enrichment analysis

Molecular interactions were obtained from the STRING 11.5 database [66]. All the potentially shared SLE and severe COVID-19 associated genes from the cross-trait meta-analysis (**Table 1**) were mapped to the whole network, connected nodes are shown in **Fig M in S1 Text** with indication of the type of interaction evidence. Gene Ontology (GO) process [67], local STRING network clusters [66], WikiPathways [68], KEGG pathway classification [69], Reactome Knowledgebase [70], and DISEASES database [71] were used for pathway enrichment of all the potentially shared genes.

Supporting information

S1 Text. Supplementary Results.
(DOCX)

Acknowledgments

We thank Nick Dand for reviewing draft versions of this paper.

Author Contributions

Conceptualization: Yuxuan Wang, Deborah S. Cunninghame Graham, David L. Morris, Timothy J. Vyse.

Data curation: Yuxuan Wang, Suri Guga, Kejia Wu, Phil Tombleson, Mary E. Comeau, Carl D. Langefeld, David L. Morris, Timothy J. Vyse.

Formal analysis: Yuxuan Wang, Suri Guga, Kejia Wu, Zoe Khaw, Konstantinos Tzoumkas, Phil Tombleson, Mary E. Comeau, Carl D. Langefeld, David L. Morris, Timothy J. Vyse.

Funding acquisition: David L. Morris, Timothy J. Vyse.

Investigation: Yuxuan Wang, Deborah S. Cunninghame Graham, David L. Morris, Timothy J. Vyse.

Methodology: Yuxuan Wang, David L. Morris, Timothy J. Vyse.

Project administration: Deborah S. Cunninghame Graham, David L. Morris, Timothy J. Vyse.

Resources: David L. Morris, Timothy J. Vyse.

Supervision: Deborah S. Cunninghame Graham, David L. Morris, Timothy J. Vyse.

Validation: Yuxuan Wang, David L. Morris, Timothy J. Vyse.

Visualization: Yuxuan Wang, David L. Morris, Timothy J. Vyse.

Writing – original draft: Yuxuan Wang, Deborah S. Cunninghame Graham, David L. Morris, Timothy J. Vyse.

Writing – review & editing: Yuxuan Wang, Deborah S. Cunninghame Graham, David L. Morris, Timothy J. Vyse.

References

1. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic mechanisms of critical illness in COVID-19. Nature. 2021; 591(7848):92–8. Epub 2020/12/12. https://doi.org/10.1038/s41586-020-03065-y PMID: 33307546.

2. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med. 2008; 358(9):900–9. https://doi.org/10.1056/NEJMoa0707865 PMID: 18204098.

3. Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet. 2008; 40(2):204–10. https://doi.org/10.1038/ng.81 WOS:000252732900019. PMID: 18204446

4. Bentham J, Morris DL, Cunninghame Graham DS, Pinder CL, Tombleson P, Behrens TW, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. Nat Genet. 2015; 47(12):1457–64. https://doi.org/10.1038/ng.3434 PMID: 26502338; PubMed Central PMCID: PMC4668589.

5. Langefeld CD, Ainsworth HC, Cunninghame Graham DS, Kelly JA, Comeau ME, Marion MC, et al. Transancestral mapping and genetic load in systemic lupus erythematosus. Nat Commun. 2017; 8:16021. https://doi.org/10.1038/ncomms16021 PMID: 28714469; PubMed Central PMCID: PMC5520018.

6. Initiative C-HG. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. Eur J Hum Genet. 2020; 28(6):715–8. Epub 2020/05/15. https://doi.org/10.1038/s41431-020-0636-6 PMID: 32404885; PubMed Central PMCID: PMC7220587.

7. Wallace C. Statistical testing of shared genetic control for potentially related traits. Genet Epidemiol. 2013; 37(8):802–13. Epub 2013/11/15. https://doi.org/10.1002/gepi.21765 PMID: 24227294; PubMed Central PMCID: PMC4158901.

8. Sigurdsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G, et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet. 2005; 76(3):528–37. Epub 2005/01/20. https://doi.org/10.1086/428480 PMID: 15657875; PubMed Central PMCID: PMC1196404.

9. Cunninghame Graham DS, Morris DL, Bhangale TR, Criswell LA, Syvanen AC, Ronnblom L, et al. Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus. PLoS Genet. 2011; 7(10):e1002341. Epub 2011/11/03. https://doi.org/10.1371/journal.pgen.1002341 PMID: 22046141; PubMed Central PMCID: PMC3203198.

10. Suarez-Gestal M, Calaza M, Endreffy E, Pullmann R, Ordi-Ros J, Sebastiani GD, et al. Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study. Arthritis Res Ther. 2009; 11(3):R69. Epub 2009/05/16. https://doi.org/10.1186/ar2698 PMID: 19442287; PubMed Central PMCID: PMC2714115.

11. Hellquist A, Jarvinen TM, Koskenmies S, Zucchelli M, Orsmark-Pietras C, Berglind L, et al. Evidence for genetic association and interaction between the TYK2 and IRF5 genes in systemic lupus erythematosus. J Rheumatol. 2009; 36(8):1631–8. Epub 2009/07/02. https://doi.org/10.3899/jrheum.081160 PMID: 19567624.

12. Cunninghame Graham DS, Akil M, Vyse TJ. Association of polymorphisms across the tyrosine kinase gene, TYK2 in UK SLE families. Rheumatology (Oxford). 2007; 46(6):927–30. Epub 2007/03/27. https://doi.org/10.1093/rheumatology/kel449 PMID: 17384181.

13. Dendrou CA, Cortes A, Shipman L, Evans HG, Attfield KE, Jostins L, et al. Resolving TYK2 locus geno-type-to-phenotype differences in autoimmunity. Sci Transl Med. 2016; 8(363):363ra149. Epub 2016/11/04. https://doi.org/10.1126/scitranslmed.aag1974 PMID: 27807284; PubMed Central PMCID: PMC5737835.

14. Võsa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, et al. Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. BioRxiv. 2018:447367.

15. Li Z, Rotival M, Patin E, Michel F, Pellegrini S. Two common disease-associated TYK2 variants impact exon splicing and TYK2 dosage. PLoS One. 2020; 15(1):e0225289. Epub 2020/01/22. https://doi.org/10.1371/journal.pone.0225289 PMID: 31961910; PubMed Central PMCID: PMC6974145.

16. El-Sherbiny YM, Psarras A, Md Yusof MY, Hensor EMA, Tooze R, Doody G, et al. A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. Sci Rep. 2018; 8(1):5793. Epub 2018/04/13. https://doi.org/10.1038/s41598-018-24198-1 PMID: 29643425; PubMed Central PMCID: PMC5895784.

17. Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurle MR, et al. Phenome-wide Mendelian ran-domization mapping the influence of the plasma proteome on complex diseases. Nat Genet. 2020; 52 (10):1122–31. Epub 2020/09/09. https://doi.org/10.1038/s41588-020-0682-6 PMID: 32895551; PubMed Central PMCID: PMC7610464.

18. Nogueira M, Puig L, Torres T. JAK Inhibitors for Treatment of Psoriasis: Focus on Selective TYK2 Inhib-itors. Drugs. 2020; 80(4):341–52. Epub 2020/02/06. https://doi.org/10.1007/s40265-020-01261-8 PMID: 32020553.

19. Boisson-Dupuis S, Ramirez-Alejo N, Li Z, Patin E, Rao G, Kerner G, et al. Tuberculosis and impaired IL-23-dependent IFN-gamma immunity in humans homozygous for a common TYK2 missense variant. Sci Immunol. 2018; 3(30). Epub 2018/12/24. https://doi.org/10.1126/sciimmunol.aau8714 PMID: 30578352; PubMed Central PMCID: PMC6341984.

20. Kerner G, Ramirez-Alejo N, Seeleuthner Y, Yang R, Ogishi M, Cobat A, et al. Homozygosity for TYK2 P1104A underlies tuberculosis in about 1% of patients in a cohort of European ancestry. Proc Natl Acad Sci U S A. 2019; 116(21):10430–4. Epub 2019/05/10. https://doi.org/10.1073/pnas.1903561116 PMID: 31068474; PubMed Central PMCID: PMC6534977.

21. Zuo H, Cattani-Cavalieri I, Musheshe N, Nikolaev VO, Schmidt M. Phosphodiesterases as therapeutic targets for respiratory diseases. Pharmacol Ther. 2019; 197:225–42. Epub 2019/02/14. https://doi.org/10.1016/j.pharmthera.2019.02.002 PMID: 30759374.

22. Castel SE, Cervera A, Mohammadi P, Aguet F, Reverter F, Wolman A, et al. Modified penetrance of coding variants by cis-regulatory variation contributes to disease risk. Nat Genet. 2018; 50(9):1327–34. Epub 2018/08/22. https://doi.org/10.1038/s41588-018-0192-y PMID: 30127527; PubMed Central PMCID: PMC6119105.

23. Coperchini F, Chiovato L, Croce L, Magri F, Rotondi M. The cytokine storm in COVID-19: An overview of the involvement of the chemokine/chemokine-receptor system. Cytokine Growth Factor Rev. 2020; 53:25–32. Epub 2020/05/25. https://doi.org/10.1016/j.cytogfr.2020.05.003 PMID: 32446778; PubMed Central PMCID: PMC7211650.

24. Thomson TM, Toscano-Guerra E, Casis E, Paciucci R. C1 esterase inhibitor and the contact system in COVID-19. Br J Haematol. 2020; 190(4):520–4. Epub 2020/06/13. https://doi.org/10.1111/bjh.16938 PMID: 32531085; PubMed Central PMCID: PMC7323335.

25. Stappers MHT, Clark AE, Aimanianda V, Bidula S, Reid DM, Asamaphan P, et al. Recognition of DHN-melanin by a C-type lectin receptor is required for immunity to Aspergillus. Nature. 2018; 555 (7696):382–6. Epub 2018/03/01. https://doi.org/10.1038/nature25974 PMID: 29489751; PubMed Cen-tral PMCID: PMC5857201.

26. Lopez Robles MD, Pallier A, Huchet V, Le Texier L, Remy S, Braudeau C, et al. Cell-surface C-type lec-tin-like receptor CLEC-1 dampens dendritic cell activation and downstream Th17 responses. Blood Adv. 2017; 1(9):557–68. Epub 2018/01/04. https://doi.org/10.1182/bloodadvances.2016002360 PMID: 29296975; PubMed Central PMCID: PMC5728597 interests.

27. Hambleton S, Salem S, Bustamante J, Bigley V, Boisson-Dupuis S, Azevedo J, et al. IRF8 mutations and human dendritic-cell immunodeficiency. N Engl J Med. 2011; 365(2):127–38. Epub 2011/04/29. https://doi.org/10.1056/NEJMoa1100066 PMID: 21524210; PubMed Central PMCID: PMC3136554.

28. Calabrese LH, Winthrop K, Strand V, Yazdany J, Walter JE. Type I interferon, anti-interferon antibodies, and COVID-19. Lancet Rheumatol. 2021. Epub 2021/03/04. https://doi.org/10.1016/S2665-9913(21) 00034-5 PMID: 33655222; PubMed Central PMCID: PMC7906736.

29. Broggi A, Ghosh S, Sposito B, Spreafico R, Balzarini F, Lo Cascio A, et al. Type III interferons disrupt the lung epithelial barrier upon viral recognition. Science. 2020; 369(6504):706–12. Epub 2020/06/13. https://doi.org/10.1126/science.abc3545 PMID: 32527925; PubMed Central PMCID: PMC7292499.

30. Hadjadj J, Yatim N, Barnabei L, Corneau A, Boussier J, Smith N, et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science. 2020; 369(6504):718–24. Epub 2020/07/15. https://doi.org/10.1126/science.abc6027 PMID: 32661059; PubMed Central PMCID: PMC7402632.

31. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A. 2003; 100(5):2610–5. Epub 2003/02/27. https://doi.org/10.1073/pnas.0337679100 PMID: 12604793; PubMed Central PMCID: PMC151388.

32. Kirou KA, Lee C, George S, Louca K, Peterson MG, Crow MK. Activation of the interferon-alpha pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. Arthritis Rheum. 2005; 52(5):1491–503. Epub 2005/05/10. https://doi.org/10.1002/art.21031 PMID: 15880830.

33. Kirou KA, Lee C, George S, Louca K, Papagiannis IG, Peterson MG, et al. Coordinate overexpression of interferon-alpha-induced genes in systemic lupus erythematosus. Arthritis Rheum. 2004; 50 (12):3958–67. Epub 2004/12/14. https://doi.org/10.1002/art.20798 PMID: 15593221.

34. Morand EF, Furie R, Tanaka Y, Bruce IN, Askanase AD, Richez C, et al. Trial of Anifrolumab in Active Systemic Lupus Erythematosus. N Engl J Med. 2020; 382(3):211–21. Epub 2019/12/19. https://doi.org/10.1056/NEJMoa1912196 PMID: 31851795.

35. Ronnblom L, Eloranta ML, Alm GV. The type I interferon system in systemic lupus erythematosus. Arthritis Rheum. 2006; 54(2):408–20. Epub 2006/02/01. https://doi.org/10.1002/art.21571 PMID: 16447217.

36. Okin D, Medzhitov R. Evolution of inflammatory diseases. Curr Biol. 2012; 22(17):R733–40. Epub 2012/09/15. https://doi.org/10.1016/j.cub.2012.07.029 PMID: 22975004; PubMed Central PMCID: PMC3601794.

37. Zhernakova A, Elbers CC, Ferwerda B, Romanos J, Trynka G, Dubois PC, et al. Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. Am J Hum Genet. 2010; 86(6):970–7. Epub 2010/06/22. https://doi.org/10.1016/j.ajhg.2010.05.004 PMID: 20560212; PubMed Central PMCID: PMC3032060.

38. Barreiro LB, Quintana-Murci L. From evolutionary genetics to human immunology: how selection shapes host defence genes. Nat Rev Genet. 2010; 11(1):17–30. Epub 2009/12/03. https://doi.org/10.1038/nrg2698 PMID: 19953080.

39. Schoenborn JR, Wilson CB. Regulation of interferon-gamma during innate and adaptive immune responses. Adv Immunol. 2007; 96:41–101. Epub 2007/11/06. https://doi.org/10.1016/S0065-2776(07)96002-2 PMID: 17981204.

40. Willcocks LC, Carr EJ, Niederer HA, Rayner TF, Williams TN, Yang W, et al. A defuncting polymorphism in FCGR2B is associated with protection against malaria but susceptibility to systemic lupus erythematosus. Proc Natl Acad Sci U S A. 2010; 107(17):7881–5. Epub 2010/04/14. https://doi.org/10.1073/pnas.0915133107 PMID: 20385827; PubMed Central PMCID: PMC2867866.

41. Mogensen TH. Pathogen recognition and inflammatory signaling in innate immune defenses. Clin Microbiol Rev. 2009; 22(2):240–73, Table of Contents. Epub 2009/04/16. https://doi.org/10.1128/CMR.00046-08 PMID: 19366914; PubMed Central PMCID: PMC2668232.

42. Bentham J, Morris DL, Graham DSC, Pinder CL, Tombleson P, Behrens TW, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. Nature genetics. 2015; 47(12):1457–64. https://doi.org/10.1038/ng.3434 PMID: 26502338; PubMed Central PMCID: PMC4668589.

43. International Consortium for Systemic Lupus Erythematosus G, Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet. 2008; 40(2):204–10. Epub 2008/01/22. https://doi.org/10.1038/ng.81 PMID: 18204446; PubMed Central PMCID: PMC3712260.

44. Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. Haplotype estimation using sequencing reads. Am J Hum Genet. 2013; 93(4):687–96. Epub 2013/10/08. https://doi.org/10.1016/j.ajhg.2013.09.002 PMID: 24094745; PubMed Central PMCID: PMC3791270.

45. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet. 2010; 11(7):499–511. Epub 2010/06/03. https://doi.org/10.1038/nrg2796 PMID: 20517342.

46. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007; 39(7):906–13. Epub 2007/06/19. https://doi.org/10.1038/ng2088 PMID: 17572673.

47. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015; 526(7571):68–74. Epub 2015/10/04. https://doi.org/10.1038/nature15393 PMID: 26432245; PubMed Central PMCID: PMC4750478.

48. Morris DL, Sheng YJ, Zhang Y, Wang YF, Zhu ZW, Tombleson P, et al. Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. Nat Genet. 2016; 48(8):940–+. https://doi.org/10.1038/ng.3603 WOS:000380755100019. PMID: 27399966

49. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26(17):2190–1. Epub 2010/07/10. https://doi.org/10.1093/bioinformatics/btq340 PMID: 20616382; PubMed Central PMCID: PMC2922887.

50. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, et al. Assessing the impact of population stratification on genetic association studies. Nat Genet. 2004; 36(4):388–93. Epub 2004/03/31. https://doi.org/10.1038/ng1333 PMID: 15052270.

51. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015; 47(3):291–5. Epub 2015/02/03. https://doi.org/10.1038/ng.3211 PMID: 25642630; PubMed Central PMCID: PMC4495769.

52. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015; 47(11):1236–41. Epub 2015/09/29. https://doi.org/10.1038/ng.3406 PMID: 26414676; PubMed Central PMCID: PMC4797329.

53. Shi H, Mancuso N, Spendlove S, Pasaniuc B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. Am J Hum Genet. 2017; 101(5):737–51. Epub 2017/11/04. https://doi.org/10.1016/j.ajhg.2017.09.022 PMID: 29100087; PubMed Central PMCID: PMC5673668.

54. Replication DIG, Meta-analysis C, Asian Genetic Epidemiology Network Type 2 Diabetes C, South Asian Type 2 Diabetes C, Mexican American Type 2 Diabetes C, Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in muylti-Ethnic Samples C, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat Genet. 2014; 46(3):234–44. Epub 2014/02/11. https://doi.org/10.1038/ng.2897 PMID: 24509480; PubMed Central PMCID: PMC3969612.

55. COVID-19 Host Genetics Initiative, Ganna A. Mapping the human genetic architecture of COVID-19: an update. medRxiv. 2022:2021.11.08.21265944. https://doi.org/10.1101/2021.11.08.21265944

56. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010; 26(18):2336–7. Epub 2010/07/17. https://doi.org/10.1093/bioinformatics/btq419 PMID: 20634204; PubMed Central PMCID: PMC2935401.

57. Consortium GT. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020; 369(6509):1318–30. Epub 2020/09/12. https://doi.org/10.1126/science.aaz1776 PMID: 32913098; PubMed Central PMCID: PMC7737656.

58. Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. Nat Commun. 2017; 8:14357. Epub 2017/02/28. https://doi.org/10.1038/ncomms14357 PMID: 28240269; PubMed Central PMCID: PMC5333359.

59. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature. 2018; 558(7708):73–9. Epub 2018/06/08. https://doi.org/10.1038/s41586-018-0175-2 PMID: 29875488; PubMed Central PMCID: PMC6697541.

60. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ATC, Replication DIG, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012; 44(4):369–75, S1-3. Epub 2012/03/20. https://doi.org/10.1038/ng.2213 PMID: 22426310; PubMed Central PMCID: PMC3593158.

61. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011; 88(1):76–82. Epub 2010/12/21. https://doi.org/10.1016/j.ajhg.2010.11.011 PMID: 21167468; PubMed Central PMCID: PMC3014363.

62. Kousathanas A, Pairo-Castineira E, Rawlik K, Stuckey A, Odhams CA, Walker S, et al. Whole genome sequencing reveals host factors underlying critical Covid-19. Nature. 2022. Epub 2022/03/08. https://doi.org/10.1038/s41586-022-04576-6 PMID: 35255492.

63. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3):559–75. Epub 2007/08/19. https://doi.org/10.1086/519795 PMID: 17701901; PubMed Central PMCID: PMC1950838.

64. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Stat Med. 2017; 36(5):855–75. Epub 2016/10/18. https://doi.org/10.1002/sim.7141 PMID: 27747915; PubMed Central PMCID: PMC5297998.

65. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford). 2017;2017. Epub 2017/06/13. https://doi.org/10.1093/database/bax028 PMID: 28605766; PubMed Central PMCID: PMC5467550.

66. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021; 49(D1):D605–D12. Epub 2020/11/26. https://doi.org/10.1093/nar/gkaa1074 PMID: 33237311; PubMed Central PMCID: PMC7779004.

67. Gene Ontology C. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. 2021; 49(D1):D325–D34. Epub 2020/12/09. https://doi.org/10.1093/nar/gkaa1113 PMID: 33290552; PubMed Central PMCID: PMC7779012.

68. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. Nucleic Acids Res. 2021; 49(D1):D613–D21. Epub 2020/11/20. https://doi.org/10.1093/nar/gkaa1024 PMID: 33211851; PubMed Central PMCID: PMC7779061.

69. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. Nucleic Acids Res. 2019; 47(D1):D590–D5. Epub 2018/10/16. https://doi.org/10.1093/nar/gky962 PMID: 30321428; PubMed Central PMCID: PMC6324070.

70. Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. Nucleic Acids Res. 2022; 50(D1):D687–D92. Epub 2021/11/18. https://doi.org/10.1093/nar/gkab1028 PMID: 34788843; PubMed Central PMCID: PMC8689983.

71. Pletscher-Frankild S, Palleja A, Tsafou K, Binder JX, Jensen LJ. DISEASES: text mining and data integration of disease-gene associations. Methods. 2015; 74:83–9. Epub 2014/12/09. https://doi.org/10.1016/j.ymeth.2014.11.020 PMID: 25484339.