

mGWAS-Explorer Tutorial

-- COVID 19 & T2D Case Studies



Computer and Browser Requirements

- A personal computer with an Internet connection
- An up-to-date web browser that supports HTML5 with JavaScript enabled, such as Google Chrome (v50+), Firefox (3.0+), and Internet Explorer (9.0+)
- We recommend a ≥ 2 GHz CPU, 4 GB physical RAM with at least 2 GB free and a minimum of a 15-inch screen with a screen resolution of 1,280 × 800 or higher
- A mouse with scrolling support is required for network visualization

Goal for this tutorial

- Demonstrate the practical application of mGWAS-Explorer by using two case studies:
 - COVID-19
 - Type 2 diabetes

COVID-19 Case Study

Motivation

- The host genetic variation is known to influence the severity of SARS-CoV-2 infection [1].
- And the blood metabolome reveals biomarkers for disease prediction and classification [2].
- However, understanding mechanisms that link genetic variation to metabolism and clinical endpoint remains an important challenge.

1. 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, doi:10.1038/s41586-022-04576-6.
2. 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 reports. Medicine* 2021, 2, 100369, doi:10.1016/j.xcrm.2021.100369.

Objectives

Therefore, we applied mGWAS-Explorer to a list of SNPs identified from a GWAS of severe COVID-19 [1] to provide insights into shared genetic architecture of diseases and intermediate metabolic phenotypes.

1. 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. *The New England journal of medicine* 2020, 383, 1522-1534, doi:10.1056/NEJMoa2020283.

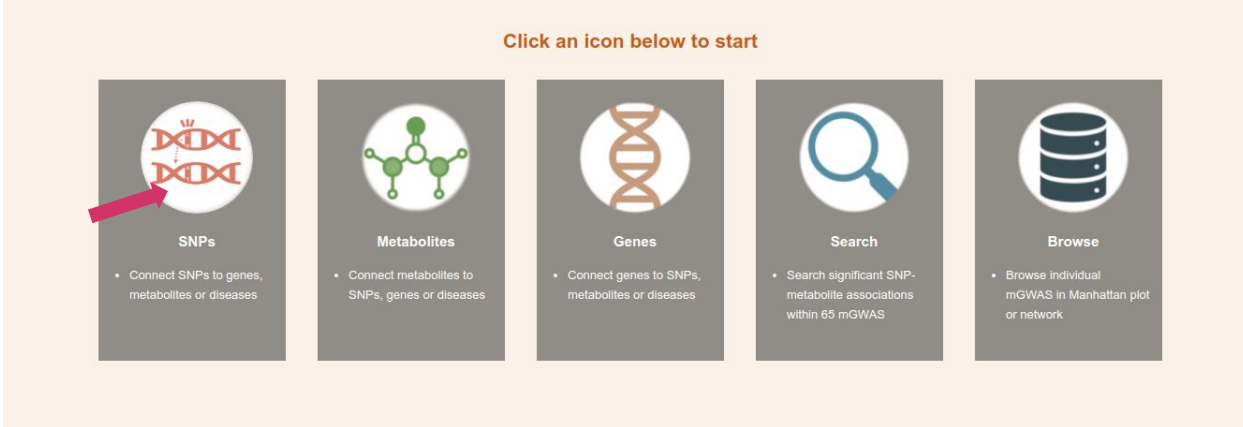
Methods

- Pre-processing:
 - Suggestive significant association p-value threshold (1×10^{-5})
 - LD clumping were performed to identify the independent signals by using the *ieugwasr* package
- Analysis overview:
 - Upload a list of SNPs
 - Network building
 - Network analytics
 - Cross-phenotype association analysis

Starting up

- Go to the mGWAS-Explorer homepage (www.mgwas.ca).
- For SNP list input, click the “SNPs” button to enter the data upload page.

Click an icon below to start



- SNPs**
 - Connect SNPs to genes, metabolites or diseases
- Metabolites**
 - Connect metabolites to SNPs, genes or diseases
- Genes**
 - Connect genes to SNPs, metabolites or diseases
- Search**
 - Search significant SNP-metabolite associations within 65 mGWAS
- Browse**
 - Browse Individual mGWAS in Manhattan plot or network

Data upload

1. Click the **“Try Examples”** button
2. Click **“Yes”** button in the pop-up window (default is SNP list 1)
3. Click **“Submit”** button to upload the data
4. Click **“Proceed”** button to the Network Builder page

Enter a list of SNPs below:

Biofluid: Unspecified
Population: Unspecified
ID type: rsID
Network type: SNP-Metabolite, SNP-Gene, SNP-Disease
SNP list (one entry per line)

1 Try Examples

2

Example SNP List

Data	Metadata	Description
<input checked="" type="radio"/> SNP list 1	ID Type: rsID	Example SNP list from COVID-19 GWAS after LD clumping (Ellinghaus et al. 2020)
<input type="radio"/> SNP list 2	ID Type: rsID	Example SNP list from type 2 diabetes GWAS after LD clumping (Scott et al. 2017)

Yes Cancel

Submit

Enter a list of SNPs below:

Biofluid: Unspecified
Population: Unspecified
ID type: rsID
Network type: SNP-Metabolite, SNP-Gene, SNP-Disease
SNP list (one entry per line)

3

4

Submit

Try Examples

Proceed

Did You Know?

You can also perform LD proxy search by specifying the population of interest and r2 value while setting the parameters for SNP to gene annotation.

- Note: here are the parameters automatically entered for the COVID-19 case study:

- Biofluid: “Unspecified”
- Population: “Unspecified”
- ID type: “rsID”
- SNP annotation:
 - “EUR” and “ $r^2=0.8$ ” are entered for LD proxy search
 - Include PPI
- Network type:
 - “SNP-Metabolite”
 - “SNP-Gene”
 - “SNP-Disease”

Biofluid	<input type="text" value="Unspecified"/>
Population	<input type="text" value="Unspecified"/>
ID type	<input type="text" value="rsID"/>
Network type	<input checked="" type="checkbox"/> SNP-Metabolite Statistical associations based on curated <u>mGWAS datasets</u>
	<input checked="" type="checkbox"/> SNP-Gene Set parameters
	<input checked="" type="checkbox"/> SNP-Disease Curated and literature-based associations from <u>DisGeNET</u>

SNP to gene annotation ×

HaploReg: LD proxies r^2

VEP: genes within kb

VEP: top nearest gene(s) within 50kb

PPI: include direct interactions [PPI Databases](#)

Network building

[Home](#) > [Upload](#) > Network Builder

Tables

The pair-wise tables together with the supporting information are listed below. You can click the table name to download the complete table, or browse the tables (max. 1000 entries). Click the **Proceed** button at the bottom to directly explore the results in a network context.

↓ snp2met	[SNP:12, Metabolite: 9]	▶ Browse
↓ snp2dis	[SNP:34, Disease: 69]	▶ Browse
↓ snp2gene	[SNP:34, Gene: 2]	▶ Browse
↓ protein2protein	[protein:8]	▶ Browse

Note: the output is based on the results after LD proxy search.

Networks

In some cases, multiple isolated networks will be generated, with a big 'continent' containing most of queries, and several small 'islands' containing one or a few queries. These networks will be available for visual analysis in the next step.

Networks	Queries	Nodes	Edges	Topology
↓ mgwas1	1	117	274	View
↓ mgwas2	1	4	4	View

Note: this is a reasonable size, therefore we don't need to perform network filtering.

Click "Proceed"

Network Tools: ?

Degree Filter

Betweenness Filter

Shortest Path Filter

Manual Batch Filter

Minimum Network

Steiner Forest Network

Reset Network

[<< Previous](#)

[>> Proceed](#)

Network overview

Here, we perform some basic customization to better visualize the network.

Note: input SNPs are indicated by the check mark.

The screenshot shows the Network Builder interface with a network visualization. The network consists of nodes and edges. Nodes are represented by colored squares and circles, and edges are represented by lines. The background is a gradient. The layout is Force Atlas. The node size is increased. The edge thickness is increased.

1) Change the background to "Gradient (dark)"

2) Change the layout to "Force Atlas"

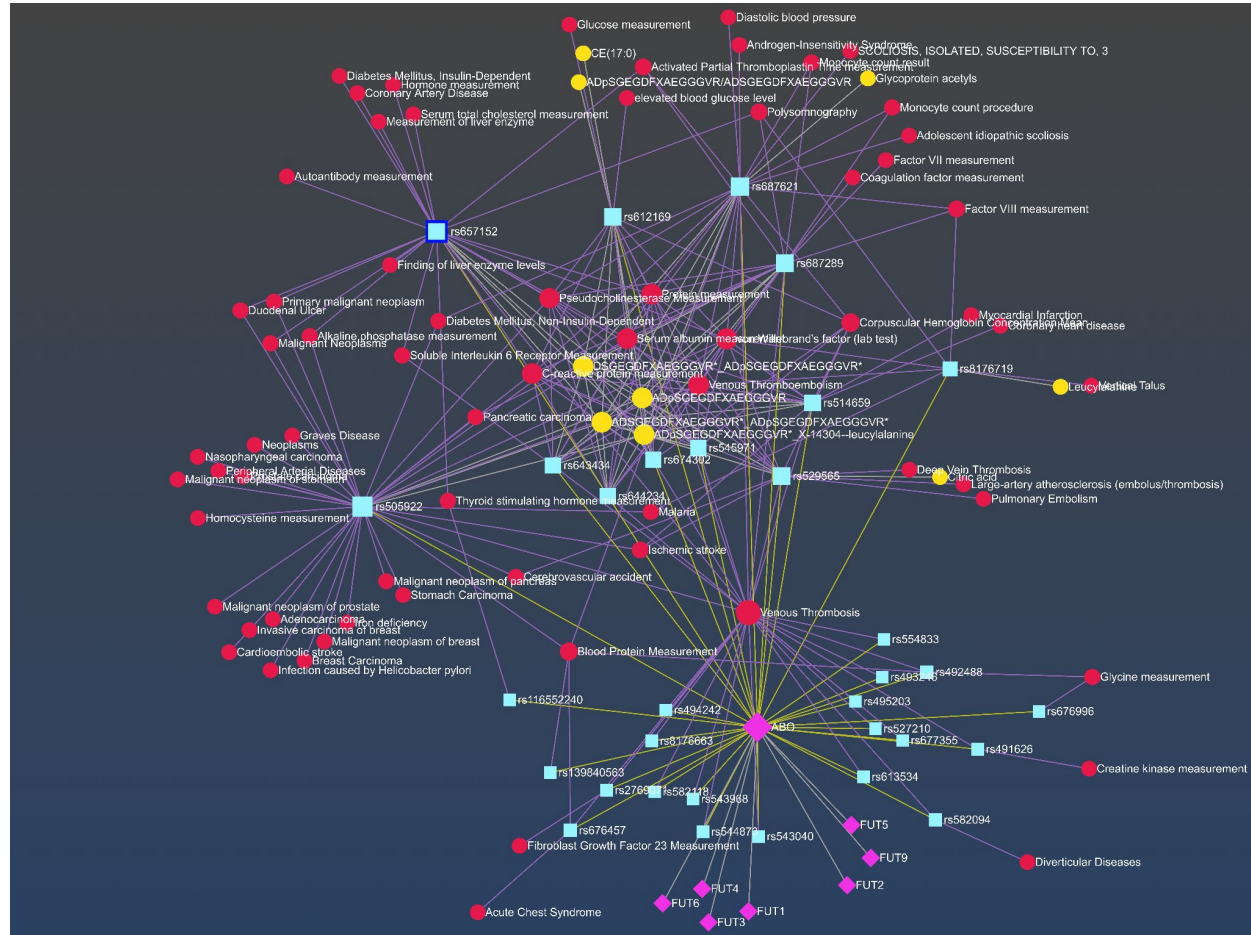
3) Increase the node size

4) Increase the edge thickness

ID	Name	Degree	Betweenness	Input
<input type="checkbox"/>	rs505922	39	2297.054	
<input type="checkbox"/>	ABO	39	2347.011	
<input type="checkbox"/>	Venous Thromb	28	1075.687	
<input type="checkbox"/>	rs657152	27	1201.567	<input checked="" type="checkbox"/>
<input type="checkbox"/>	rs687621	24	872.9469	
<input type="checkbox"/>	rs687289	20	519.5094	
<input type="checkbox"/>		19	574.4336	
<input type="checkbox"/>		17	505.5447	
<input type="checkbox"/>		15	300.7738	
<input type="checkbox"/>		13	51.54467	
<input type="checkbox"/>		13	85.01544	
<input type="checkbox"/>		13	51.54467	
<input type="checkbox"/>	rs644254	12	37.29362	
<input type="checkbox"/>	Venous Thromb	12	200.3215	
<input type="checkbox"/>	von Willebrandf	12	200.3215	
<input type="checkbox"/>	ADpSGEGDFX	11	133.2678	
<input type="checkbox"/>	ADpSGEGDFX	11	133.2678	
<input type="checkbox"/>	ADSGEGDFXA	11	133.2678	
<input type="checkbox"/>	DSGEGDFXAE	11	133.2678	

- mGWAS-Explorer discovered that the input SNP **rs657152** at the **ABO** locus is in high LD ($r^2 > 0.8$) with multiple other SNPs in this region.

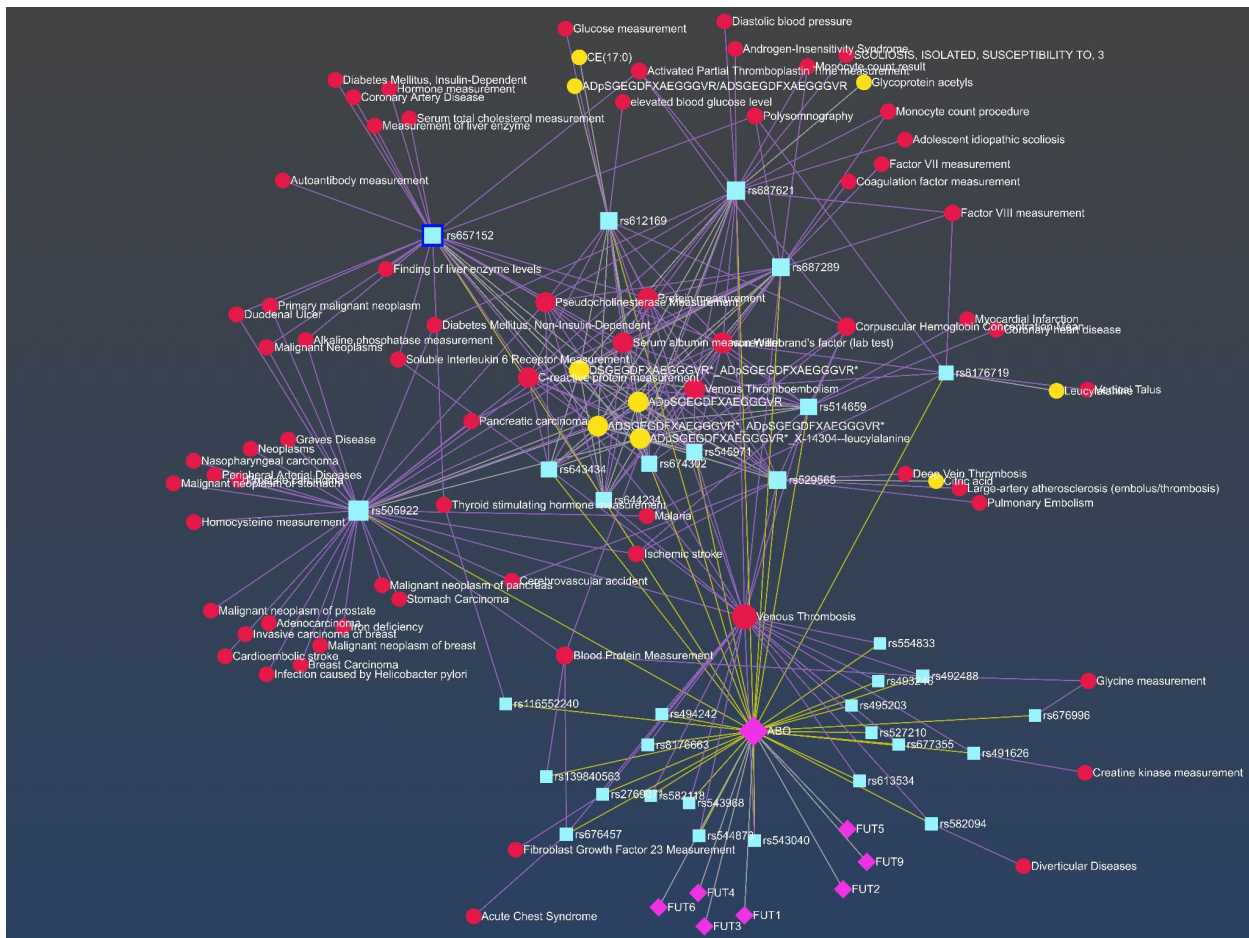
- The input SNP and SNPs in high LDs are associated with multiple metabolites and other human diseases, such as citric acid, malaria, ischemic stroke, and venous thrombosis.



- Interestingly, mGWAS-Explorer identified shared **ABO** variants between **COVID-19, fibrinogen A- α peptides** (e.g., ADpSGEGDFXAEGGGVR) and **venous thromboembolism**.

- Fibrinogen play a role in forming the blood clots.

- Therefore, the associations 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.



How to identify the shared SNPs between a disease and a metabolite?

1) Pick a color of interest

2) Select disease and metabolite of interest

3) Select "Shared" from the dropdown menu

4) Click "Submit"

Name	Degree	Betweenness
rs505922	39	225.00000
rs505922	39	225.00000
rs505922	28	100.00000
rs505922	27	12.00000
rs505922	24	872.00000
rs687289	20	519.5094
rs529565	19	574.4336
rs612169	17	505.5447
rs514659	15	300.7738
rs545971	13	51.54467
rs643434	13	85.01544
rs674302	13	51.54467
rs644234	12	37.29362
<input checked="" type="checkbox"/> Venous Thromb 12	200.3215	
von Willebrand's 12	200.3215	
<input checked="" type="checkbox"/> ADpSGEGDFX 11	133.2678	
ADpSGEGDFX 11	133.2678	
ADSGEGDFXA 11	133.2678	
DSGEGDFXAE 11	133.2678	
C-reactive prote 11	133.2678	

Current Selections

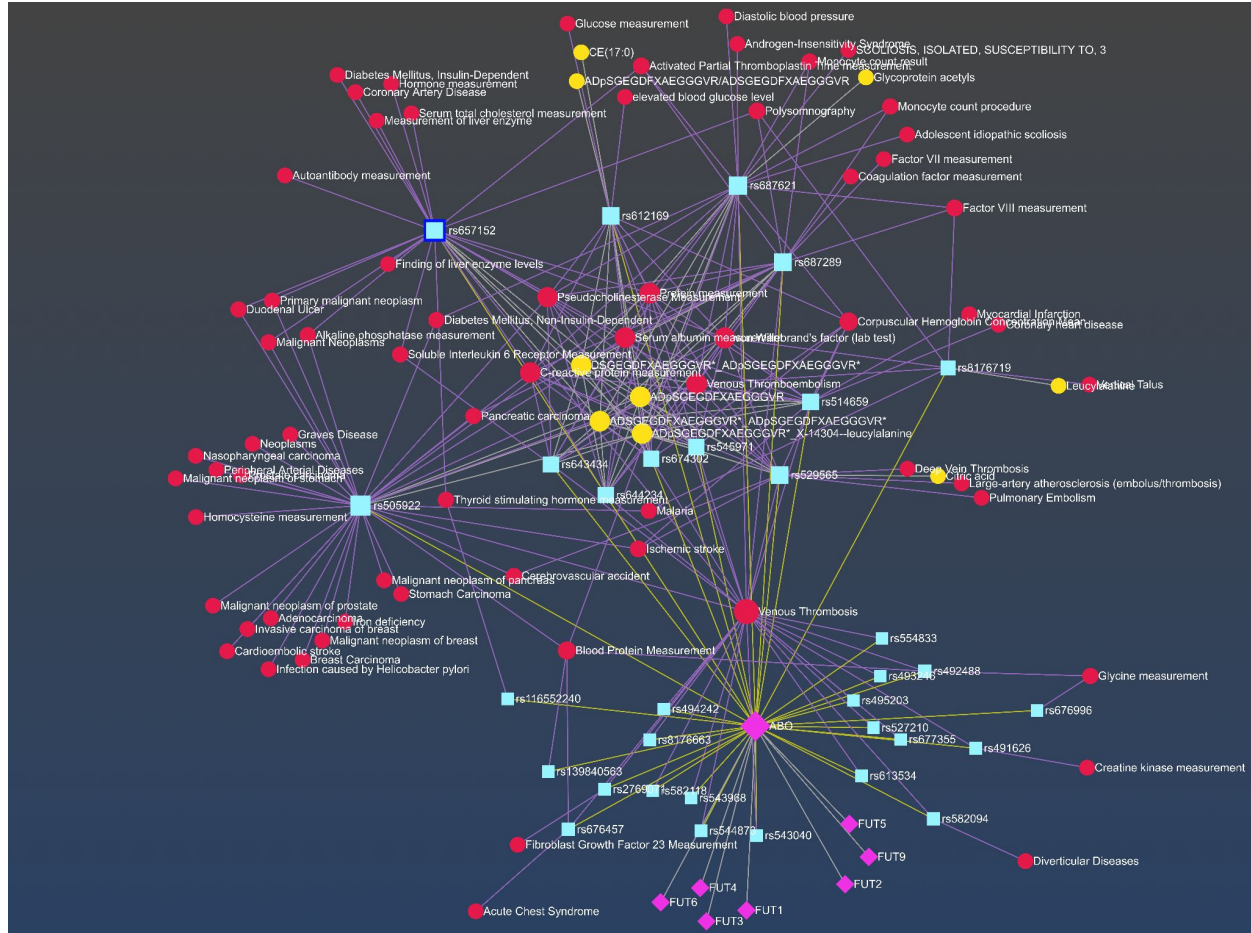
- Node: ADpSGEGDFXAEAGGVR

Cross-phenotype
association
analysis

- Indeed, studies have reported that COVID-19 is associated with an increased risk of thromboembolism.

- Therefore, we sought to investigate whether the association between **fibrinogen A- α peptides** associated loci could provide additional insights into underlying pathophysiologic mechanisms of COVID-19.

Let's take the list of the **fibrinogen A- α peptides** and start from the **Metabolite** module.



Start from the Metabolite Module

Follow similar workflow as the SNP module:

- Data Upload page: Click “Try Example” → “Yes” → “Submit” → “Proceed”
- Network Builder page: Click “Proceed”

Network result

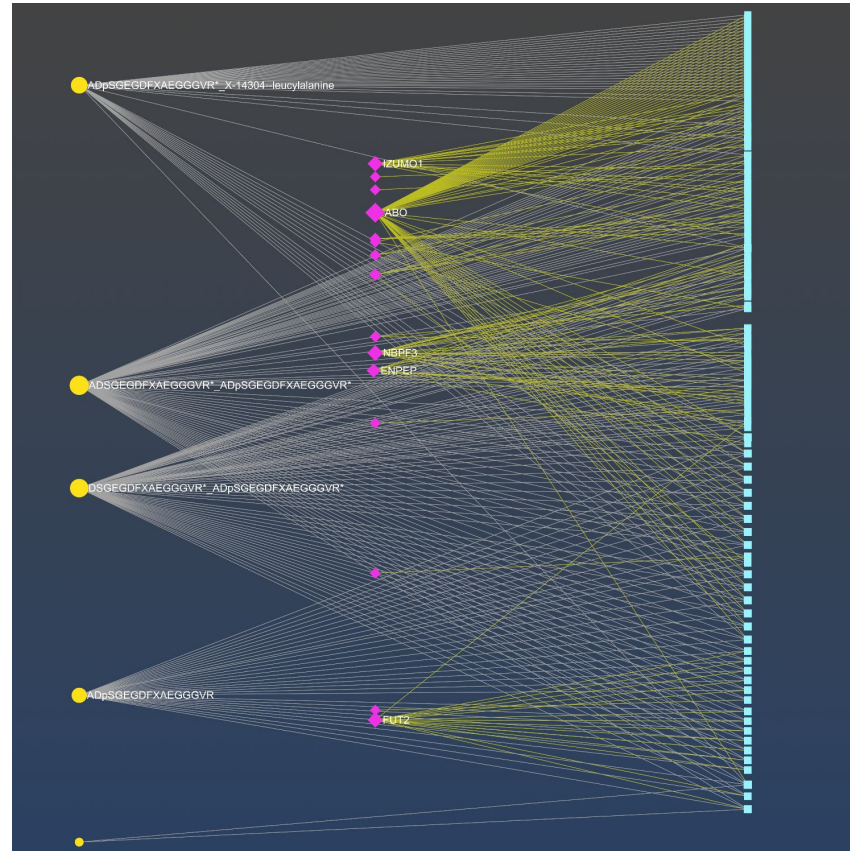
- mGWAS-Explorer revealed variants at *ENPEP* and *FUT2* genes are associated with levels and/or ratios of fibrinogen A- α peptides.

- In fact, *ENPEP* was discovered to be a candidate co-receptor for the coronavirus SARS-CoV-2 [1].

- And, individuals with an inactivating *FUT2* mutations were more likely to develop a less severe form of the COVID-19 disease [2].

1) 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. *Journal of medical virology* 2020

2) Mankelov, 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



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.

Type 2 Diabetes Case Study

Motivation

- Around 250 genomic regions have been associated to type 2 diabetes (T2D) susceptibility in genome-wide association studies [1].
- Some studies have highlighted the link to metabolomic profiles [2].

1. Langenberg, C.; Lotta, L.A. Genomic insights into the causes of type 2 diabetes. *Lancet (London, England)* 2018, 391, 2463-2474, doi:10.1016/s0140-6736(18)31132-2.
2. 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. *Nature Genetics* 2021, 53, 54-64, doi:10.1038/s41588-020-00751-5.

Objectives

Therefore, we applied mGWAS-Explorer to a list of SNPs from a published GWAS of T2D [1] with an attempt to examine shared genetic signals with circulating metabolites.

1. 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, doi:10.2337/db16-1253.

Methods

- Pre-processing:
 - Significant association p-value threshold (5×10^{-8})
 - LD clumping were performed to identify the independent signals by using the *ieugwasr* package
- Analysis overview: similar workflow as the COVID-19 case study
 - Upload a list of SNPs → select SNP list 2
 - Network building
 - Network analytics
 - Cross-phenotype association analysis

Network results

Here, we perform some basic customization to better visualize the network.

The screenshot shows a network visualization tool with a central network graph and several side panels. Five callout boxes point to specific settings:

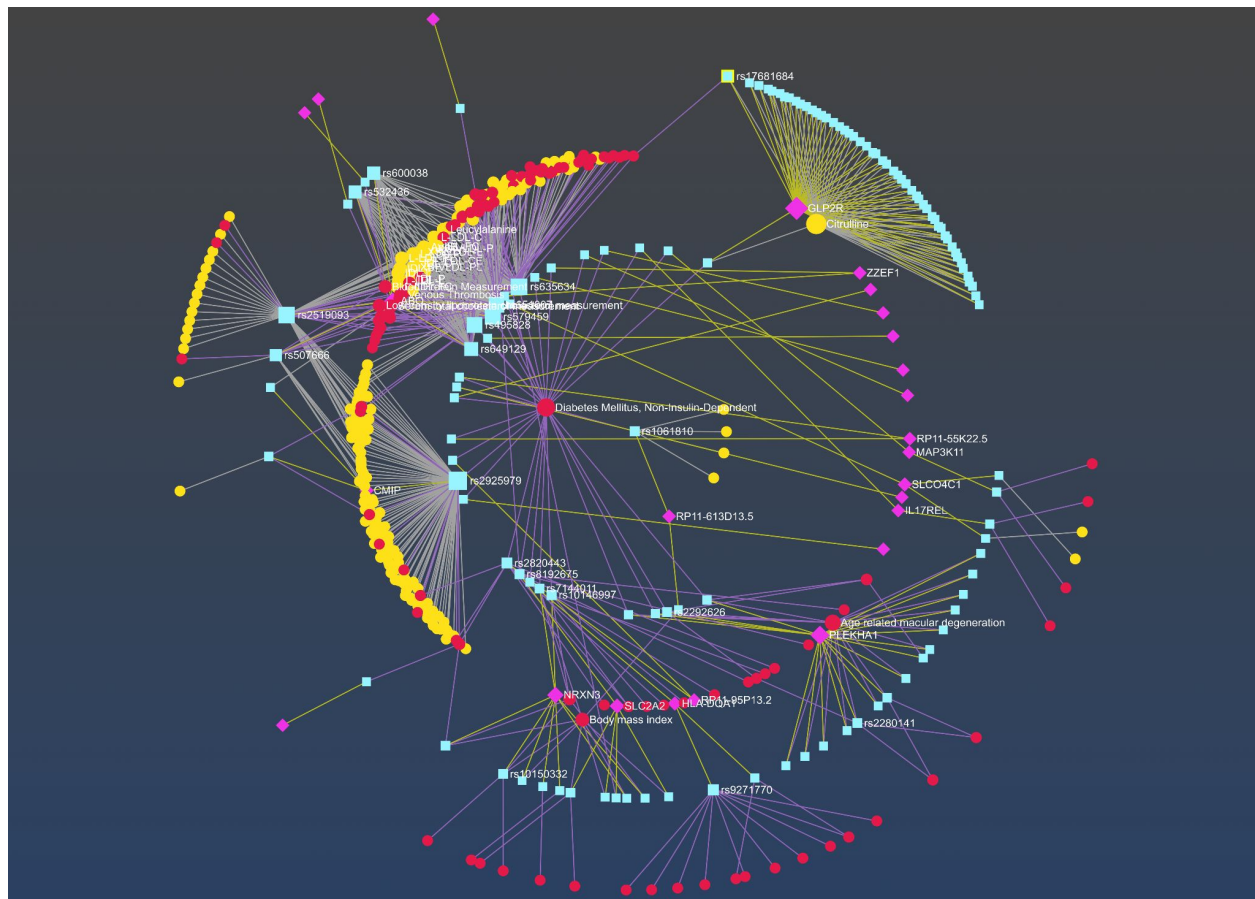
- 1) Change the background to "Gradient (dark)"
- 2) Change the layout to "Concentric circle"
- 3) Increase the node size
- 4) Increase the edge thickness
- 5) Adjust the node color by node type

The interface includes a top toolbar with options for Network, Background, View, Layout, Node, Edge, Scope, and Download. On the left, there are panels for Global Node Styles (with sliders for SNP, Metabolite, Disease), Node Explorer (with a table of node statistics), and a list of current selections. On the right, there are panels for Function Explorer and Module Explorer. The network graph itself is a complex, multi-colored network with nodes of varying sizes and colors (blue, yellow, pink, purple) connected by edges of varying thicknesses.

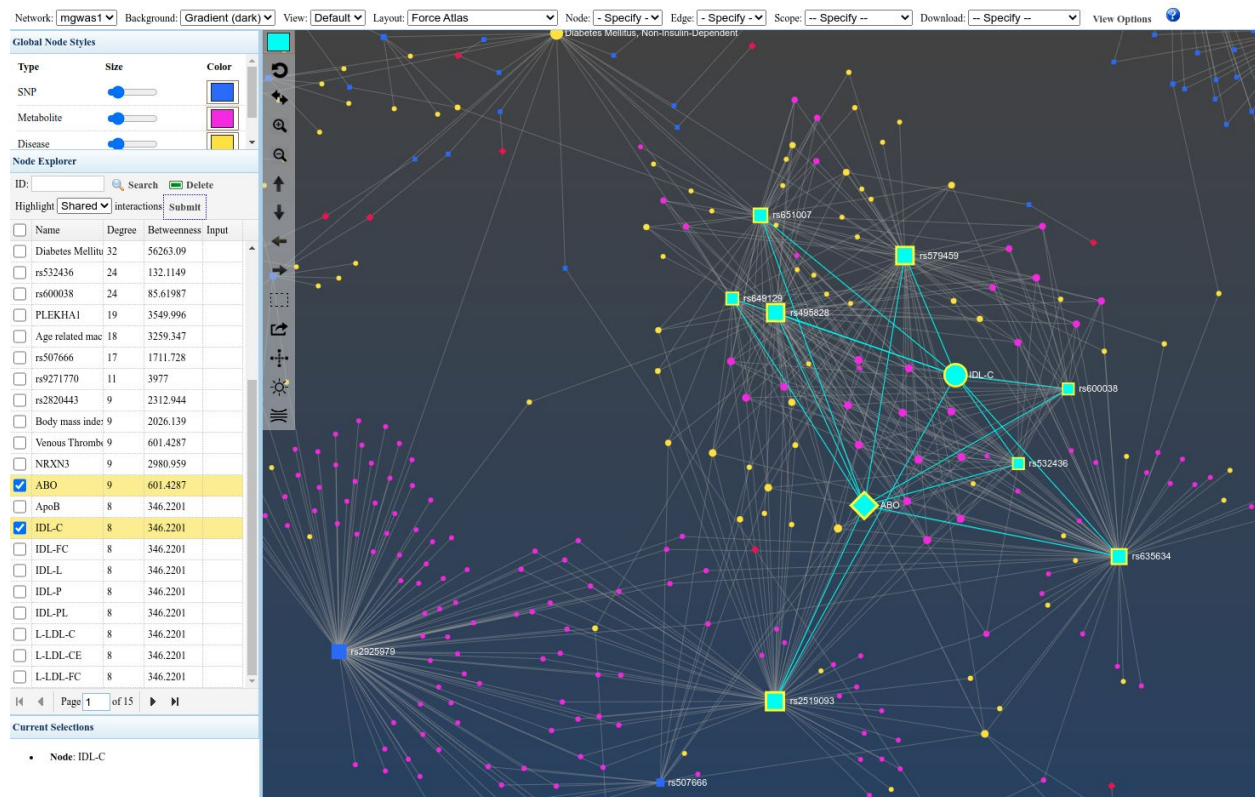
ID	Name	Degree	Betweenness
rs2925979		105	34849.21
rs635634			17707.77
rs649129		32	2220.576
Diabetes Mellitu	32		56263.09
rs532436		24	132.1149
rs600038		24	85.61987
PLEKHA1	19		3549.996
Age related mac	18		3259.347
rs507666		17	1711.728
rs9271770		11	3977
rs2820443		9	2312.944
Body mass index	9		2026.139
Venous Thromb	9		601.4287
NRXN3	9		2980.959

- mGWAS-Explorer confirmed the cross-phenotype associations between **citrulline** metabolite, **T2D**, and **body mass index** and identified the missense **rs17681684** variant for **citrulline** in the **GLP2R** gene as reported by Lotta et al. [1].

1) Lotta, L.A.; Pietzner, M.; Stewart, I.D.; Wittmanns, 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. Nature Genetics 2021, 53, 54-64, doi:10.1038/s41588-020-00751-5.

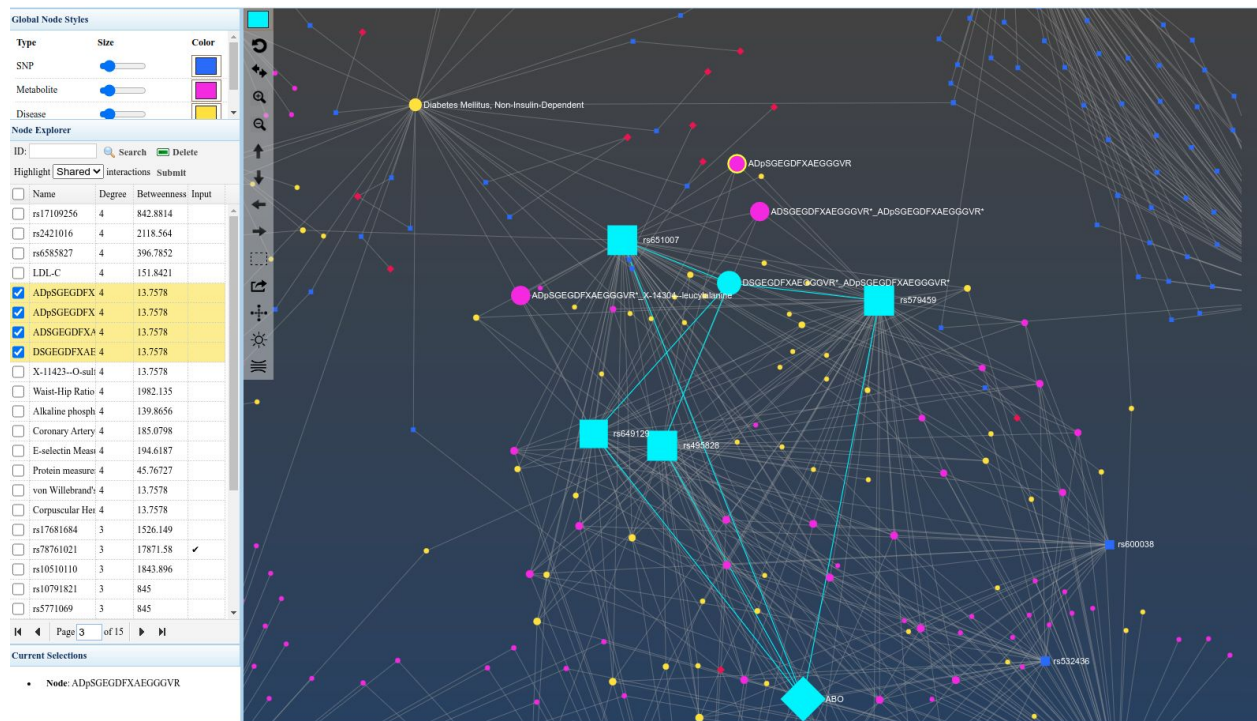


- Additionally, we identified shared genetic signals between **T2D, coronary artery disease** and **cholesterol levels** at ***ABO*** locus.



- Furthermore, mGWAS-Explorer also revealed **metabolites** levels and their ratios identified in the previous **COVID-19** case study shared associations with **T2D** associated SNPs at **ABO** locus.

- In fact, multiple studies have reported comorbidity of T2D and COVID-19 [1].



1) Rajpal, A.; Rahimi, L.; Ismail-Beigi, F. Factors leading to high morbidity and mortality of COVID-19 in patients with type 2 diabetes. Journal of diabetes 2020, 12, 895-908, doi:10.1111/1753-0407.13085.

In summary

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.

==The End==