Article

mGWAS-Explorer: Linking SNPs, Genes, Metabolites, and Diseases for Functional Insights

Le Chang [1], Guangyan Zhou [2], Huiting Ou [1] and Jianguo Xia [1,2,*]

1   Department of Human Genetics, McGill University, Montreal, QC H3A 0C7, Canada;
     le.chang@mail.mcgill.ca (L.C.); huiting.ou@mail.mcgill.ca (H.O.)
2   Institute of Parasitology, McGill University, Montreal, QC H9X 3V9, Canada; guangyan.zhou@mail.mcgill.ca
*   Correspondence: jeff.xia@mcgill.ca; Tel.: +1-(514)-398-8668

Abstract: Tens of thousands of single-nucleotide polymorphisms (SNPs) have been identified to be significantly associated with metabolite abundance in over 65 genome-wide association studies with metabolomics (mGWAS) to date. Obtaining mechanistic or functional insights from these associations for translational applications has become a key research area in the mGWAS community. Here, we introduce mGWAS-Explorer, a user-friendly web-based platform to help connect SNPs, metabolites, genes, and their known disease associations via powerful network visual analytics. The application of the mGWAS-Explorer was demonstrated using a COVID-19 and a type 2 diabetes case studies.

Keywords: mGWAS; SNP; mQTL; metabolomics; pleiotropy; cross-phenotype association analysis; network

1. Introduction

Genome-wide association studies (GWAS) have identified hundreds of thousands of genetic loci associated with complex diseases. These associations have improved our understanding of the genetic architecture of human diseases [1]. However, translations of these associations into biomedical or pharmaceutical applications have been limited, as the majority of the disease-associated loci reside in the non-coding regions of the genome with no obvious gene targets [2]. Technology advancements in mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy have allowed GWAS to be carried out with metabolomics (mGWAS) to study genetically influenced metabotypes (GIMs) [3,4]. mGWAS have been very successful in identifying metabolite quantitative trait loci (mQTLs). An mQTL is a locus that is associated with variations in metabolite abundance [3]. In addition to having larger effects compared to loci identified in GWAS of clinical phenotypes in general, many mQTLs can map to genes encoding enzymes or transporters, providing biochemical context for these variations [3,5]. Leveraging these mQTLs to improve our knowledge of metabolism and metabolic disorders for translational applications has become a key research area in the mGWAS community.

mQTLs are characterized by polygenicity and pleiotropy [6,7]. Polygenicity means a single trait is influenced by multiple genes, whereas pleiotropy refers to the phenomenon in which genetic variants affect multiple traits or diseases [8,9]. For instance, one single-nucleotide polymorphism (SNP) can directly affect multiple traits, or different SNPs in high linkage disequilibrium (LD) may exist for more than one trait. Pleiotropy may provide insights into the cause of trait comorbidity and help determine the direction of causal relationships by pointing to shared genetic mechanisms [8,10]. Various strategies have been developed to examine genetic relationships between multiple phenotypes [11–18]. For example, LD score regression is a popular method to assess the genetic correlations of pairwise traits using GWAS summary statistics [13]. Colocalization is another strategy aiming to identify causal variants at two overlapping association signals [14]. These methods have successfully identified pleiotropic genomic regions and addressed fundamental

research questions regarding the polygenicity of traits, but it is challenging to scale up these methods to study hundreds of traits at once.

Comprehensive annotations are necessary in order to gain functional insights into SNP–metabolite associations. Many resources are available to support SNP to gene annotation, such as VEP and SNiPA [19,20]. For metabolite annotation, there is a wealth of biochemical knowledge on enzymatic reactions as well as transporters and their substrates. In addition, mapping GWAS results to the protein–protein interaction (PPI) network can potentially augment the association signals [21].

Recently, cross-phenotype association analysis has gained increasing attention [22–26]. It takes a specific SNP and searches for associations across a range of molecular or disease phenotypes, which allows for elucidations of complex networks between phenotypes and their genetic loci. A variety of databases currently exist to store the genotype–phenotype association datasets, including GWAS Catalog [27], PhenoScanner [28], OpenGWAS [29], Open Targets Genetics [30], PheLiGe [31], DisGeNET [32], as well as specific tools for mGWAS, such as the metabolomics GWAS server [33,34]. Valuable tools currently exist to allow users to perform cross-phenotype analysis [35–40]. However, these tools do not offer extra support beyond displaying and visualization, and none of them are dedicated to mGWAS.

There is a clear demand for dedicated bioinformatics resources to support mGWAS data analysis and interpretation. Our overall assumption is that by developing a centralized place for mGWAS datasets and performing deep annotation of the underlying SNPs and metabolites, users can gain valuable functional insights into the statistical associations identified from mGWAS results.

A network is a valuable approach to depict mGWAS results and allows the dissection of polygenicity and pleiotropy. Heterogeneous networks comprising various types of nodes (e.g., SNPs, genes, metabolites, and diseases) and edges (e.g., statistical or biochemical associations) have been remarkably useful in depicting the complex interplay across biological entities [41]. These network-based approaches have the potential to identify and prioritize therapeutic candidates to generate new hypotheses [42].

Here, we introduce mGWAS-Explorer (https://www.mgwas.ca (accessed on 1 May 2022), a user-friendly web-based platform for network-based integrative analysis and visual exploration of SNPs, genes, metabolites, and diseases. Its key features include:

- Comprehensive collection and deep annotation of SNP–metabolite associations based on data from the 65 mGWAS to date.
- Support for SNP-based, gene-based, and metabolite-based network generation to facilitate interpreting results.
- Powerful network visual analytics system facilitating interactive exploration and built-in topological and functional enrichment analysis.

mGWAS-Explorer also includes a comprehensive list of frequently asked questions (FAQs) and detailed tutorials. Together, these features comprise a powerful platform for functional interpretation and cross-phenotype association analysis of mGWAS datasets.

2. Results

2.1. Overview of the Curated mGWAS Datasets

Since the first study in 2008 [5], mGWAS with increasing sample sizes and various populations have been conducted, resulting in a continued increase in SNP–metabolite associations. We systematically curated the public mGWAS datasets to date. A summary table of these mGWAS datasets can be found in Table 1. Please note the *p*-value cutoffs are based on significance thresholds of the original studies, as the *p*-values and effect sizes of SNP-metabolite associations may differ across different studies due to the differences in sample sizes, population types, or the metabolomics platforms [4,7].

Table 1. A summary of the mGWAS datasets in mGWAS-Explorer.

Sample Type	Study #	* Metabolite #	** Metabolite Ratio #	SNP #	SNP–Metabolite Associations #
Blood	57	3992	1265	67,570	30,3090
Urine	5	271	1123	6877	9647
Saliva	1	14	0	1364	1454
Cerebrospinal fluid (CSF)	1	15	0	1178	1182
Mitochondria	1	0	390	194	404
Sum (unique)	65	4147	2388	73,737	313,720

* Metabolite number includes both targeted (compound names) and untargeted measures (feature IDs, such as '391.2859_3.774' based on mass to charge ratio and retention time. The total number of such feature IDs is 2464). ** Metabolite ratios (metabolite A/metabolite B) can be useful as they may reflect the biochemical conversion of metabolites and thus enhance the association signals. The # sign indicates size or total number [43].

2.2. Overview of the mGWAS-Explorer

The main workflow of mGWAS-Explorer is summarized in Figure 1. There are three major steps—data input, network creation, and network visual analytics. To begin, users can enter through one of the five modules based on input type. The 'SNP' module allows users to explore SNP–gene, SNP–metabolite, or SNP–disease networks. We provide support for LD proxy search to maximize the search by looking for SNPs in LDs with the input SNPs. After SNP to gene mapping, users can choose to include PPIs in the networks. The 'Gene' module maps genes to SNPs that are significant in the mGWAS datasets, or to metabolites (i.e., through encoding enzymes or transporters), or known associated diseases. The 'Metabolite' module maps metabolites to associated SNPs, genes, or diseases. The 'Search' module allows users to search known SNP–gene associations in mGWAS datasets, while the 'Browse' module allows users to browse individual mGWAS data in a 3D Manhattan plot or a network view. To start the analysis, users must click a circular button from the mGWAS-Explorer homepage to enter the corresponding data upload page. Various functions are available to allow users to refine the networks. In the last step, the results are shown as interactive networks for visual exploration. Users can easily search, explore, highlight, or perform functional enrichment analysis on the nodes of interest. For instance, double-clicking an edge will display the evidence supporting the relationships. The network results can be downloaded in PNG, SVG format, or as graph files.
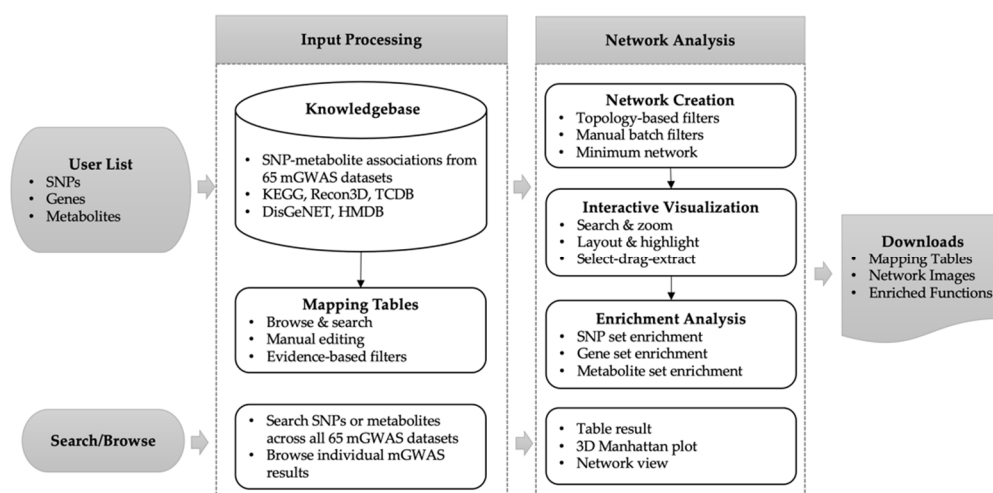

Figure 1. Overview of mGWAS-Explorer workflow. Users can upload different data types. The input will be mapped to the underlying knowledgebases to create mapping tables and networks. The visualization page allows users to intuitively explore the networks to identify important associations as well as to perform topology or functional analysis.

2.3. Analysis Workflow

There are five modules in mGWAS-Explorer corresponding to the five different types of input. Users can upload a list of SNPs, metabolites, or genes (Figure 2a–c); browse individual mGWAS dataset in a 3D Manhattan plot (Figure 2d); or search significant SNP-metabolite associations across all mGWAS datasets (Figure 2e).
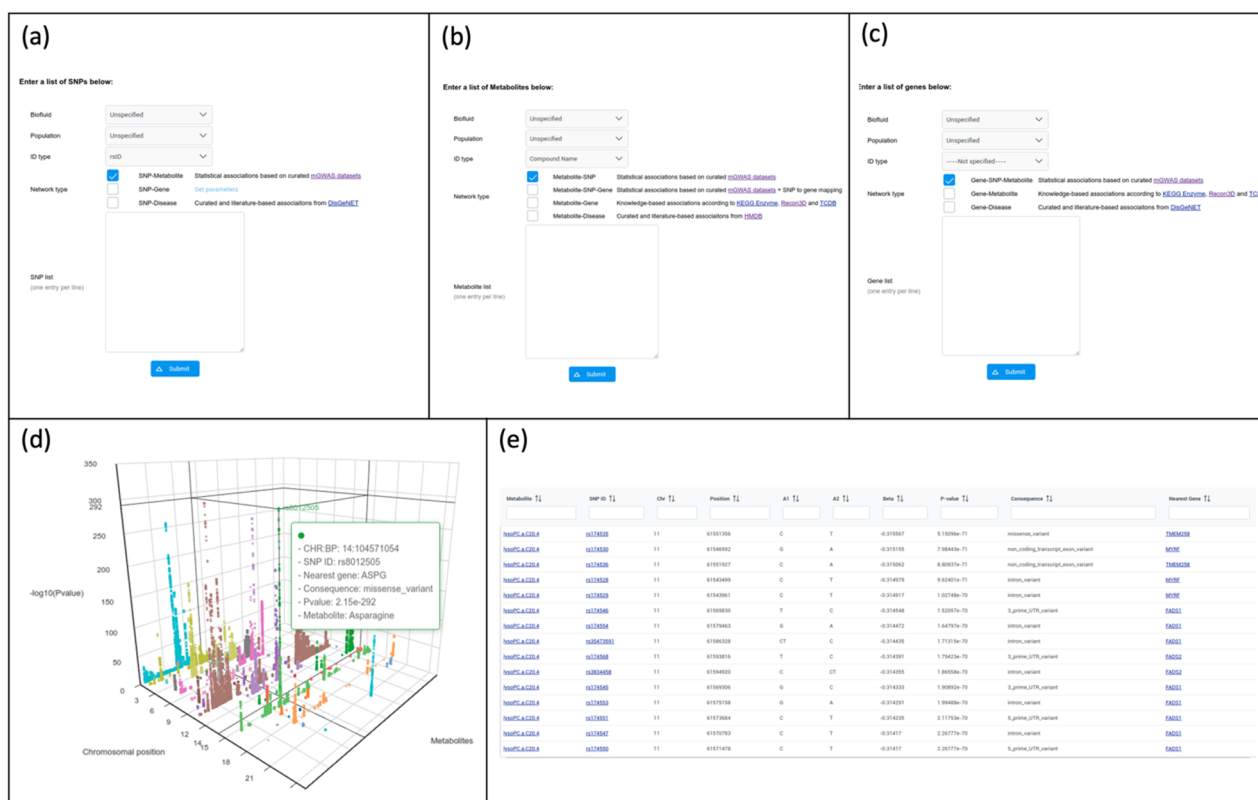

Figure 2. Screenshots of upload pages for SNP (**a**), metabolite (**b**), and gene (**c**) modules. (**d**) A screenshot of a 3D Manhattan plot from the 'Browse' module based on the data from Lotta et al. [7]. The x-axis is the position of the SNPs, and the y-axis represents different metabolites, while the z-axis corresponds to the significance of the association; (**e**) a screenshot showing the table view in the 'Browse' module.

2.3.1. Search and Browse

The 'Search' module supports searching the association results of the curated 65 mGWAS publications. Meanwhile, the 'Browse' module allows users to visually explore the data in a 3D Manhattan plot or network view. The 3D Manhattan plot strengthens the exploration of metabolome-wide pleiotropy at the genome-wide level [39]. Users can mouse-over a dot to see the SNP annotation, including the rsID, CHR:BP, nearest gene, most severe consequence, *p*-value, and metabolite name (Figure 2d).

2.3.2. From SNPs to Networks

Network-based approaches have become increasingly applied to identify shared genetic underpinnings (i.e., pleiotropy) in GWAS, where nodes are SNPs or phenotypes (e.g., metabolites or diseases) and edges represent significant associations [11]. The 'SNPs' module supports SNP–metabolite (or metabolite ratios), and SNP–disease network analysis. Optionally, LD proxy search (i.e., population type and r^2) could be performed to maximize the results. Users also have the flexibility to include PPI networks, as well as to filter on biofluid or population types (Figure 2a).

2.3.3. From Metabolites to Networks

For many GIMs, metabolites can be functionally connected to enzymes, or transporters [3]. The 'Metabolites' module allows users to perform either statistical-based or knowledge-based metabolite-gene associations as well as metabolite-disease associations. Users can upload a list of metabolites from the upload page (Figure 2b). mGWAS-Explorer currently accepts either HMDB ID, KEGG ID or compound name. The uploaded list is then mapped to genes, SNPs, or diseases for network creation and subsequent visualization.

2.3.4. From Genes to Networks

The nearest gene mapping approach is suggested to be an effective indicator of true positive genes for mQTLs [44]. In mGWAS-Explorer, users can upload their gene lists in the 'Genes' module, the reversed nearest-gene mapping will be automatically performed and return SNPs that are significant in the mGWAS (Figure 2c). The network output will include genes, SNPs, and metabolites. Alternatively, users can perform gene–metabolite mapping via biochemical knowledge or to the associated diseases.

2.4. Case Studies

2.4.1. COVID-19 Case Study

The host genetic variation is known to influence the severity of SARS-CoV-2 infection [45–49] and the blood metabolomics can reveal biomarkers for disease diagnosis and prognosis [50,51]. However, understanding mechanisms that link genetic variation to metabolism and clinical endpoints remains an important challenge. Therefore, we applied mGWAS-Explorer to a list of SNPs identified from a GWAS of severe COVID-19 [47] to provide insights into the shared genetic architecture of diseases and intermediate metabolic phenotypes. We used a suggestive significant association *p*-value threshold (1×10^{-5}) for mGWAS-Explorer, resulting in 19 SNPs after LD clumping. mGWAS-Explorer revealed that the SNPs at the *ABO* (alpha 1-3-N-acetylgalactosaminyltransferase and alpha 1-3-galactosyltransferase) locus were in high LD ($r^2 > 0.8$) with numerous other SNPs in this region associated with multiple metabolites and other human diseases, such as leucylalanine, citric acid [51], malaria [52], ischemic stroke [53], and venous thrombosis [54]. The blood type locus *ABO* has been linked to the risk of COVID-19 in several studies [47,55]. Multiple hypotheses have been proposed to explain the mechanism, such as anti-A and/or anti-B antibodies against corresponding antigens, or the glycosyltransferase activity [56]. mGWAS-Explorer provided insights into these possible mechanisms, which identified associations of *ABO* variants with levels and the ratios of fibrinogen A-α peptides (e.g., ADpSGEGDFXAEGGGVR) and venous thromboembolism (Figure 3a). Fibrinogen plays a role in blood clotting [57]. Therefore, the association between *ABO* variants with fibrinogen may suggest that *ABO* influences COVID-19 via regulating thrombosis, which provided a functional explanation for the observed association of *ABO* with COVID-19 risk. Indeed, studies have reported that COVID-19 is associated with an increased risk of thromboembolism [58]. Therefore, we sought to investigate whether the association between fibrinogen A-α peptide-associated loci could provide additional insights into the underlying pathophysiology of COVID-19. Interestingly, mGWAS-Explorer revealed variants in *ENPEP* (glutamyl aminopeptidase) and *FUT2* (fucosyltransferase 2) genes are associated with levels and/or ratios of fibrinogen A-α peptides. Additionally, *FUT2* gene was also identified in the PPI network with the *ABO* gene (Figure 3a). In fact, *ENPEP* was discovered to be a candidate co-receptor for the coronavirus SARS-CoV-2 [59] and individuals with an inactivating *FUT2* mutations were more likely to develop a less severe form of the COVID-19 disease [60]. In summary, mGWAS-Explorer supports the evidence that *ABO*, *ENPEP*, and *FUT2* may be candidate genes and discovered fibrinogen A-α peptides as potential biomarkers for COVID-19 disease.
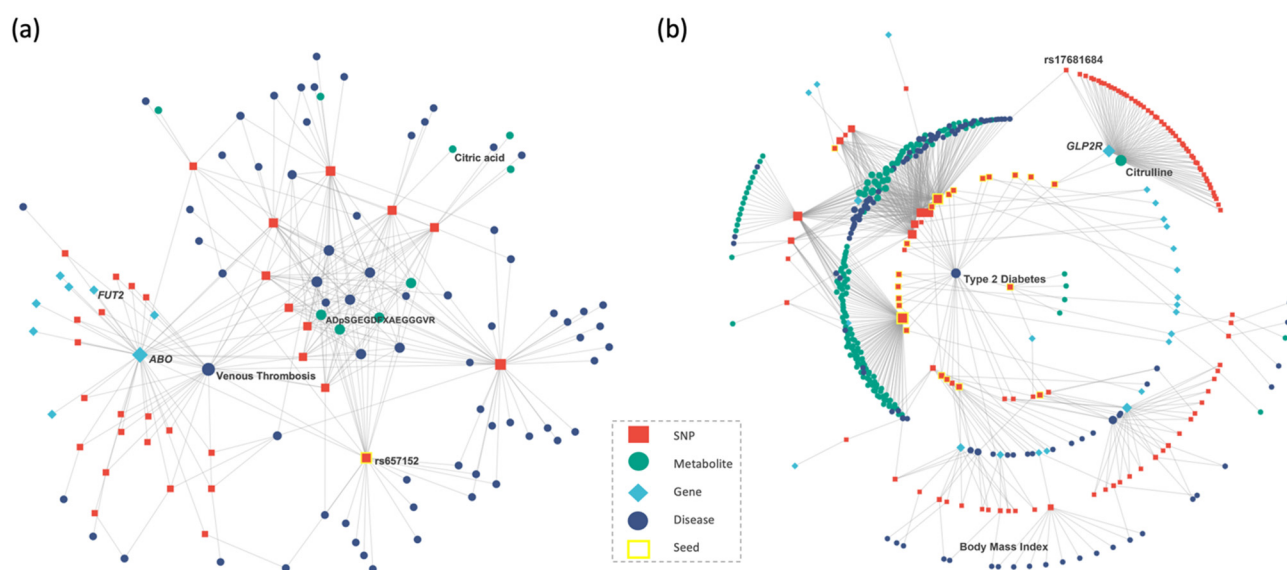
Figure 3. Screenshots of network results for the (**a**) COVID-19 case study and (**b**) type 2 diabetes case study. Each node represents either a SNP (orange square), a metabolite (green circle), a gene (blue diamond), a disease (dark blue circle), or a seed node (yellow square outline). Each edge is either an association between one SNP and one metabolite, an association between one SNP and one disease, a positional mapping of SNP to gene, or a protein–protein interaction. The size of a node is proportional to the number of other nodes connected to it.

2.4.2. Type 2 Diabetes Case Study

Around 250 genomic regions have been associated with type 2 diabetes (T2D) susceptibility in genome-wide association studies, some studies have highlighted the link to metabolomic profiles [7,61,62]. We applied mGWAS-Explorer to a list of SNPs from a published GWAS of T2D [63] in an attempt to examine shared genetic signals with circulating metabolites. Notably, mGWAS-Explorer confirmed the associations between citrulline metabolites, T2D, body mass index (Figure 3b), and identified the missense rs17681684 variant for citrulline in the *GLP2R* (glucagon like peptide 2 receptor) gene as reported by Lotta et al. [7]. Additionally, it identified shared genetic signals between T2D, coronary artery disease, and cholesterol levels at the *ABO* locus. Indeed, previous epidemiological studies have reported that the associations of *ABO* group with coronary artery diseases are mediated by cholesterol [64], although the evidence regarding associations between *ABO* blood group with type 2 diabetes were not consistent [65–68]. Thus, further studies are required to identify the associations between *ABO* variants, T2D, and cholesterol levels. Furthermore, mGWAS-Explorer also revealed metabolites levels and their ratios identified in the previous COVID-19 case study shared associations with T2D loci. In fact, multiple studies have reported the comorbidity of T2D and COVID-19 [69–71]. In brief, analyzing the T2D cross-phenotype associations with metabolites and other diseases highlighted comorbid conditions with shared genetic signals, illustrating the usefulness of mGWAS-Explorer.

2.5. Comparison with Other Tools

Table 2 provides detailed comparisons of mGWAS-Explorer with several bioinformatics resources that can be used for mGWAS, including Metabolomics GWAS Server [33,34], PheWeb [35], NETMAGE [37], and GePhEx [72]. The metabolomics GWAS server supports searching the results of two genome-wide association studies on the blood and urine metabolome in 7824 and 3861 individuals with European ancestry [33,34]. PheWeb is an excellent tool for developers to build a website to explore and visualize large-scale genetic associations [35]. NETMAGE focuses on visualizing disease–disease networks from sum-

mary statistics [37], and GePhEx allows visualization and interpretation of relationships across multiple traits with genetic associations evidence [72].

Table 2. Comparison of the main features of mGWAS-Explorer with other web-based tools. Symbols used for feature evaluations with '√' for present, '−' for absent, and '+' for a more quantitative assessment (more '+' symbols indicate better support).

Tool Name	mGWAS-Explorer	Metabolomics GWAS Server	PheWeb	NETMAGE	GePhEx
Data input and processing					
SNP	√	√	√	√	√
LD proxy search	√	√	−	−	√
Gene	√	√	√	−	√
Metabolite	√	√	√	−	√
Enrichment analysis					
SNP-set	√	−	−	−	−
Gene-set	√	−	−	−	√
Metabolite-set	√	−	−	−	−
Cross-phenotype exploration	√	√	√	√	√
Visual analytics					
Network visualization	+++	−	−	+	−
Network customization	+++	−	−	+	−
Integration with PPI network	√	−	−	−	−
Subnetwork extraction	√	−	−	√	−
Topology-based filtering	√	−	−	−	−
3D Manhattan plot	√	−	−	−	−

URL links:

- Metabolomics GWAS Server: http://metabolomics.helmholtz-muenchen.de/gwas/ (accessed on 1 May 2022).
- PheWeb: https://github.com/statgen/pheweb (accessed on 1 May 2022)
- NETMAGE: https://hdpm.biomedinfolab.com/netmage/ (accessed on 1 May 2022) (accept PheWAS summary statistics)
- GePhEx: https://gephex.ega-archive.org/ (accessed on 1 May 2022).

3. Discussion

Establishing meaningful connections between diseases and deciphering molecular mechanisms that underpin shared genetic architectures are among the key objectives of GWAS. Our work shows that integrating mGWAS summary statistics, LD proxy search, and visual analytics can rapidly reveal multiple associations across metabolites and diseases, which can be utilized to better understand the ongoing global health crisis, such as the COVID-19 pandemic and type 2 diabetes.

When looking at the cross-phenotype associations between metabolites and diseases, it is important to investigate the shared SNPs identified in the mGWAS-Explorer output to examine where the SNPs are located on the genome and the extent of overlapping of the SNPs. In fact, we consider mGWAS-Explorer as the initial stage in a pipeline for an in-depth mechanistic understanding of mGWAS before moving on to similarity analysis [26], colocalization analysis [14] or Mendelian randomization studies [73] to further investigate shared genetic signals in the same locus and to identify causal links. Ultimately, experimental studies in model organisms and human clinical studies are required to test the generated hypothesis to fully understand the mechanisms.

While our first case study highlighted shared genetic variants regulating metabolite abundance (e.g., citric acid and fibrinogen A-α peptides) and COVID-19 at *ABO*, much work needs to be done to fully understand the underlying mechanisms. Citric acid acts

as a bridge between carbohydrate and fatty acid metabolism, promoting the growth and development of immune cells [74]. Additionally, citric acid is an important component in the TCA cycle. TCA cycle metabolites play key roles in signaling regulations of the innate and adaptive immune systems [75], which may be involved in COVID-19 pathogenesis. mGWAS-Explorer was also able to identify *ENPEP* and *FUT2* as potential candidate genes for COVID-19, although the association signal of these two genes were below the genome-wide significance threshold in the original study [47]. Many follow-up studies have reported *ABO* and *ENPEP* as COVID-19 risk genes; however, the evidence for the *FUT2* gene is conflicting [60,76,77]. Indeed, a recent whole-genome sequencing study identified variants in *FUT2* associated with critical COVID-19 diseases [45]. *FUT2* is responsible for the expression of histo-blood group antigens on the mucosal surface of gastrointestinal, genitourinary, and respiratory tracts. Inactivating *FUT2* mutations lead to a non-secretor status, which confer resistance to norovirus and rotavirus gastrointestinal infections [78,79]. Furthermore, the minor allele frequency of the stop–gain variant rs601338 in the *FUT2* gene is drastically different between the European population (0.441) and the East Asian population (0.004) [80,81], which might explain the differences in host response to SARS-CoV-2 among different populations. However, the mechanism by which secretor status influencing COVID-19 pathogenesis is not fully understood. Therefore, it may be valuable to perform colocalization analysis and Mendelian randomization studies to identify the causal link between *FUT2* variants, fibrinogen A-α peptides, and COVID-19.

Increases in throughput and decreases in cost will enable a growing number of mGWAS to be conducted in the near future. The web server will be regularly updated to incorporate the most up-to-date mGWAS datasets, disease associations, and additional SNP annotation data (e.g., eQTLs or chromatin interactions) to serve as a valuable bioinformatics platform for mGWAS researchers. With this context, we also intend to add support for peak annotations of untargeted metabolomics data obtained from high-resolution mass spectrometry.

4. Materials and Methods

4.1. Knowledgebase Curation

(a) mGWAS papers were searched from PubMed, Web of Science, bioRxiv, and medRxiv, resulting in 65 publications as of December 2021. The summary statistics were either downloaded from public databases or supplementary data of the original publications. Statistical associations between metabolites and SNPs were summarized and pre-filtered using study-specific significance thresholds. In addition to *p*-values and effect sizes of SNP–metabolite associations, we have included metadata from each publication, such as the type of biofluid, sample size, population type, genotyping platform, metabolomics platform, etc. (b) For SNP annotation, three options are provided, including HaploReg [82], PhenoScanner [28], and VEP [20] by using the Application Programming Interface (API) service of each database. For the first two options, users can also perform an LD proxy search based on different populations and r^2 values. With VEP, users can select either a specific distance or the nearest number of genes for SNP annotation. (c) SNP–disease and gene–disease associations were downloaded from the DisGeNET database [32]. HMDB database was used to obtain metabolite–disease associations [83]. (d) KEGG, Recon3D, and Transporter Classification Database (TCDB) were used to curate knowledge-based gene–metabolite association information [84–86]. (e) The protein–protein interaction information is based on several well-established PPI databases [87–89]. (f) The libraries for enrichment analysis were curated from seven well-known databases, including GO, Reactome, KEGG, Orphanet, DrugMatrix, DisGeNET and DSigDB. The detailed description of these databases and their links can be found in the Supplementary Material.

4.2. Input Processing and Connection Identification

SNPs are identified by rsIDs, genes are identified by Entrez IDs, and metabolites are identified by HMDB IDs, platform-specific IDs, or feature tags (m/z_retention_time).

Additional identifiers have also been included, such as genomic coordinates for SNPs (after lifting all SNPs to GRCh37 assembly using LiftOver [90]), Ensemble ID, and gene symbol for genes, as well as KEGG ID and common name for metabolites.

There are two general types of relationships, including inter-omics and phenotype-specific links. Inter-omics connections are based on statistical associations (based on mGWAS), or knowledge-based associations (based on positional mapping for SNP–gene annotation or encoding enzymes/transporters for metabolite–gene connections). Phenotype-specific links allow users to identify variants that are associated with disease phenotypes. This information is obtained from DisGeNET [32] based on case-control genome-wide association studies or via text mining from the literature.

4.3. Implementation

The backend analysis was implemented using the R programming language (version 4.1.3). The whole framework was built based on the PrimeFaces component library (version 11.0.0). The integrated data are stored in a relational database using SQLite. The interactive visualization was developed based on the sigma.js and echarts.js JavaScript libraries for network view and 3D Manhattan plot, respectively.

4.4. Data Collection for Case Studies

The datasets of COVID-19 and type 2 diabetes case studies were downloaded from their respective original publications [47,63]. LD clumping were performed to identify the independent signals by using the *ieugwasr* package with default parameters [29] prior to the analysis. Specifically, the SNPs with the lowest *p*-value are retained, where SNPs in LD within a certain window are removed in LD clumping [91]. In both case studies, European population and $r^2 = 0.8$ were set as input parameters for LD proxy search.

4.5. Network Visual Analytics

4.5.1. Network Creation and Customization

The default networks are built by querying for the direct mapping from the knowledgebase. Optionally, users can choose to expand the network by including PPIs in the SNP module. However, the result may suffer from the 'hairball' effect, which severely limits the usefulness and interpretability. Therefore, mGWAS-Explorer offers support to refine large networks based on node degree or betweenness values, batch filtering, or the shortest paths, as well as by computing minimum subnetworks based on the prize-collecting Steiner Forest (PCSF) algorithm [92]. The detailed instructions on how to navigate the network visualization system can be found in our Supplementary Material.

4.5.2. Functional Enrichment Analysis

The combination of network visualization and functional enrichment analysis is a valuable tool for gaining key biological insights. For SNP input, two types of enrichment approaches have been implemented—(1) directly testing in SNP-set library, or (2) testing on mapped genes for enrichment using hypergeometric tests. When the input is a gene or metabolite, the associated gene-set or metabolite-set enrichment analysis can be performed. The result tables will be displayed under the Function Explorer panel. Notably, clicking a row of the table will highlight the nodes contained in the corresponding function/pathway within the network. In addition, mGWAS-Explorer also permits enrichment analysis on the selected nodes of interests, for instance, from the batch selection panel.

4.5.3. Other Advanced Features

The Network Viewer page contains multiple advanced features for network visual exploration, including Network Layout, Global Node Styles, Module Explorer, Batch Selection, and Path Finder. Ten different network layout algorithms are available, including Force-Atlas, Fruchterman–Reingold, Circular, Graphopt, Large Graph, Random, Circular Bipartite/Tripartite, Linear Bipartite/Tripartite, Concentric, and Backbone layout. Three

module detection algorithms are offered in the Module Explorer, including the WalkTrap, InfoMap, and the Label Propagation algorithms based on the igraph R package [93]. These options can be combined to obtain a better visualization experience. Users can find details of these algorithms on our FAQs page.

5. Conclusions

We have developed mGWAS-Explorer to allow users to easily explore the published mGWAS datasets, and to provide contextualized analysis for a given list of SNPs, genes, or metabolites. As demonstrated by our case studies of COVID-19 and type 2 diabetes, mGWAS-Explorer can facilitate hypothesis generation and reveal functional insights into the genetic basis of human metabolism to permit translational discoveries.

References

1. Visscher, P.M.; Wray, N.R.; Zhang, Q.; Sklar, P.; McCarthy, M.I.; Brown, M.A.; Yang, J. 10 years of GWAS discovery: Biology, function, and translation. *Am. J. Hum. Genet.* **2017**, *101*, 5–22. [CrossRef] [PubMed]
2. Cano-Gamez, E.; Trynka, G. From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Front. Genet.* **2020**, *11*, 424. [CrossRef] [PubMed]
3. Kastenmüller, G.; Raffler, J.; Gieger, C.; Suhre, K. Genetics of human metabolism: An update. *Hum. Mol. Genet.* **2015**, *24*, R93–R101. [CrossRef] [PubMed]
4. Hagenbeek, F.A.; Pool, R.; van Dongen, J.; Draisma, H.H.M.; Jan Hottenga, J.; Willemsen, G.; Abdellaoui, A.; Fedko, I.O.; den Braber, A.; Visser, P.J.; et al. Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. *Nat. Commun.* **2020**, *11*, 39. [CrossRef] [PubMed]
5. Gieger, C.; Geistlinger, L.; Altmaier, E.; De Angelis, M.H.; Kronenberg, F.; Meitinger, T.; Mewes, H.-W.; Wichmann, H.-E.; Weinberger, K.M.; Adamski, J. Genetics meets metabolomics: A genome-wide association study of metabolite profiles in human serum. *PLoS Genet.* **2008**, *4*, e1000282. [CrossRef]
6. Gallois, A.; Mefford, J.; Ko, A.; Vaysse, A.; Julienne, H.; Ala-Korpela, M.; Laakso, M.; Zaitlen, N.; Pajukanta, P.; Aschard, H. A comprehensive study of metabolite genetics reveals strong pleiotropy and heterogeneity across time and context. *Nat. Commun.* **2019**, *10*, 4788. [CrossRef]
7. Lotta, L.A.; Pietzner, M.; Stewart, I.D.; Wittemans, L.B.L.; Li, C.; Bonelli, R.; Raffler, J.; Biggs, E.K.; Oliver-Williams, C.; Auyeung, V.P.W.; et al. A cross-platform approach identifies genetic regulators of human metabolism and health. *Nat. Genet.* **2021**, *53*, 54–64. [CrossRef]
8. Solovieff, N.; Cotsapas, C.; Lee, P.H.; Purcell, S.M.; Smoller, J.W. Pleiotropy in complex traits: Challenges and strategies. *Nat. Rev. Genet.* **2013**, *14*, 483. [CrossRef]
9. Visscher, P.M.; Yang, J. A plethora of pleiotropy across complex traits. *Nat. Genet.* **2016**, *48*, 707. [CrossRef]
10. Watanabe, K.; Stringer, S.; Frei, O.; Umićević Mirkov, M.; de Leeuw, C.; Polderman, T.J.C.; van der Sluis, S.; Andreassen, O.A.; Neale, B.M.; Posthuma, D. A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **2019**, *51*, 1339–1348. [CrossRef]

11. Weighill, D.; Jones, P.; Bleker, C.; Ranjan, P.; Shah, M.; Zhao, N.; Martin, M.; DiFazio, S.; Macaya-Sanz, D.; Schmutz, J.; et al. Multi-Phenotype Association Decomposition: Unraveling Complex Gene-Phenotype Relationships. *Front. Genet.* **2019**, *10*, 417. [CrossRef] [PubMed]

12. Julienne, H.; Laville, V.; McCaw, Z.R.; He, Z.; Guillemot, V.; Lasry, C.; Ziyatdinov, A.; Nerin, C.; Vaysse, A.; Lechat, P.; et al. Multitrait GWAS to connect disease variants and biological mechanisms. *PLoS Genet.* **2021**, *17*, e1009713. [CrossRef] [PubMed]

13. Bulik-Sullivan, B.K.; Loh, P.R.; Finucane, H.K.; Ripke, S.; Yang, J.; Patterson, N.; Daly, M.J.; Price, A.L.; Neale, B.M. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **2015**, *47*, 291–295. [CrossRef] [PubMed]

14. Giambartolomei, C.; Vukcevic, D.; Schadt, E.E.; Franke, L.; Hingorani, A.D.; Wallace, C.; Plagnol, V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **2014**, *10*, e1004383. [CrossRef] [PubMed]

15. Majumdar, A.; Haldar, T.; Bhattacharya, S.; Witte, J.S. An efficient Bayesian meta-analysis approach for studying cross-phenotype genetic associations. *PLoS Genet.* **2018**, *14*, e1007139. [CrossRef] [PubMed]

16. Han, B.; Pouget, J.G.; Slowikowski, K.; Stahl, E.; Lee, C.H.; Diogo, D.; Hu, X.; Park, Y.R.; Kim, E.; Gregersen, P.K.; et al. A method to decipher pleiotropy by detecting underlying heterogeneity driven by hidden subgroups applied to autoimmune and neuropsychiatric diseases. *Nat. Genet.* **2016**, *48*, 803–810. [CrossRef] [PubMed]

17. Trochet, H.; Pirinen, M.; Band, G.; Jostins, L.; McVean, G.; Spencer, C.C.A. Bayesian meta-analysis across genome-wide association studies of diverse phenotypes. *Genet. Epidemiol.* **2019**, *43*, 532–547. [CrossRef]

18. Li, X.; Zhu, X. Cross-Phenotype Association Analysis Using Summary Statistics from GWAS. *Methods Mol. Biol.* **2017**, *1666*, 455–467. [CrossRef]

19. Arnold, M.; Raffler, J.; Pfeufer, A.; Suhre, K.; Kastenmüller, G. SNiPA: An interactive, genetic variant-centered annotation browser. *Bioinformatics* **2014**, *31*, 1334–1336. [CrossRef]

20. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.; Thormann, A.; Flicek, P.; Cunningham, F. The ensembl variant effect predictor. *Genome Biol.* **2016**, *17*, 122. [CrossRef]

21. Carlin, D.E.; Fong, S.H.; Qin, Y.; Jia, T.; Huang, J.K.; Bao, B.; Zhang, C.; Ideker, T. A Fast and Flexible Framework for Network-Assisted Genomic Association. *iScience* **2019**, *16*, 155–161. [CrossRef] [PubMed]

22. Bastarache, L.; Denny, J.C.; Roden, D.M. Phenome-Wide Association Studies. *JAMA* **2022**, *327*, 75–76. [CrossRef]

23. Tanha, H.M.; Sathyanarayanan, A.; Nyholt, D.R. Genetic overlap and causality between blood metabolites and migraine. *Am. J. Hum. Genet.* **2021**, *108*, 2086–2098. [CrossRef] [PubMed]

24. Kaur, Y.; Wang, D.X.; Liu, H.Y.; Meyre, D. Comprehensive identification of pleiotropic loci for body fat distribution using the NHGRI-EBI Catalog of published genome-wide association studies. *Obes. Rev.* **2019**, *20*, 385–406. [CrossRef] [PubMed]

25. Guo, Y.; Rist, P.M.; Daghlas, I.; Giulianini, F.; Kurth, T.; Chasman, D.I. A genome-wide cross-phenotype meta-analysis of the association of blood pressure with migraine. *Nat. Commun.* **2020**, *11*, 3368. [CrossRef] [PubMed]

26. George, G.; Huang, Y.; Gan, S.; Nar, A.S.; Ha, J.; Venkatesan, R.; Mohan, V.; Wang, H.; Brown, A.; Palmer, C.N.A.; et al. iPheGWAS: An intelligent computational framework to integrate and visualise genome-phenome wide association results. *bioRxiv* **2022**. [CrossRef]

27. Buniello, A.; MacArthur, J.A.L.; Cerezo, M.; Harris, L.W.; Hayhurst, J.; Malangone, C.; McMahon, A.; Morales, J.; Mountjoy, E.; Sollis, E.; et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **2019**, *47*, D1005–D1012. [CrossRef]

28. Kamat, M.A.; Blackshaw, J.A.; Young, R.; Surendran, P.; Burgess, S.; Danesh, J.; Butterworth, A.S.; Staley, J.R. PhenoScanner V2: An expanded tool for searching human genotype-phenotype associations. *Bioinformatics* **2019**, *35*, 4851–4853. [CrossRef]

29. Elsworth, B.; Lyon, M.; Alexander, T.; Liu, Y.; Matthews, P.; Hallett, J.; Bates, P.; Palmer, T.; Haberland, V.; Smith, G.D.; et al. The MRC IEU OpenGWAS data infrastructure. *bioRxiv* **2020**. [CrossRef]

30. Ghoussaini, M.; Mountjoy, E.; Carmona, M.; Peat, G.; Schmidt, E.M.; Hercules, A.; Fumis, L.; Miranda, A.; Carvalho-Silva, D.; Buniello, A.; et al. Open Targets Genetics: Systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res.* **2021**, *49*, D1311–D1320. [CrossRef]

31. Shashkova, T.I.; Pakhomov, E.D.; Gorev, D.D.; Karssen, L.C.; Joshi, P.K.; Aulchenko, Y.S. PheLiGe: An interactive database of billions of human genotype-phenotype associations. *Nucleic Acids Res.* **2021**, *49*, D1347–D1350. [CrossRef] [PubMed]

32. Piñero, J.; Ramírez-Anguita, J.M.; Saüch-Pitarch, J.; Ronzano, F.; Centeno, E.; Sanz, F.; Furlong, L.I. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res* **2020**, *48*, D845–D855. [CrossRef] [PubMed]

33. Shin, S.-Y.; Fauman, E.B.; Petersen, A.-K.; Krumsiek, J.; Santos, R.; Huang, J.; Arnold, M.; Erte, I.; Forgetta, V.; Yang, T.-P. An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **2014**, *46*, 543. [CrossRef] [PubMed]

34. Raffler, J.; Friedrich, N.; Arnold, M.; Kacprowski, T.; Rueedi, R.; Altmaier, E.; Bergmann, S.; Budde, K.; Gieger, C.; Homuth, G.; et al. Genome-Wide Association Study with Targeted and Non-targeted NMR Metabolomics Identifies 15 Novel Loci of Urinary Human Metabolic Individuality. *PLoS Genet.* **2015**, *11*, e1005487. [CrossRef]

35. Gagliano Taliun, S.A.; VandeHaar, P.; Boughton, A.P.; Welch, R.P.; Taliun, D.; Schmidt, E.M.; Zhou, W.; Nielsen, J.B.; Willer, C.J.; Lee, S.; et al. Exploring and visualizing large-scale genetic associations by using PheWeb. *Nat. Genet.* **2020**, *52*, 550–552. [CrossRef]

36. Wang, L.; Balmat, T.J.; Antonia, A.L.; Constantine, F.J.; Henao, R.; Burke, T.W.; Ingham, A.; McClain, M.T.; Tsalik, E.L.; Ko, E.R.; et al. An atlas connecting shared genetic architecture of human diseases and molecular phenotypes provides insight into COVID-19 susceptibility. *Genome Med.* **2021**, *13*, 83. [CrossRef]

37. Sriram, V.; Shivakumar, M.; Jung, S.H.; Nam, Y.; Bang, L.; Verma, A.; Lee, S.; Choe, E.K.; Kim, D. NETMAGE: A human disease phenotype map generator for the network-based visualization of phenome-wide association study results. *Gigascience* **2022**, *11*, giac002. [CrossRef]

38. Strayer, N.; Shirey-Rice, J.K.; Shyr, Y.; Denny, J.C.; Pulley, J.M.; Xu, Y. PheWAS-ME: A web-app for interactive exploration of multimorbidity patterns in PheWAS. *Bioinformatics* **2021**, *37*, 1778–1780. [CrossRef]

39. George, G.; Gan, S.; Huang, Y.; Appleby, P.; Nar, A.S.; Venkatesan, R.; Mohan, V.; Palmer, C.N.A.; Doney, A.S.F. PheGWAS: A new dimension to visualize GWAS across multiple phenotypes. *Bioinformatics* **2020**, *36*, 2500–2505. [CrossRef]

40. Zhu, Z.; Anttila, V.; Smoller, J.W.; Lee, P.H. Statistical power and utility of meta-analysis methods for cross-phenotype genome-wide association studies. *PLoS ONE* **2018**, *13*, e0193256. [CrossRef]

41. Lee, B.; Zhang, S.; Poleksic, A.; Xie, L. Heterogeneous Multi-Layered Network Model for Omics Data Integration and Analysis. *Front. Genet.* **2019**, *10*, 1381. [CrossRef] [PubMed]

42. Sadegh, S.; Skelton, J.; Anastasi, E.; Bernett, J.; Blumenthal, D.B.; Galindez, G.; Salgado-Albarrán, M.; Lazareva, O.; Flanagan, K.; Cockell, S.; et al. Network medicine for disease module identification and drug repurposing with the NeDRex platform. *Nat. Commun.* **2021**, *12*, 6848. [CrossRef] [PubMed]

43. Petersen, A.-K.; Krumsiek, J.; Wägele, B.; Theis, F.J.; Wichmann, H.-E.; Gieger, C.; Suhre, K. On the hypothesis-free testing of metabolite ratios in genome-wide and metabolome-wide association studies. *BMC Bioinform.* **2012**, *13*, 120. [CrossRef]

44. Stacey, D.; Fauman, E.B.; Ziemek, D.; Sun, B.B.; Harshfield, E.L.; Wood, A.M.; Butterworth, A.S.; Suhre, K.; Paul, D.S. ProGeM: A framework for the prioritization of candidate causal genes at molecular quantitative trait loci. *Nucleic Acids Res.* **2018**, *47*, e3. [CrossRef] [PubMed]

45. Kousathanas, A.; Pairo-Castineira, E.; Rawlik, K.; Stuckey, A.; Odhams, C.A.; Walker, S.; Russell, C.D.; Malinauskas, T.; Wu, Y.; Millar, J.; et al. Whole genome sequencing reveals host factors underlying critical COVID-19. *Nature* **2022**. [CrossRef]

46. Pairo-Castineira, E.; Clohisey, S.; Klaric, L.; Bretherick, A.D.; Rawlik, K.; Pasko, D.; Walker, S.; Parkinson, N.; Fourman, M.H.; Russell, C.D.; et al. Genetic mechanisms of critical illness in COVID-19. *Nature* **2021**, *591*, 92–98. [CrossRef]

47. Ellinghaus, D.; Degenhardt, F.; Bujanda, L.; Buti, M.; Albillos, A.; Invernizzi, P.; Fernández, J.; Prati, D.; Baselli, G.; Asselta, R.; et al. Genomewide Association Study of Severe COVID-19 with Respiratory Failure. *N. Engl. J. Med.* **2020**, *383*, 1522–1534. [CrossRef]

48. Mapping the human genetic architecture of COVID-19. *Nature* **2021**, *600*, 472–477. [CrossRef]

49. Zhang, Q.; Bastard, P.; Liu, Z.; Le Pen, J.; Moncada-Velez, M.; Chen, J.; Ogishi, M.; Sabli, I.K.D.; Hodeib, S.; Korol, C.; et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **2020**, *370*, eabd4570. [CrossRef]

50. Sindelar, M.; Stancliffe, E.; Schwaiger-Haber, M.; Anbukumar, D.S.; Adkins-Travis, K.; Goss, C.W.; O'Halloran, J.A.; Mudd, P.A.; Liu, W.C.; Albrecht, R.A.; et al. Longitudinal metabolomics of human plasma reveals prognostic markers of COVID-19 disease severity. *Cell Rep. Med.* **2021**, *2*, 100369. [CrossRef]

51. Shi, D.; Yan, R.; Lv, L.; Jiang, H.; Lu, Y.; Sheng, J.; Xie, J.; Wu, W.; Xia, J.; Xu, K.; et al. The serum metabolome of COVID-19 patients is distinctive and predictive. *Metab. Clin. Exp.* **2021**, *118*, 154739. [CrossRef] [PubMed]

52. Timmann, C.; Thye, T.; Vens, M.; Evans, J.; May, J.; Ehmen, C.; Sievertsen, J.; Muntau, B.; Ruge, G.; Loag, W.; et al. Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature* **2012**, *489*, 443–446. [CrossRef] [PubMed]

53. Li, H.; Cai, Y.; Xu, A.D. Association study of polymorphisms in the ABO gene and their gene-gene interactions with ischemic stroke in Chinese population. *J. Clin. Lab. Anal.* **2018**, *32*, e22329. [CrossRef] [PubMed]

54. Germain, M.; Saut, N.; Oudot-Mellakh, T.; Letenneur, L.; Dupuy, A.M.; Bertrand, M.; Alessi, M.C.; Lambert, J.C.; Zelenika, D.; Emmerich, J.; et al. Caution in interpreting results from imputation analysis when linkage disequilibrium extends over a large distance: A case study on venous thrombosis. *PLoS ONE* **2012**, *7*, e38538. [CrossRef] [PubMed]

55. Zhao, J.; Yang, Y.; Huang, H.; Li, D.; Gu, D.; Lu, X.; Zhang, Z.; Liu, L.; Liu, T.; Liu, Y.; et al. Relationship between the ABO Blood Group and the Coronavirus Disease 2019 (COVID-19) Susceptibility. *Clin. Infect. Dis.* **2021**, *73*, 328–331. [CrossRef]

56. Goel, R.; Bloch, E.M.; Pirenne, F.; Al-Riyami, A.Z.; Crowe, E.; Dau, L.; Land, K.; Townsend, M.; Jecko, T.; Rahimi-Levene, N.; et al. ABO blood group and COVID-19: A review on behalf of the ISBT COVID-19 Working Group. *Vox Sang.* **2021**, *116*, 849–861. [CrossRef]

57. Kattula, S.; Byrnes, J.R.; Wolberg, A.S. Fibrinogen and Fibrin in Hemostasis and Thrombosis. *Arterioscler. Thromb. Vasc. Biol.* **2017**, *37*, e13–e21. [CrossRef]

58. Malas, M.B.; Naazie, I.N.; Elsayed, N.; Mathlouthi, A.; Marmor, R.; Clary, B. Thromboembolism risk of COVID-19 is high and associated with a higher risk of mortality: A systematic review and meta-analysis. *EClinicalMedicine* **2020**, *29*, 100639. [CrossRef]

59. Lange, C.; Wolf, J.; Auw-Haedrich, C.; Schlecht, A.; Boneva, S.; Lapp, T.; Horres, R.; Agostini, H.; Martin, G.; Reinhard, T.; et al. Expression of the COVID-19 receptor ACE2 in the human conjunctiva. *J. Med. Virol.* **2020**, *92*, 2081–2086. [CrossRef]

60. Mankelow, T.J.; Singleton, B.K.; Moura, P.L.; Stevens-Hernandez, C.J.; Cogan, N.M.; Gyorffy, G.; Kupzig, S.; Nichols, L.; Asby, C.; Pooley, J.; et al. Blood group type A secretors are associated with a higher risk of COVID-19 cardiovascular disease complications. *EJHaem* **2021**, *2*, 175–187. [CrossRef]

61. Langenberg, C.; Lotta, L.A. Genomic insights into the causes of type 2 diabetes. *Lancet* **2018**, *391*, 2463–2474. [CrossRef]

62. Fuchsberger, C.; Flannick, J.; Teslovich, T.M.; Mahajan, A.; Agarwala, V.; Gaulton, K.J.; Ma, C.; Fontanillas, P.; Moutsianas, L.; McCarthy, D.J.; et al. The genetic architecture of type 2 diabetes. *Nature* **2016**, *536*, 41–47. [CrossRef] [PubMed]

63. Scott, R.A.; Scott, L.J.; Mägi, R.; Marullo, L.; Gaulton, K.J.; Kaakinen, M.; Pervjakova, N.; Pers, T.H.; Johnson, A.D.; Eicher, J.D.; et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* **2017**, *66*, 2888–2902. [CrossRef] [PubMed]

64. Chen, Y.; Chen, C.; Ke, X.; Xiong, L.; Shi, Y.; Li, J.; Tan, X.; Ye, S. Analysis of circulating cholesterol levels as a mediator of an association between ABO blood group and coronary heart disease. *Circulation. Cardiovasc. Genet.* **2014**, *7*, 43–48. [CrossRef]

65. Li, S.; Schooling, C.M. A phenome-wide association study of ABO blood groups. *BMC Med.* **2020**, *18*, 334. [CrossRef]

66. Meo, S.A.; Rouq, F.A.; Suraya, F.; Zaidi, S.Z. Association of ABO and Rh blood groups with type 2 diabetes mellitus. *Eur. Rev. Med. Pharmacol. Sci.* **2016**, *20*, 237–242.

67. Fagherazzi, G.; Gusto, G.; Clavel-Chapelon, F.; Balkau, B.; Bonnet, F. ABO and Rhesus blood groups and risk of type 2 diabetes: Evidence from the large E3N cohort study. *Diabetologia* **2015**, *58*, 519–522. [CrossRef]

68. Yahaya, T.O.; Oladele, E.O.; Mshelia, M.B.; Sifau, M.O.; Fashola, O.D.; Bunza, M.; Nathaniel, J. Influence of ABO blood groups and demographic characteristics on the prevalence of type 2 diabetes in Lagos, southwest Nigeria. *Bull. Natl. Res. Cent.* **2021**, *45*, 144. [CrossRef]

69. Zhu, L.; She, Z.G.; Cheng, X.; Qin, J.J.; Zhang, X.J.; Cai, J.; Lei, F.; Wang, H.; Xie, J.; Wang, W.; et al. Association of Blood Glucose Control and Outcomes in Patients with COVID-19 and Pre-existing Type 2 Diabetes. *Cell Metab.* **2020**, *31*, 1068–1077.e1063. [CrossRef]

70. Metwally, A.A.; Mehta, P.; Johnson, B.S.; Nagarjuna, A.; Snyder, M.P. COVID-19-Induced New-Onset Diabetes: Trends and Technologies. *Diabetes* **2021**, *70*, 2733–2744. [CrossRef]

71. Rajpal, A.; Rahimi, L.; Ismail-Beigi, F. Factors leading to high morbidity and mortality of COVID-19 in patients with type 2 diabetes. *J. Diabetes* **2020**, *12*, 895–908. [CrossRef] [PubMed]

72. Farré, X.; Spataro, N.; Haziza, F.; Rambla, J.; Navarro, A. Genome-phenome explorer (GePhEx): A tool for the visualization and interpretation of phenotypic relationships supported by genetic evidence. *Bioinformatics* **2019**, *36*, 890–896. [CrossRef] [PubMed]

73. Hemani, G.; Zheng, J.; Elsworth, B.; Wade, K.H.; Haberland, V.; Baird, D.; Laurin, C.; Burgess, S.; Bowden, J.; Langdon, R.; et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **2018**, *7*, e34408. [CrossRef] [PubMed]

74. Williams, N.C.; O'Neill, L.A.J. A Role for the Krebs Cycle Intermediate Citrate in Metabolic Reprogramming in Innate Immunity and Inflammation. *Front. Immunol.* **2018**, *9*, 141. [CrossRef]

75. Martínez-Reyes, I.; Chandel, N.S. Mitochondrial TCA cycle metabolites control physiology and disease. *Nat. Commun.* **2020**, *11*, 102. [CrossRef]

76. Delanghe, J.R.; De Buyzere, M.L.; Speeckaert, M.M. Genetic Polymorphisms in the Host and COVID-19 Infection. *Adv. Exp. Med. Biol.* **2021**, *1318*, 109–118. [CrossRef]

77. Matzhold, E.M.; Berghold, A.; Bemelmans, M.K.B.; Banfi, C.; Stelzl, E.; Kessler, H.H.; Steinmetz, I.; Krause, R.; Wurzer, H.; Schlenke, P.; et al. Lewis and ABO histo-blood types and the secretor status of patients hospitalized with COVID-19 implicate a role for ABO antibodies in susceptibility to infection with SARS-CoV-2. *Transfusion* **2021**, *61*, 2736–2745. [CrossRef]

78. Lindesmith, L.; Moe, C.; Marionneau, S.; Ruvoen, N.; Jiang, X.; Lindblad, L.; Stewart, P.; LePendu, J.; Baric, R. Human susceptibility and resistance to Norwalk virus infection. *Nat. Med.* **2003**, *9*, 548–553. [CrossRef]

79. Payne, D.C.; Currier, R.L.; Staat, M.A.; Sahni, L.C.; Selvarangan, R.; Halasa, N.B.; Englund, J.A.; Weinberg, G.A.; Boom, J.A.; Szilagyi, P.G.; et al. Epidemiologic Association between FUT2 Secretor Status and Severe Rotavirus Gastroenteritis in Children in the United States. *JAMA Pediatrics* **2015**, *169*, 1040–1045. [CrossRef]

80. Cunningham, F.; Allen, J.E.; Allen, J.; Alvarez-Jarreta, J.; Amode, M.R.; Armean, I.M.; Austine-Orimoloye, O.; Azov, A.G.; Barnes, I.; Bennett, R.; et al. Ensembl 2022. *Nucleic Acids Res.* **2021**, *50*, D988–D995. [CrossRef]

81. Ferrer-Admetlla, A.; Sikora, M.; Laayouni, H.; Esteve, A.; Roubinet, F.; Blancher, A.; Calafell, F.; Bertranpetit, J.; Casals, F. A natural history of FUT2 polymorphism in humans. *Mol. Biol. Evol.* **2009**, *26*, 1993–2003. [CrossRef] [PubMed]

82. Ward, L.D.; Kellis, M. HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* **2015**, *44*, D877–D881. [CrossRef]

83. Wishart, D.S.; Guo, A.; Oler, E.; Wang, F.; Anjum, A.; Peters, H.; Dizon, R.; Sayeeda, Z.; Tian, S.; Lee, B.L.; et al. HMDB 5.0: The Human Metabolome Database for 2022. *Nucleic Acids Res.* **2022**, *50*, D622–D631. [CrossRef] [PubMed]

84. Kanehisa, M. Enzyme annotation and metabolic reconstruction using KEGG. *Protein Funct. Predict. Methods Protoc.* **2017**, *1611*, 135–145.

85. Saier, M.H.; Reddy, V.S.; Moreno-Hagelsieb, G.; Hendargo, K.J.; Zhang, Y.; Iddamsetty, V.; Lam, K.J.K.; Tian, N.; Russum, S.; Wang, J.; et al. The Transporter Classification Database (TCDB): 2021 update. *Nucleic Acids Res.* **2021**, *49*, D461–D467. [CrossRef]

86. Brunk, E.; Sahoo, S.; Zielinski, D.C.; Altunkaya, A.; Dräger, A.; Mih, N.; Gatto, F.; Nilsson, A.; Preciat Gonzalez, G.A.; Aurich, M.K.; et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.* **2018**, *36*, 272–281. [CrossRef]

87. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [CrossRef]

88. Breuer, K.; Foroushani, A.K.; Laird, M.R.; Chen, C.; Sribnaia, A.; Lo, R.; Winsor, G.L.; Hancock, R.E.; Brinkman, F.S.; Lynn, D.J. InnateDB: Systems biology of innate immunity and beyond—Recent updates and continuing curation. *Nucleic Acids Res.* **2013**, *41*, D1228–D1233. [CrossRef]

89. Rolland, T.; Taşan, M.; Charloteaux, B.; Pevzner, S.J.; Zhong, Q.; Sahni, N.; Yi, S.; Lemmens, I.; Fontanillo, C.; Mosca, R.J.C. A proteome-scale map of the human interactome network. *Cell* **2014**, *159*, 1212–1226. [CrossRef]

90. Lee, B.T.; Barber, G.P.; Benet-Pagès, A.; Casper, J.; Clawson, H.; Diekhans, M.; Fischer, C.; Gonzalez, J.N.; Hinrichs, A.S.; Lee, C.M.; et al. The UCSC Genome Browser database: 2022 update. *Nucleic Acids Res.* **2022**, *50*, D1115–D1122. [CrossRef]

91. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef] [PubMed]

92. Akhmedov, M.; Kedaigle, A.; Chong, R.E.; Montemanni, R.; Bertoni, F.; Fraenkel, E.; Kwee, I. PCSF: An R-package for network-based interpretation of high-throughput data. *PLoS Comput. Biol.* **2017**, *13*, e1005694. [CrossRef] [PubMed]

93. Csardi, G.; Nepusz, T.J.I. The igraph software package for complex network research. *InterJournal Complex Syst.* **1695**, *5*, 1–9.

94. UniProt Consortium, T. UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* **2018**, *46*, 2699. [CrossRef]

95. Landrum, M.J.; Kattman, B.L. ClinVar at five years: Delivering on the promise. *Hum. Mutat.* **2018**, *39*, 1623–1630. [CrossRef]

96. MacArthur, J.; Bowler, E.; Cerezo, M.; Gil, L.; Hall, P.; Hastings, E.; Junkins, H.; McMahon, A.; Milano, A.; Morales, J.; et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **2017**, *45*, D896–D901. [CrossRef]

97. Li, M.J.; Liu, Z.; Wang, P.; Wong, M.P.; Nelson, M.R.; Kocher, J.P.; Yeager, M.; Sham, P.C.; Chanock, S.J.; Xia, Z.; et al. GWASdb v2: An update database for human genetic variants identified by genome-wide association studies. *Nucleic Acids Res.* **2016**, *44*, D869–D876. [CrossRef]

98. Luck, K.; Kim, D.K.; Lambourne, L.; Spirohn, K.; Begg, B.E.; Bian, W.; Brignall, R.; Cafarelli, T.; Campos-Laborie, F.J.; Charloteaux, B.; et al. A reference map of the human binary protein interactome. *Nature* **2020**, *580*, 402–408. [CrossRef]

99. Boyle, E.I.; Weng, S.; Gollub, J.; Jin, H.; Botstein, D.; Cherry, J.M.; Sherlock, G. GO::TermFinder—Open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **2004**, *20*, 3710–3715. [CrossRef]