KGG: A systematic biological <u>K</u>nowledge-based mining system for <u>G</u>enome-wide <u>G</u>enetic studies (Version 4.1)

User Manual

Miao-Xin Li, Lin Jiang and Jiang Li

Laboratory of Precision Medical Genomics Zhongshan School of Medicine Sun Yat-Sen University, Guangzhou City, Guangdong Province, 510080, PRC

1. Introduction and References	3
1.1 Introduction	3
1.2 References	3
2. Installation	4
2.1 Install Java Runtime Environment (JRE)	4
2.2 Install and initiate KGG	4
3. Interface and main functions	5
3.1 Project	5
3.2 Data	6
3.4 Gene	6
3.5 BioModule	6
3.6 Power	6
3.5 Tools	6
3.6 Window	6
4. Input files	6
4.1 Input file 1 (GWAS results)	6
4.2 Input file 2 (Candidate Gene list)	7
5. Tutorial of knowledge-based secondary association analysis	7
5.1: Data preparation	8
5.2 Secondary knowledge-based association analysis1	2
5.2.1 Gene-based association analysis by GATES1	.2
5.2.2 Gene-based association analysis by ECS1	.4
5.2.3 Conditional associational analysis by ECS on significant or interested genes1	.4
5.2.4 Gene-pair-based association analysis by HYST1	.5
5.2.5 Multivariate gene-based association analysis by MGAS1	.7
5.2.6 Estimating driver-tissues by selective expression of genes associated with comple	ЭХ
diseases or traits2	0
6. Power estimation of set-based tests by SPS2	1
7. Updates from KGG3.5 to KGG4.02	8

Content

Hints for large GWAS dataset (around or over 2.5 million SNPs)

Set or change large memory for KGG4 say, 2000MB, by *Tools->Set System Memory*.

1. Introduction and References

1.1 Introduction

 $KGG^{[1]}$ (<u>K</u>nowledge-based mining system for <u>G</u>enome-wide <u>G</u>enetic studies) is a software tool to perform knowledge-based secondary analysis with summary statistics from genome-wide association studies (GWAS). At present, the version 4 has been equipped with main functions to perform 5 types secondary Knowledge-based analysis by using SNP p-values from GWAS:

- Gene-based association^[2,3],
- Conditional gene-based association^[2],
- Multivariate gene-based association^[4],
- Gene-pair interaction-based association^[5],
- Geneset based association ^[6].

In addition, KGG has provided direct hyperlinks to several useful bioinformatics annotation databases on sequence variants (<u>http://jjwanglab.org/gwasrap</u>), genes (GeneCards, <u>http://www.genecards.org/</u>) and pathways (MsigDB, <u>http://www.broadinstitute.org/gsea/msigdb</u>). A number of functions to model emerging epigenomic regulatory data for prioritizing association signals are still under development.

1.2 References

- Li MX, Sham PC, Cherny SS, Song YQ. A knowledge-based weighting framework to boost the power of genome-wide association studies. PLoS One. 2010 Dec 31;5(12):e14480.
- 2. Li MX, Gui HS, Kwan JS, Sham PC. GATES: A rapid and powerful gene-based association test using extended Simes procedure. Am J Hum Genet. 2011 Mar 11;88(3):283-293.
- 3. Li et al. A powerful approach isolates independently associated genes of Schizophrenia with summary statistics from large-scale whole genome association meta-analysis (Submitted)
- 4. Sluis et al. MGAS: a powerful tool for multivariate gene-based genome-wide association analysis. Bioinformatics. 2015 Apr 1;31(7):1007-15.
- 5. Li MX*, Kwan JS*, Sham PC. HYST: A hybrid set-based test for genome-wide association studies, with application to protein-protein interaction-based association analysis. Am J Hum Genet. 2012 Sep 7;91(3):478-88.
- 6. Gui et al. Genome-wide gene- and gene-set-based association analyses identify novel patterns of genetic sharing across complex phenotypes(Submitted)
- 7. Li J, Sham PC, Song Y, Li M. SPS: a simulation tool for calculating power of set-based genetic association tests. Genet Epidemiol. 2015;39(5):395-7

2. Installation

2.1 Install Java Runtime Environment (JRE)

The Java Runtime Environment (JRE) v1.7 (or higher version) is required to run KGG4 on any operating systems (OS). It can be downloaded from <u>http://java.sun.com/javase/downloads/index.jsp</u> for free. Installing the JRE is very easy on Windows OS and Mac OS X.

On Linux, you have more work to do. Details of the installation can be found at <u>http://www.java.com/en/download/help/linux_install.xml</u>. In Ubuntu, if you have an error message like: "Exception in thread "AWT-EventQueue-0" java.awt.HeadlessException ...", then please installs the Sun Java Running Environment (JRE) first.

To install the Sun JRE on Ubuntu(10.04), please use the following commands: sudo add-apt-repository "deb http://archive.canonical.com/ lucid partner" sudo apt-get update sudo apt-get install sun-java7-jre sun-java7-plugin sun-java7-fonts

Detailed explanation of above commands can be found at <u>http://www.ubuntugeek.com/how-install-sun-java-runtime-environment-jre-in-ubuntu-10-04-lucid-lynx.html</u>.

Note: After completing Java installation, please make sure that not only the java is executable but also the extracted jre/bin directory is added to the PATH, otherwise KGG4 would not start properly. This is easily achieved by executing the following command on the terminal:

echo 'export PATH=/path/to/installed/jre/bin:\$PATH' >> ~/.bashrc && source ~/.bashrc

Thanks Attila Pulay for the suggestion!

2.2 Install and initiate KGG

To simplify the installation, we still keep KGG as a green tool (i.e., no formal installation procedure guided by an installation wizard). After decompressing the kgg4.zip file, you will see a "bin" folder where there are 3 script files to initiate KGG4. On Microsoft Windows, please double click kgg4.exe or kgg464.exe file. On Linux, Mac OS X and Solaris, please type the kgg4 in a Command-line Terminal.

If you have over 4 million variants, you are suggested set a larger memory for KGG. The default setting is 4GB. To do it, please click $\underline{\text{Tools}} \rightarrow \underline{\text{Set system memory}}$



Figure 2.0 Set system memory

3. Interface and main functions

Figure 3.1 shows a typical interface of KGG with an active project.



Figure 3.1 A typical KGG interface

Illustration:

Frame 1: tree-structured branches to manage input data and analysis results of a KGG project;

Frame 2: view of input data or output results;

Frame 3: running log of KGG analysis results;

The graphic dialogs of KGG are self-explanatory. Therefore, we will not elaborate the function of each buttons.

3.1 Project

- **Create Project**: create a new KGG project.
- > **Open Project:** open an existing KGG project.
- > Close Project: close the current project.
- **Exit:** exit the KGG application.

3.2 Data

- Load P value file: import your association summary results (e.g., the plink output).
- Define Seed Genes: tell KGG the known causal genes of the disease you are studying.
- Build Analysis Genome: build an analysis genome in which KGG maps all SNPs to their gene features and calculates the r-square or genotypic correlation of SNPs within genes.

3.4 Gene

- > Univariate Association: conduct a univariate gene-based association scan.
- > Multivariate Association: conduct a multivariate gene-based association scan.
- > View Genes: view and export gene-based association results.
- > LD Plot: view the LD pattern of variants within genes.
- **Conditional Association:** conduct a conditional gene-based association scan.

3.5 BioModule

- Gene-pair-based Association: conduct an association scan at gene pairs.
- **View gene-pairs:** view association p-values of gene pairs.
- Geneset-based Association: conduct an association scan at genesets.
- > View Geneset: view p-values of geneset-based association analysis.

3.6 Power

Calculator: SPS-a simulation tool for calculating power of set-based genetic association tests.

3.5 Tools

Set System Memory: set the memory of KGG.

3.6 Window

- > Analysis Log: show some summary results and log.
- > **Project:** show the structure of the working project.
- **Resource:** show the resource that KGG contains.
- > **PassedResultViewer:** show the log of a secondary analysis.
- RunningResultViewer: show the real-time running log when performing a secondary analysis.
- **TableViewer:** display the content of input p-value file and annotation file.
- > **Output:** show the IDE results.

4. Input files

4.1 Input file 1 (GWAS results)

KGG focuses on secondary analysis of GWAS p-values. The major input of KGG is the association p-values (produced by conventional statistical genetic methods, such as PLINK) in a text file. KGG supports a user-customized format for the association p-values. Once chromosome number (or chromosome number and physical position) and p-values columns are available in a file, you can define the column order by yourselves on KGG. The input files are allowed to include more than one p-value column. The following is an example.

Example input format of KGG:

CHR	SNPID	SNPPOS	P-value1	P-value2	P-value3	•••
4	Snp1	100001	0.02301	0.8815	0.007688	•••
4	Snp2	110011	0.4384	0.9575	0.006112	•••
4	Snp3	120001	0.002688	0.007688	0.4893	•••
4	Snp4	130011	0.01115	0.006112	0.119	•••
4	Snp5	140001	0.005892	0.4893	0	
	•••	•••	•••	•••	•••	•••

Additionally, a p-value column could include values of different models. During building the analysis genome, KGG can recognize this format ?(which format? The follow one?) with the input format "multiple tests per column".

Example a more complex input format of KGG:

CHR	SNP	P-value1	Test-Mode	P-value2	•••
4	rs1513559	0.02301	additive	0.007688	
4	rs1513559	0.4384	recessive	0.006112	
4	rs1513559	0.002688	dominant	0.4893	
4	rs1841043	0.01115	additive	0.119	
4	rs1841043	0.005892	recessive	0	•••
•••					•••

4.2 Input file 2 (Candidate Gene list)

Candidate genes could be loaded one by one or imported from a TXT file. The input file has only one column without header, in which one row contains only one gene (symbol or ID).

5. Tutorial of knowledge-based secondary association analysis

We use real dataset of Crohn's disease (available а at http://grass.cgs.hku.hk/limx/kgg/download/KGGSample.zip) as an example to demonstrate how to use KGG for a series of knowledge-based secondary association analysis of conventional p-values from GWAS. This dataset was originally downloaded from a public domain released by (Barrett, et al., 2008) and have SNP ID conversion by SNPTracker (http://grass.cgs.hku.hk/limx/snptracker/) for coordinates of Hg19. It includes 7 columns, as CHR, SNP, POS, RISK, NONRISK, META-Z and META-P. The effective input data in the input summary statistics file are chromosome (CHR), coordinate (POS) and variants' p-values (META-P). The main analysis procedure is illustrated in Figure 5.0.



Figure 5.0 Pipeline chart of KGG analysis (version 4)

Notes: Circle nodes stand for data and files (input, output), rectangles denote an analytical procedure, a dashed line stand for virtual relationship between a dataset and an analysis.

5.1: Data preparatio	n		
<u>To create a new pro</u> 'CrohnDisease', and s	<u>oject,</u> please c set the project	lick the menu <u>Project</u> \rightarrow <u>Create</u> path at C:\KGG (or other path d	<u>Project</u> , with a name lefined by users).
	🌆 Create KGG Proje	ct	×
	Project Name: Working Folder: Description:	CrohnDisease D:\KGG\ The knowledge-based downstream genetic/geno mic statistical analysis for XXX disease	
Load the p value f	Figure 5	.1.1 Dialog of creating project	ata J ood P Valua Fila

<u>Load the p-value file</u> into the project. Please select the menu <u>Data</u> \rightarrow <u>Load P Value File</u> and choose the file 'CrohnGWASresultHg19.txt' containing whole-genome association p-values for Crohn disease in KGGSample folder.

rojec Resource	🎹 TableViewe	er Window 🗙 🎬	RunningResul	tViewer \times			• • •	•
CrohnDisease	CHR	SNP	BP	RISK	NONRISK	MET A-Z	META-P	
P-value Files:	1	rs3094315	752566	A	G	1.208042	0.227031	
🐜 🚮 CrohnGWASresultHg19	1	rs4040617	779322	A	G	0.5591984	0.5760264	٦
	1	rs2980300	785989	С	Т	0.5241999	0.6001394	-
	1	rs4075116	1003629	T	С	2.66553	0.007686718	٦
	1	rs3934834	1005806	T	С	1.319292	0.18707166	٦
	1	rs3737728	1021415	G	A	2.474539	0.01334083	٦
	1	rs6687776	1030565	Т	с	2.292393	0.02188298	
	1	rs9651273	1031540	G	A	0.7116839	0.4766606	
	1	rs4970405	1048955	G	A	1.140031	0.2542734	
	1	rs12726255	1049950	G	A	1.580504	0.1139915	
	1	rs2298217	1064979	c	Т	0.09809688	0.9218554	
	1	rs4970362	1094738	G	A	0.02632069	0.9790016	٦
	1	r s9442385	1097335	Т	G	0.2917067	0.770511	٦
	1	rs9660710	1099342	A	C	0.1359162	0.8918876	
	1	rs4970420	1106473	A	G	2.418564	0.015581894	٦
	1	rs1320565	1119858	Т	C	1.229683	0.218816	
	1	rs11260549	1121794	A	G	2.183678	0.02898592	
	1	rs10907175	1130727	C	A	1.449057	0.14732166	
	1	rs9729550	1135242	C	A	3.072463	0.002123002	٦
	1	rs11721	1152631	A	C	2.538362	0.011137286	٦
	1	rs2887286	1156131	С	Т	1.392902	0.16364952	
								-

Figure 5.1.2 Input GWAS original result file

<u>Define a number of candidate genes</u>. Click the menu <u>Data</u> \rightarrow <u>Define Seed Genes</u> to import file 'CrohnCandidateGeneSet.txt' as input of candidate genes. Define all genes as seed genes and save them as candidategeneset_crohn. Note: this step is optional and the seed genes will be only used to highlight gene pairs and gene sets.

		To Be Selecte	ed					
Gene Symbol	\sim	Source	Symbol	EntrezID	Name	Chromosome	As Seed	
1.0.0.								
lateGeneSet.								None
Genes								
								A11
								Remove
ndidate Gene	View							Add
ndidate Gene	View es Symbol		EntrezID	Name		Chromosome	As Seed	Add
ndidate Geno purce put	View s Symbol STAT3		EntrezID 6774	Name signal trar	isducer and	Chromosome 17 o21	As Seed	Add
ndidate Geno purce put put	View es Symbol TLR4		EntrezID 6774 7099	Name signal tran toll like r	isducer and	Chromosome 17q21 9q33.1	As Seed	bh A
ndidate Gena purce put put put	View Symbol SIAT3 TLR4 TIFRSF6B		EntrezID 6774 7099 8771	Name signal trar toll like r tuno n ecro	isducer and eceptor 4 sis factor	Chromosome 17q21 9q33.1 20q13.33	As Seed	bbA V
ndidate Geno purce put put put put	View Symbol STAT3 TLR4 THFRSF0E THFRSF0E		EntrezID 6774 7099 8771 9966	Name signal trar toll like r tumor necro tumor necro	isducer and receptor 4 osis factor sis factor	Chromosome 17q21 2q33.1 20q13.33 2q32	As Seed	Add
ndidate Geno purce put put put put put	View Symbol STAT3 TLR4 THFSF6E THFSF6E XSF1		EntrezID 6774 7099 8771 9966 7494	Nume signal trar toll like r tumor necro tumor necro K-box bind	isducer and ecceptor 4 sis factor sis factor ng protein 1	Chromosome 17g21 9g33.1 20g13.33 9g32 22g12.1	As Seed	Add V V V V
ndidate Geno purce put put put put put put put	view Symbol STAT3 TLR4 TMFRSF6B TMFSF15 XBP1 ORMDL3		EntrezID 6774 7099 8771 9966 7494 94103	Name signal trar toll like r tumor necro X-box bindi ORMDL sphir	sducer and eceptor 4 sis factor ng protein 1 golipid bio	Chromosome 17q21 9q33.1 20q13.33 9q32 22q12.1 17q12	As Seed	bbA bbA V V V V V
ndidate Geno vurce put put put put put put put put put put	View Symbol SIAI3 TLR4 THFRSF6E THFSF15 XBF1 ORMDL3 CCR6		EntrezID 6774 7099 8771 9966 7494 94103 1235	Name signal trax toll like x tumor necro X-box bindi ORMDL spin C-C motifo	sducer and eceptor 4 sis factor sis factor ng protein 1 ugolipid bio hemokine re	Chromosome 17q21 9q33.1 20q13.33 9q32 22q12.1 17q12 6q27	As Seed	
ndidate Geno put put put put put put put put put put	View Symbol STAT3 TLR4 THFRSF6E THFSF15 XBP1 ORMDL3 CCR6 CARD9		EntrezID 6774 7099 8771 9966 7494 94103 1235 64170	Name signal trar toll like r tumor necro X-box bindi ORMDL sphir C-C motif c caspase rec	receptor 4 osis factor ng protein 1 rgolipid bio hemokine re ruitment do	Chromosome 17q21 9q33.1 20q13.33 9q32 22q12.1 17q12 6q27 9q34	As Seed	Add V V V V V V V V V V
ndidate Geno purce put put put put put put put put put put	View Symbol STAT3 TLR4 THFSP56 THFSP56 THFSP56 THFSP56 THFSP56 CR5 CR6 CARD9 ICOSL6		EntrezID 6774 7099 8771 9966 7494 94103 1235 64170 23308	Name signal trar toll like r tumor necro tumor necro K-box bind ORMDL sphir C-C motif c caspase rec inducible T	sducer and eceptor 4 sis factor ng protein 1 ngolipid bio hemokine re "roell co-st	Chromosome 17q21 9q33.1 20q13.33 9q32 22q12.1 17q12 5q27 9q34 21q22.3	As Seed	Add V V V V V V V V V V V V V

Figure 5.1.3 Input candidate gene set for crohn's disease

<u>Build an analysis genome.</u> Click the Menu <u>Data</u> \rightarrow <u>Build Analysis Genome.</u>

Download ancestry matched genotypes in 1000 Genomes Project to adjust LD between variants. Click the 'Download' hyperlink on the Dialog to go to a web-page for downloading the genotypes at <u>http://grass.cgs.hku.hk/limx/kgg/phasedgty.html</u>. Choose 'EUR(495 subjects)' and click the hyperlink <u>lkg.phase3.v5.shapeit2.eur.hg19.tar.gz</u>. Unpack the data into 23 compressed VCF files corresponding to 23 chromosomes. Click **i** to load all the "vcf.gz"files on the "Build Analysis Genome" dialog.

Select <u>META-P</u> for building analysis genome with the default setting and clink the button **build** to build analysis genome. It will take a round 15 miniutes on a notebook computer with a CPU with 2.0GHz.

Genome Name: genome1	Px	value files:	CrohnGWASresu	ıltHg19.txt			~
CHR Select SNP OWE or EP WORE RISK NONRISK column(s) "Ctrl" for META-2	CHR SNP 1 rs3094315 1 rs4040617 1 rs2980300 1 rs4075116 1 rs3934834	BP 752566 779322 785989 1003629 1005806	RISK A C T T	NONRISK G G T C C	META-Z 1.208042 0.5591984 0.5241999 2.66553 1.319292	META-P 0.227031 0.5760264 0.6001394 0.007686718 0.18707166	^
selection Extended gene region length Ignore Low LD: r2 < 0.005	1 rs5/3/728 1 rs6687776 1 5 kb at 5' : [1021415 1030565 5 kb at	: 3'	A C t Correlation Setting	2.474539 2.292393 n matrix of p	0.01334083 0.02188298 henotypes for m	v ultivariate analysis only
Adjust by genomic control Gene defintion database Re	(divided by chi−square efGene ∽	median) 🗌	Chr Mar Imp	omsome Colum ker ID Colum utation Qual	nn: CHR nn: SNP .ity Column:	✓ Marker Po ✓ Marker (Optional) ✓	osition Column: BP V Position Version: hg19 V
LD Data Haplotypes (VCF format) Ge Genome Coordinates Versio D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s D:\KGG\EUR\lkg.phase3.v5.s	enotypes (Plink formet) on hg19 shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr] shapeit2.eur.hg19.chr]	ownload	No Inp Mis Regi Mod	positions?! ut type p-v ut format S sing data 1s ons (Options el: Exclusio	Get positions alues ingle test per ibel IIA 1) m ~ In fro t	s of SHPs by SHF v r column v Has T chromosome v m o	PTracker itle Row in the association file bp bp bp Cancel

Figure 5.1.4 Select META-P to build analysis genome and name the genome as genome_crohn



5.2 Secondary knowledge-based association analysis

5.2.1 Gene-based association analysis by GATES

Click <u>Gene</u> \rightarrow <u>Univariate Association</u> to set the parameters as Figure 5.5.1. Set the Scan name as 'genome_crohn', and select SNP p-values to integrate the analysis genome, then choose 'Extended Simes test (GATES, more powerful for a gene with one or a few independent causal variants' method. It should be noted that exported Manhattan plots and QQ plots will be shown in "Running Result Viewer Window" (Figure 5.5.2).

		Mannattan plot display
alysis Genomes Gene (Groups	Label genes with p-values <= 1E-6 Width 1200
enome Set genome1	🗸 🔽 v Snip	Label SNPs with p-values <= 5E-8 Height 500
P Value Name	Select	Minimal p-value 1E-10 Manhanttan plot SNPs outside genes
IETA-P		00 plot display
		QQ plot SNPs inside genes 🗹 Width 600
		QQ plot SNP outside genes V Height 400
his analysis genome h	as NO phenotype correlation matrix	! Minimal p-value 1E-10
his analysis genome h	as NO phenotype correlation matrix	! Minimal p-value 1E-10
his analysis genome h ethods	as NO phenotype correlation matrix	! Minimal p-value 1E-10
his analysis genome h thods Extended Simes test (as NO phenotype correlation matrix GATES, more powerful for a gene wi	! Minimal p-value 1E-10
his analysis genome h thods Extended Simes test (☑ Ignore single-nucle	as NO phenotype correlation matrix GATES, more powerful for a gene wi cotide polymorphisms (SNPs) without	Minimal p-value 1E-10 th one or a few independent risk variants) t linkage disequilibrium (LD) information
his analysis genome h ethods Extended Simes test () ✓ Ignore single-nucl Reference Li MX. Gui HS. Kw	as NO phenotype correlation matrix GATES, more powerful for a gene wi eotide polymorphisms (SNPs) without an JS, Sham PC, GATES: A rapid	Minimal p-value 1E-10 th one or a few independent risk variants) t linkage disequilibrium (LD) information d and powerful gene-based association
his analysis genome h ethods Extended Simes test (✓ Ignore single-nucl Reference Li MX, Gui HS, Kwa test using extende	as NO phenotype correlation matrix GATES, more powerful for a gene wi eotide polymorphisms (SNPs) without an JS, Sham PC. GATES: A rapid ed Simes procedure. Am J Hum G	! Minimal p-value IE-10 th one or a few independent risk variants) t linkage disequilibrium (LD) information d and powerful gene-based association Genet. 2011 Mar 11;88(3):283-293
his analysis genome h ethods Extended Simes test (✓ Ignore single-nucl Reference Li MX, Gui HS, Kwa test using extende	as NO phenotype correlation matrix GATES, more powerful for a gene wi eotide polymorphisms (SNPs) without an JS, Sham PC. GATES: A rapid ed Simes procedure. Am J Hum G	Minimal p-value 1E-10 th one or a few independent risk variants) t linkage disequilibrium (LD) information d and powerful gene-based association Genet. 2011 Mar 11;88(3):283-293



Figure 5.2.1.2 The display after gene-based scan

If you want to view the detail results, please click the "Show: Detailed Results" node under "Genome Scan" in the left frame. The new tab named "ViewGenes" will be created to provide you more information about the result (Figure 5.5.3). You can also export the results you want in this tab.

ohnDisease		esuitviewer	^ <u>1ab1</u>	eViewer Wi	ndow 🗙 🎱 V	ie≢ Genes ≻	۲.				+ >
	Gene Info							SNP In	nfo	Search gene:	0R2#1
P-value Files:	Symbol	NominalP	Corrected	Chromo	Start_P	Group			Position	Retrieve	ZSCAW31
🚮 CrohnGWASresultHg19.txt	IL23R	1.63E-34	3. 778873	1	67632168	protein		0010	01000131		
Genome: genome1	NOD2	2.32E-31	2.679584	16	50733260	protein-		39675	67722216	-Multiple Ter	
🖋 Source : CrohnGWASres	ltHg19.txt5	3.07E-20	2.364267	2	234184371	non-codi		.841	67694202	multiple les	
🖋 Versien : he19	ATG16L1	6.08E-20	3.511987	. 2	234160216	protein		1819	67670213	Method:	Benjamini & Hochberg (1995) 🗸
Cons. databasa - RofG	Clorf141	3.59E-16	1.662082.	1	67557858	protein		3151	67719129	Farmer Pater	0.0E Apply
Icod columns - Durns.	PIGFRN	3.91E-12	1.508359	. 1	117452544	protein		689	67659421	birdr Rate.	Аррту
- usea columns : [MEIA	CYLD	8.28E-12	2.734493	16	50776028	protein		39630	67662622		
Gene 5 extension :	D. IRGM	1.38E-10	3.974699	. 5	150226084	protein		331	67676922	Export Setti	ng
Gene 3' extension :	5. RRP15	4.59E-10	1.179195	. 1	218458628	protein		39671	67705726	Content:	Variants inside genes $$
SNP LD : 1KG Haplotyp	s LINC01475	6.41E-10	1.346762	10	101286106	non-codi)9026	67705958	_	- 1/
🗸 Adjust P Value : No	NKX2-3	6.41E-10	1.346762	10	101292689	protein		39629	67688349	Format:	Excel(V Export
Gene Scan: genome crohn	PTPN2	7.94E-10	1.530658	18	12792300	protein-		333	67678993	Come a val	
Genome : genomel	RFTN1	1.20E-9	2.139483.	. 3	16357351	protein		332	67678922	Gene p var	
Wultinguista Tart :	GLT1D1	2.27E-9	3.743640.	12	129337978	protein-		1056	67627260 ¥	Variants p	values <= 5E-2
	SAG	6 148-9	9 471889	2	234216308	motein-	~	<	>		
P value sources : []	514										
Test : GATES	Genome Bro	wser									
🖉 Show : Detailed Result	s					Como: T	1 2 2		25 24)		
						Gene: 1	_23	K(1.0.	55-34)		
	• [0.0-0.2)	• [0.2-0.4)	<mark>•</mark> [0.4-0.6) •	[0.6-0.8)	0.8-1.0) • 1	0 • ref			Exon 🗢 Intro	n-UTR-UpDownS	stream 🔷 ncRNA 🔺 InterGene 🔻 Oth
	• [0.0-0.2)	• [0.2-0.4)	<mark>-</mark> [0.4-0.6) -	[0.6-0.8)	0.8-1.0) • 1	0 • ref			Exon • Intro	n-UTR-UpDownS	stream 🔶 ncRNA 🔺 InterGene 🔻 Oth
	• [0.0-0.2) 35	• [0.2-0.4)	<mark>-</mark> [0.4-0.6) -	[0.6-0.8)	0.8-1.0) • 1	0 ● ref			Exon • Intro	n-UTR-UpDownS	stream 🔶 ncRNA 🔺 InterGene 🔻 Otr
	• [0.0-0.2) 35 -	• [0.2-0.4)	<mark>•</mark> (0.4-0.6) •	[0.6-0.8)	0.8-1.0) • 1	0 ● ref			Exon • Intro	n-UTR-UpDownS	stream ♦ ncRNA ▲ InterGene ♥ Otr
	• [0.0-0.2) 35 30	• [0.2-0.4)	• [0.4-0.6) •	(0.6-0.8)	0.8-1.0) • 1	0 ● ref			Exon • Intro	n-UTR-UpDownS	tream ◆ ncRNA ▲ InterGene ♥ Otr
	• [0.0-0.2) 35	• [0.2-0.4)	● [0.4-0.6) ●	0.6-0.8	0.8-1.0) • 1	0 ● ref			Exon • Intro	n-UTR-UpDownS	tream ◆ ncRNA ▲ InterGene ♥ Oth
	(0.0-0.2) 35 30 30 25	• [0.2-0.4)	● [0.4-0.6) ●	0.6-0.8	0.8-1.0) • 1	0 • ref			Exon Intro	n-UTR-UpDownS	tream ◆ ncRNA ▲ InterGene ♥ Oth
	• [0.0-0.2) 35 30 (a) 25 [b)	• [0.2-0.4)	● [0.4-0.6) ●	0.6-0.8)	0.8-1.0) • 1	0 • ref			Exon • Intro	n-UTR-UpDownS	tream ● ncKNA ▲ interGene ♥ Otr
	• [0.0-0.2) 35 30 (a) 25 (b) 20	• [0.2-0.4)	• [0.4-0.6) •	(0.6-0.8)	0.8-1.0) • 1	0 • ref	•		Exon Intro	n-UTR-UpDownS	
	• [0.0-0.2] 35 30 (an) 25 (b) 20 (c) 15	• [0.2-0.4)	0 [0.4-0.6)	• [0.6-0.8) ●	0.8-1.0) • 1	0 • ref	•		Exon • Intro	e	itream ● ncKNA ▲ interGene ♥ Otr
	• [0.0-0.2] 35 30 • 25 • 20 • 0 • 15	• [0.2-0.4)	(0.4-0.6)	• [0.6-0.8) •	• [0.8-1.0) • 1	0 • ref	•		Exon • Intro	e	

5.2.2 Gene-based association analysis by ECS

Similarly, you can use another approach, an effective chi-squared test (ECS), to perform the gene-based association analysis. Effective chi-squared test (ECS, more powerful for a gene with multiple dense independent risk variants). The ECS is more powerful than GATES when a gene has multiple dense independent risk variants. ECS also lends itself for a conditional gene-based association analysis.

5.2.3 Conditional associational analysis by ECS on significant or interested genes

Click the menu <u>Gene</u> \rightarrow <u>Conditional Associational</u> to set the parameters as Figure 5.6.1. Click Load Genes and choose 'Method: Benjamini & Hochberg (1995)' to selected significant genes. By default, genes within 5 MB form a group. After the analysis, the significance of the conditional associational analysis will be shown in the column "ISScore", which is a p-value-like measurement of the significance.

Gene Symbols Selec Method: Benjamini Error Rate: 0.05	ted Regions & Hochberg	Sig. Genes (1995)	~		Gene As	sociation	Set: genome_	crohn_e	cs
Method: Benjamini Error Rate: 0.05	& Hochberg	(1995)	~						
Error Rate: 0.05		/			P Volute	Source	META-P		- C
Error Rate: 0.05					i varue	oource.	MUTA 1		
	Group	Distance: 50	00000 bp		Load	Genes	Update Ra	nks	Reset Ran
Group Gene	Chromo	StartP	OriginalP	#SNP	Rankin	Select	ISScore		a 1
PERS	1	7844488	4 718-5	19	4 718-5		2		SelectAll
UTS2	1	7907671	2.13E-6	5	2.13E-6		?		UnselectAl
2 DMRTA2	1	50883222	0.00013	1	0.00013		?		
MIER1	1	67395925	7.32E-7	14	7.32E-7		?		Run
C1 or f141	1	67557858	5.77E-24	25	5.77E-24		?		Franch
IL23R	1	67632168	1.36E-59	31	1.36E-59		?	-	Expor
PHTF1	1	114239823	6.68E-5	5	6.68E-5		?	-	
KSBN1	1	114304453	0.00010	12	0.00010		2	-	
AP481-AS1	1	114399256	7 998-5	24	7 998-5		?	-	
ARHGEF2	1	155916629	0.00048	4	0.00048		?		
CD48	1	160650211	3.71E-6	13	3.71E-6		?		
i ITLN1	1	160846329	3.02E-7	2	3.02E-7		?		
5 L0C1019	. 1	160902254	4.62E-5	5	4.62E-5		?		
i ITLN2	1	160914815	0.00020	5	0.00020		?	_	
SWT1	1	185126290	0.00021	21	0.00021		?	-	
Clorf106	1	200863948	0.00019	7	0.00019		?	-	
KIF21B	1	200938513	1.198-5	7	1.198-5		?	-	
ACBD3	1	226332379	8.09E-5	12	8 098-5		2		
0 L0C1019	2	10589853	0.00027	2	0.00027		?		
1 DIMT3A	2	25504320	0.00012	16	0.00012		?		
			+	-				-	
	Gene Gene PER3 PER3 UTS2 UMRTA2 MIER1 Clorfl41 IL23R PHTP1 PHTP1 RSBH1 PHTP1 AP4B1-AS1 ARNGEP2 CD48 ITLL1 STT1 Clorfl06 KIF21B RRP15 ACBD3 0 LOC10191	Gene Chrome PER3 1 UTS2 1 DMRTA2 1 Immediate 1	Gene Chromo StartF PER3 1 7844488 UTS2 1 7907671 JMRTA2 1 50833222 MIERI 1 67357858 Clorfl41 1 67557858 IL23R 1 67632168 PHTF1 1 114329823 RSBN1 1 114304433 PHTF12 1 114399256 AF4B1-AS1 1 114399256 ARHOFP2 1 155916629 CD49 1 160650211 ITLM1 1 160946329 CO49 1 16090254 ITLM2 1 16094815 SWT1 1 153126290 Clorf106 200863948 KIP21B Q00938513 RRP15 1 218458628 ACED3 1 226332379 0 LOC1019 2 10589633 1	Gene Chromo StartP OriginalP PER3 1 7844488 4.718-5 UTS2 1 7907671 2.138-6 DMRTA2 1 50883222 0.00013 MIER1 1 67357858 5.778-24 IL33R 1 67652168 1.38E-6 PHTF1 1 14329823 6.68E-5 RSBM1 1 14356432 0.00010 PTFN22 1 14356432 0.00016 PTFN22 1 14356432 0.00016 PTFN22 1 155916629 0.00048 CD48 1 160950211 3.71E-6 ITLM1 1 160946329 3.02E-7 LOC1019 1 160946329 3.002E-7 ITLM1 1 160946329 3.02E-7 ITLM2 1 160914815 0.00020 SWT1 1 185126290 0.00020 SWT1 1 185126290 0.00021 <t< td=""><td>Gene Chromo StartP OriginalP #SNP PER3 1 7844488 4.71E-5 19 UTS2 1 7907671 2.13E-6 5 DMRTA2 1 50883222 0.00013 1 MIER1 1 67395925 7.32E-7 14 Clorf141 1 67557858 5.77E-24 25 IL23R 1 67652168 1.3E-5 5 PHTF1 1 114304453 0.00010 9 PHTF1 1 114356432 0.00016 13 AP4B1-AS1 1 114399256 7.99E-5 24 ARNGEP2 1 155916629 0.00048 4 CD48 1 1606950211 3.71E-6 13 ITLN1 1 160904255 5 5 ITLN1 1 160904254 4.62E-5 5 ITLN1 1 160904255 0.00019 7 IL001019 1<</td><td>Gene Chromo StartP OriginalP #SNP Rankin PER3 1 7844488 4.71E~5 19 4.71E~5 UTS2 1 7907671 2.13E~6 5 2.13E~6 DMRTA2 1 5083222 0.00013 1 0.00013 MIER1 1 67395925 7.32E~7 14 7.32E~7 Clorfl41 1 67537585 5.77E~24 25 5.77E~24 IL23R 1 67632168 1.36E~59 31 1.36E~59 PHTF1 1 114394433 0.00010 9 0.00016 PHTP1 1 143945432 0.00016 13 0.00016 PHTP12 1 155916629 0.00048 4 0.00048 CD48 1 16086329 3.02E~7 2 3.02E~7 LC1019 1 160902254 4.62E~5 4.62E~5 4.62E~5 LC1111 1 160906234 0.00020 5</td><td>Gene Chromo StartP OriginalP SMP Rankin Selet PER3 1 7844488 4.71E-5 19 4.71E-5 \checkmark UTS2 1 7907671 2.13E-6 5 2.13E-6 \checkmark IMER4 1 50683222 0.0013 1 0.00013 \checkmark MIER1 1 67395925 7.32E-7 14 7.32E-7 \checkmark IL23R 1 67657858 5.77E-24 25 5.77E-24 \checkmark PHTF1 1 114239823 6.68E-5 5 6.68E-5 \checkmark RESB1 1 114356423 0.00016 9 0.00016 \checkmark PTH72 114356426 0.00048 4 0.00048 \checkmark RESB1 1 14396256 7.99E-5 24 7.99E-5 \checkmark CD48 1 160650211 3.71E-6 13 3.71E-6 \checkmark ILDC1019 1 160902254 <t< td=""><td>Gene Chromo StartP OriginalP #SNP Rankin Select ISScore PER3 1 784488 4.718-5 19 4.718-5 9 9 UTS2 1 7907671 2.138-6 5 2.138-6 9 9 IMER42 1 5083222 0.00013 1 0.00013 9 9 MIER1 1 67395925 7.328-7 14 7.328-7 9 9 9 ILI23R 1 67557858 5.77E-24 25 5.77E-24 9</td><td>Gene Chrono StartP OriginalP SNP Rankin Select ISScore PER3 1 784488 4.71E-5 19 4.71E-5 ? ? UTS2 1 7907671 2.13E-6 5 2.13E-6 ? ? IMERIA2 1 50833222 0.0013 1 0.0013 ? ? MIERI 1 67557858 5.77E-24 5 5.77E-24 ? ? ILI23R 1 67652168 1.3E-59 31 1.3EE-59 ? ? MIERI 1 14239823 6.68E-5 5 6.68E-5 ? ? MERSH1 1 114356423 0.00016 9 0.00016 ? ? PHTP12 114356423 0.00016 13 0.00048 ? ? ? ARMGEP2 1 155916629 0.0044 0.00048 ? ? ? ILLC019 1 160</td></t<></td></t<>	Gene Chromo StartP OriginalP #SNP PER3 1 7844488 4.71E-5 19 UTS2 1 7907671 2.13E-6 5 DMRTA2 1 50883222 0.00013 1 MIER1 1 67395925 7.32E-7 14 Clorf141 1 67557858 5.77E-24 25 IL23R 1 67652168 1.3E-5 5 PHTF1 1 114304453 0.00010 9 PHTF1 1 114356432 0.00016 13 AP4B1-AS1 1 114399256 7.99E-5 24 ARNGEP2 1 155916629 0.00048 4 CD48 1 1606950211 3.71E-6 13 ITLN1 1 160904255 5 5 ITLN1 1 160904254 4.62E-5 5 ITLN1 1 160904255 0.00019 7 IL001019 1<	Gene Chromo StartP OriginalP #SNP Rankin PER3 1 7844488 4.71E~5 19 4.71E~5 UTS2 1 7907671 2.13E~6 5 2.13E~6 DMRTA2 1 5083222 0.00013 1 0.00013 MIER1 1 67395925 7.32E~7 14 7.32E~7 Clorfl41 1 67537585 5.77E~24 25 5.77E~24 IL23R 1 67632168 1.36E~59 31 1.36E~59 PHTF1 1 114394433 0.00010 9 0.00016 PHTP1 1 143945432 0.00016 13 0.00016 PHTP12 1 155916629 0.00048 4 0.00048 CD48 1 16086329 3.02E~7 2 3.02E~7 LC1019 1 160902254 4.62E~5 4.62E~5 4.62E~5 LC1111 1 160906234 0.00020 5	Gene Chromo StartP OriginalP SMP Rankin Selet PER3 1 7844488 4.71E-5 19 4.71E-5 \checkmark UTS2 1 7907671 2.13E-6 5 2.13E-6 \checkmark IMER4 1 50683222 0.0013 1 0.00013 \checkmark MIER1 1 67395925 7.32E-7 14 7.32E-7 \checkmark IL23R 1 67657858 5.77E-24 25 5.77E-24 \checkmark PHTF1 1 114239823 6.68E-5 5 6.68E-5 \checkmark RESB1 1 114356423 0.00016 9 0.00016 \checkmark PTH72 114356426 0.00048 4 0.00048 \checkmark RESB1 1 14396256 7.99E-5 24 7.99E-5 \checkmark CD48 1 160650211 3.71E-6 13 3.71E-6 \checkmark ILDC1019 1 160902254 <t< td=""><td>Gene Chromo StartP OriginalP #SNP Rankin Select ISScore PER3 1 784488 4.718-5 19 4.718-5 9 9 UTS2 1 7907671 2.138-6 5 2.138-6 9 9 IMER42 1 5083222 0.00013 1 0.00013 9 9 MIER1 1 67395925 7.328-7 14 7.328-7 9 9 9 ILI23R 1 67557858 5.77E-24 25 5.77E-24 9</td><td>Gene Chrono StartP OriginalP SNP Rankin Select ISScore PER3 1 784488 4.71E-5 19 4.71E-5 ? ? UTS2 1 7907671 2.13E-6 5 2.13E-6 ? ? IMERIA2 1 50833222 0.0013 1 0.0013 ? ? MIERI 1 67557858 5.77E-24 5 5.77E-24 ? ? ILI23R 1 67652168 1.3E-59 31 1.3EE-59 ? ? MIERI 1 14239823 6.68E-5 5 6.68E-5 ? ? MERSH1 1 114356423 0.00016 9 0.00016 ? ? PHTP12 114356423 0.00016 13 0.00048 ? ? ? ARMGEP2 1 155916629 0.0044 0.00048 ? ? ? ILLC019 1 160</td></t<>	Gene Chromo StartP OriginalP #SNP Rankin Select ISScore PER3 1 784488 4.718-5 19 4.718-5 9 9 UTS2 1 7907671 2.138-6 5 2.138-6 9 9 IMER42 1 5083222 0.00013 1 0.00013 9 9 MIER1 1 67395925 7.328-7 14 7.328-7 9 9 9 ILI23R 1 67557858 5.77E-24 25 5.77E-24 9	Gene Chrono StartP OriginalP SNP Rankin Select ISScore PER3 1 784488 4.71E-5 19 4.71E-5 ? ? UTS2 1 7907671 2.13E-6 5 2.13E-6 ? ? IMERIA2 1 50833222 0.0013 1 0.0013 ? ? MIERI 1 67557858 5.77E-24 5 5.77E-24 ? ? ILI23R 1 67652168 1.3E-59 31 1.3EE-59 ? ? MIERI 1 14239823 6.68E-5 5 6.68E-5 ? ? MERSH1 1 114356423 0.00016 9 0.00016 ? ? PHTP12 114356423 0.00016 13 0.00048 ? ? ? ARMGEP2 1 155916629 0.0044 0.00048 ? ? ? ILLC019 1 160

ie Resourc —		i 📥 😒 👘								
	er 📕 P	assedResultV	iewer × 🛅	TableViewe	r Window X	O View	Genes X 🔑 (Conditions	lGeneAssoc ×	•
CrohnDisease	Gene Symb	ols Selecte	d Regions	Sig. Genes			Gene As	sociation	Set: genome crohn	ecs
CrohnGWASresultHg19	Method:	Benjamini &	t Hochberg (1995)	~		P Value	Source:	META-P	~
Candi. Gene Files:	P P .				00000 h-					
candidategeneset_cr	Error Ka		Group	Jistance. 50	00000 op		Load	Genes	Update Ranks	Reset R
Genome: genome1						1				
Source : CrohnGWA	Group	Gene	Chromo	StartP	OriginalP	#SNP	Rankin	Select	ISScore	SelectA
Version : hg19	1	PER3	1	7844488	4.71E-5	19	4. 71E-5		1.00000	
Gene database :	1	UTS2	1	7907671	2.13E-6	5	2.13E-6		2.13E-6	Unselect
🖋 Used columns :	2	DMRTA2	1	50883222	0.00013	1	0.00013		0.00013	
Gene 5' extensio	3	MIER1	1	67395925	7.32E-7	14	7.32E-7		0.42004	Run
Gene 3' extensio	3	C1 or f141	1	67557858	5.77E-24	25	5.77E-24		0.00579	France
SMP LD : 1KG Hap]	3	IL23R	1	67632168	1.36E-59	31	1.36E-59		1.36E-59	Expor
Adjust P Value	4	PHTF1	1	114239823	6.68E-5	5	6.68E-5		6.68E-5	
Gana Scan' ganome or	4	RSBN1	1	114304453	0.00010	9	0.00010		0.20434	
Concert : genome_cit	4	PTPN22	1	114356432	0.00016	13	0.00016		1.00000	
Genome . genomel	4	AP4B1-AS1	1	114399256	7.99E-5	24	7.99E-5		0.22004	
Multivariate Tes	5	ARHGEF2	1	155916629	0.00048	4	0.00048		0.00178	
P value sources	5	CD48	1	160650211	3.71E-6	13	3.718-6		3.558-5	
Test : GATES	5	ITLNI	1	160846329	3.028-7	2	3.02E-7		3.02E-7	
Show : Detailed R	5	L0C1019	1	160902254	4.628-5	5	4.628-5		0.23747	
Gene Scan: genome_cro	0	ITLN2	1	160914815	0.00020	5	0.00020		0.85507	
Genome : genome1	0	5111	1	200862048	0.00021	7	0.00021		1.00000	
	7	KIR21R	1	200803948	1 108-5	11	1 198-5		1.108-5	
Multivariate Tea	1	RRP15	1	218458628	5 668-8	7	5 668-8		5.668-8	
 Multivariate Te: P value sources 	8	100 10	*	226332370	8 098-5	12	8 098-5		8.098-5	
Multivariate Tes P value sources Test : ECS	8	ACRD3	1		0.001 0	12	0.000 0	Ľ .	0.002 0	
Multivariate Ter P value sources Test : ECS Show - Detailed P	8 9	ACBD3	1	10589853	0.00027	2	0.00027		0.00027	
Multivariate Ter P value sources Test : ECS Show : Detailed R	8 9 10	ACBD3 LOC1019	2	10589853	0.00027	2	0.00027		0.00027	

Figure 5.2.3.2 The display after conditional associational analysis

5.2.4 Gene-pair-based association analysis by HYST

Click the menu <u>BioModule</u> \rightarrow <u>Gene-pair-based Association</u> to set the parameters as Figure 5.7.1. Note: the gene-pair based association analysis should use gene-based association analysis results by GATES as an input. A QQ plot of the gene-pair-based association p-value will be shown at the end of the analysis (Figure 5.7.2).



Figure 5.2.4.2 The display after running PPI-based association scan

Click the node "Show: Detailed Results" under 'Gene-pair scan', and you will get the graph of gene pairs. You can also export the results you want in this tab.



Figure 5.2.4.3 Function of displaying the results of PPI-based association scan

5.2.5 Multivariate gene-based association analysis by MGAS

The multivariate gene-based association analysis is different from the above knowledge-based secondary association analyses that are designed for multivariate analysis. Therefore, in the example dataset (available at http://grass.cgs.hku.hk/limx/kgg/download/KGGSample.zip), we prepared another real example from a published paper [Nat Genet. 2009 Jan;41(1):35-46] to demonstrate the the KGGSample\MultiPhenos folder. there analysis. In are two files. 9MetabolicPhenotypesPhg19.txt and 9MetabolicPhenotypesCorr.txt, contains which p-values and Pearson correlation of 9 quantitative metabolic traits respectively. Similarly, you should load the p-value file in to the KGG project at first.

Compare to the univariate analysis, there is one unique setting for the multivariate analysis when the analysis genome is built. The Pearson correlation should be specified by clicking "Set Correlation matrix of phenotypes for multivariate analysis only" in the "Build Analysis Genome" Dialog (Figure 5.2.5.1).

		genomeMe	etabo	1		P value	files:	9Metaboli	cPhenotype	sPhg19	.txt		~		
	Curre	ntChr	^	0 (0)				DAT	CRR			CI II	1001	1	THE
ct	curre	ntRSID		CurrentUnr	curren	010 0161	centros	DM1	0.295	-	0.028	GLU O. E144		101	105
or	Curre	ntPos		4	rs10000	023 9573	3906	0.8791	0. 335		0.8985	0. 7722	0.34	25	0.993
lue	BMI			4	rs10000	030 1033	74154	0.3644	0.283	5	0.4969	0.7887	0.96	69	0.7353
mn(s)	DEP			2	rs10000	07 2377	52054	0.3698	0.947	6	0.4607	0.4485	0.44	80	0.2814
l″ for	GLU			4	rs10000	092 2189	5517	0.371	0.235	7	0.05847	0.4337	0.51	07	0.7881
pie ction	HDL		~	<	ive the first	121 115/5	/4035	in neger	11 1211		11 XX24	11 5445	10.05	~	>
tondad		nori en 1		L 5	lik at 5	/ . E	kh et	- 2° 🔽	Set Corre	lation	matrix of	nhanatunas	for mul	tivaria	te englycic enly
tenueu	gene i	m2 (0	o oos	n 5	Ko at 5		KO at		ile Settir		matrix or	phenotypes	101 841	LIVALIC	are analysis only
nore Lo	The c	12 \ 0	. 003						The better	*5					CurrentPos V
ljust b		Set Corre	atio	n matrix or pri	enotypes										1_−10
ne defi	Select MORE	t ONE or p-value		Correlation	Matirx f	file D:\KG	G \ KGGS ar	mple\Mult:	iPhenos\9M	etaboli	cPhenotype	esCorr.txt			1 on. ng19
Data	colum	n(s)													
Jata	Ctrl	for for												1.0-	t
plotype	selec	tion			BMI	CRP	DBP	GLU	HDL.	INS	LDL	SBP	TG		
Genome	curre	entRSID	^	5		0.22	0.3	0.18	0.31	0.52	0.23	0.3	0.32		
- chome	Curre	entPos		ĉ	0.22		0.06	0.04	0.09	0.15	٥	0.07	0.09		
	BMI			8	0.3	0.06		0.13	0.07	0.3	0.14	0.6	0.23		ne association fi
	DEB			EL	0.18	0.04	0.13	1.1	0.05	0.31	0.02	0.18	0.1	0.0	
	GLU			보	0.31	0.09	0.07	0.05		0.27	0.14	0.01	0.33		
	HDL			£	0.52	0.15	0.3	0.31	0.27	1	0.19	0.26	0.44		
	INS			Ę	0.23	0	0.14	0.02	0.14	0.19		0.09	0.26		> Build
	LDL			â	0.3	0.07	0.6	0.18	0.01	0.26	0.09		0.18		
	SBP			2	0.32	0.09	0.23	0.1	0.33	0.44	0.26	0.18	1.		
	TG		~												
						Sav	e						Canc	el	Cancel
												e			
ck tl plo	he n ot w:	nenu ill be	Fig 1 <u>G</u> e s	<u>gure 5.2</u> h <u>ene</u> → hown in	<u>Mul</u> n the	Spe tivaria Runn	ate A ingR	the co Associ Result	orrelat ation Viewe	tion to ser (F	of phe et the Figure	param 5.9.3)	es an eters . The	alysi as F deta	is ïgure 5.9.2. iiled gene-b
ck tl plo alue	he n ot wi	nenu ill be an be	Fig 1 <u>G</u> e si e v:	<u>gure 5.2</u> i <u>ene</u> → hown in iewed b	<u>Mult</u> n the by cli	Spe tivaria Runn cking	cify and the state of the state	<i>the co</i> <u>Associ</u> Result node	orrelat ation Viewe "Shov	tion to ser er (H v De	<i>of phe</i> et the Figure etailed	param 5.9.3) Resul	eters . The t" (F	alysi as F deta igure	is Tigure 5.9.2. ailed gene-b e 5.9.4).
ck tl plo alue _{Gene-1}	he n ot w: es ca	nenu ill be an be	Fig 1 <u>G</u> e si e v	$\frac{\text{dene}}{\text{hown in}} \rightarrow \frac{1}{1000}$	<u>Mult</u> n the by cli	Spe tivaria Runn cking	cify and the state of the state	the co Associ Result node	orrelation ation Viewe "Shov	tion to se er (F w De	of phe et the Figure etailed	enotypo param 5.9.3) l Resul	eters The t" (F	alysi as F deta igure	is ïgure 5.9.2. niled gene-b e 5.9.4).
ck th plo alue Gene-1	he n ot w: es ca based	nenu ill be an be associa	Fig $1 \frac{G}{G}$ $E = S$ $E = V$ 1tion	$\frac{1}{1}$ $\frac{1}$	<u>Mult</u> n the by cli	Spe tivaria Runn cking	cify ate <u>A</u> ingR the	<u>Associ</u> Associ Result node	orrelation ation Viewe "Show	tion to ser (F w De	<i>of phe</i> et the Figure etailed	param 5.9.3) l Resul	eters . The t" (F	alysi as F deta igure	is Tigure 5.9.2. ailed gene-b e 5.9.4).
ck th plo alue Gene-I	he m ot w: es ca based	nenu ill be an be associa	Fig G G G G G G G G	gure 5.2 $\frac{1}{2} \rightarrow \frac{1}{2}$ hown in iewed b	<u>Mult</u> n the by cli	Spe tivaria Runn cking	cify and the state of the state	Associ Associ Result node	ation Viewe "Shov plot disp s with p-	tion to ser (F w De	of phe et the Figure etailed	param 5.9.3) l Resul	eters . The t" (F	alysi as F deta igure	is 'igure 5.9.2. ailed gene-b e 5.9.4).
ck th plo alue Gene-I	he m ot w: cs ca based gene	menu ill be an be associa	Fis Fis $C = S$ Fis Fi	gure 5.2 $\frac{1}{2} \rightarrow \frac{1}{2}$ hown in iewed b 1 scan $\frac{1}{2}$	<u>Mul</u> n the by cli	Spe tivaria Runn cking	cify ate <u>A</u> ingR the :	Associ Result node	ation Viewe "Shov plot disp s with p-	tion to se er (F w De	of phe et the Figure etailed	param 5.9.3) l Resul	eters The t" (F)	alysi as F deta igure	is 'igure 5.9.2. ailed gene-b e 5.9.4).
ck th plo alue Gene-I a Name alysis	he m ot w es ca based based gene s Geno Set	menu ill be an be associa escanMu omes G genomeM	Fiş I <u>G</u> eS eV atior Iti	gure 5.2 <u>sene</u> → hown in iewed b iscan Groups bol	<u>Mult</u> n the by cli	Spe tivaria Runn cking	the	Associ Result node	ation Viewe "Shov plot disp s with p-v	tion to ser (F w De play values alues	of phe et the Figure etailed	param 5.9.3) l Resul	eters The t" (F	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
ck th plo alue Gene-l a Namee	he n bt W es ca based based gene s Geno Set	menu ill be an be associa escanMu omes G genomeN me	Fiş I <u>G</u> esi ev atior Itil	$\frac{\text{gure 5.2}}{\text{hown in}}$ $\frac{\text{hown in}}{\text{iewed b}}$ $\frac{\text{Groups}}{\text{bol}}$	<u>Mult</u> n the by cli	Spe tivaria Runn cking	ate A ingR the	Associ Result node	ation Viewe "Shov plot disp s with p-v value 1E-	tion to seer (F v De olay values alues	of phe et the Figure etailed s <= 1E- <= 5E- Manhantt	param 5.9.3) I Resul	eters The t" (F dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
ck th plo alue Gene-I a Name alysis enome P Valu	he n bt W based based gene s Geno Set ue Nam	menu ill be an be associa escanMu omes G genomeN	$Fi \begin{cases} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ $	$\frac{\text{gure 5.2}}{\text{hown in}}$ $\frac{\text{hown in}}{\text{iewed b}}$ $\frac{\text{Groups}}{\text{bol}}$	<u>Mult</u> n the by cli	Spe tivaria Runn cking	cify in the state of the state	Associ Result node	plot disp s with p-v value 1E-	tion to seer (F v De olay values alues	of phe et the Figure etailed s <= 1E- <	param 5.9.3) l Resul	eters The t" (F dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
ck th plo alue Gene-J a Name enome P Valu MI IRP	he n bt Wi es ca based based gene s Geno Set ue Nam	menu ill be an be associa escanMu omes G genomeN ne	Fig I <u>G</u> e S e V ation Itti Heta	$\frac{\text{gure 5.2}}{\text{hown in}}$ $\frac{\text{hown in}}{\text{iewed b}}$ $\frac{\text{Groups}}{\text{bol}}$	<u>Mult</u> n the by cli	Spe tivaria Runn cking	the s	A generation of the contract o	plot disp s with p-v value 1E- isplay	tion to seer (F v De olay values alues	of phe et the Figure etailed s <= 1E- <	param 5.9.3) l Resul	eters The t" (F dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
ck th plo alue Gene-I a Name alysis enome P Valu MI RP EP	he n bt Wi cs ca based based based c gene s Geno Set 1 ue Nam	menu ill be an be associa escanMu omes G genomeW ne	Fig I <u>G</u> e S e V atior Itti	$\frac{\text{dene}}{\text{hown in}} \rightarrow \frac{1}{\text{hown in}}$	<u>Mul</u> n the by cli	Spe tivaria Runn cking	the s	A state of the contract of the	ation Viewe "Show plot disp s with p-v value [IE- isplay UPs insid	tion to see (F v De slay values alues	of phe et the Figure etailed s <= 1E- (= 5E- Manhantti	param 5.9.3) l Resul	eters The t" (F dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
ck th plo alue Gene-l alysis enome P Valu MI RP BP HU	he m bt W S C C C C C C C C C C C C C C C C C C C	menu ill be an be associa escanMu omes G genomeW ne	Fig I <u>G</u> e S e V atior lti lti	$\frac{\text{dene}}{\text{hown in}} \rightarrow \frac{1}{\text{hown in}}$	Mult n the by cli	Spe tivaria Runn cking	the s	Associ Associ Result node	plot disp s with p-v value [IE- isplay UPs insid	tion to seer (F v De solay values alues to le gen	of phe et the Figure etailed s <= 1E- <	enotype param 5.9.3) l Resul	eters The t" (F) dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
Ck th plo alue Gene- a Name alysis enome P Valu BP LU DL	he m ot w: cs cc based based gene s Geno Set ue Nam	menu ill be an be associa escanMu genomeN me	Fig I <u>G</u> e S e V atior Iti	$\frac{\text{dene}}{\text{hown in}} \rightarrow \frac{1}{1000}$	Mult n the by cli	Spe tivaria Runn cking	the	Associ Associ Result node	plot disp s with p-v value [IE- isplay MPs insid	tion to seer (F v De slay values alues alues to le genuite	of phe et the Figure etailed $s \leftarrow IE-$ Menhentte es \square Wid es \square Hei	param 5.9.3) l Resul 6 Ti 8 He an plot Sl dth 600 i ght 400	eters The t" (F dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
Ck th plo alue Gene- a Name alysis enome P Valu BP LU DL	he m bt w based based gene s Geno Set i ue Nam	menu ill be an be associa escanMu genomeN me	Fig I <u>G</u> e S e V atior lti lti	$\frac{\text{dene}}{\text{hown in}} \rightarrow \frac{1}{1000}$	Mult n the by cli	Spe tivaria Runn cking	the s	Associ Associ Result node	ation Viewe "Show plot disp s with p- value [IE- isplay IIPs insid IIP outsid -value [tion to see er (F v De slay values alues alues to le genuite te=-10	of phe et the Figure etailed $s \leftarrow IE-$ Manhantte es \square Wid es \square Hei	param 5.9.3) l Resul 6 Ti 8 He an plot Sl dth 600 i ght 400	eters The t" (F dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
Ck th plo alue Gene-J a Name alysis enome P Valu MI RP P LU DL	he m bt W based based ; gene s Geno Set 1 ue Nam	nenu ill be an be associa escanMu omes G genome ne	Fig I <u>G</u> e S e V atior lti lti lti me l	$\frac{\text{dene}}{\text{hown in}} \rightarrow \frac{1}{\text{hown in}}$	Mulin the by cli	Spe tivaria Runn cking	n matrix	Associ Result node	ation Viewe "Show "Show "ishow value 1E- isplay WPs insid WP outsid -value [tion to see (F v De slay values alues alues le genuite te genuite te genuite	of phe et the Figure etailed s <= 1E- < <= 5E- Manhantti es 🗹 Wid es 🗹 Hei	enotype param 5.9.3) l Resul	eters The t" (F dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
Ck th plo alue Gene-J alysis enome P Valu MI RP P LU DL his ar	he m ot w based based gene s Geno Set ue Nam	nenu ill be an be associa escanMu omes G genome ne	Fig I <u>G</u> E S E V atior Itil Hene Meta	$\frac{\text{gure 5.2}}{\text{hown in}}$ $\frac{\text{hown in}}{\text{iewed b}}$ $\frac{\text{groups}}{\text{bol}}$ $\frac{\text{Sele}}{\text{HAS a pheno}}$	Mulin the by cli	Spe tivaria Runn cking	n matri	Associ Result node	ation Viewe "Show "Show "ishow value 1E- isplay WPs insid WP outsid -value [tion to see (F v De alues alues le genuite tE-10	of phe et the Figure etailed s <= 1E- < <= 5E- Manhantta es 🗹 Wid es 🗹 Hei	enotype param 5.9.3) l Resul	eters The t" (F dth 12 ight 50 Ps outs	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
Ck th plo alue Gene-J a Name alysis enome P Valu BP LU DL his ar thods	he m bt w based based gene s Geno Set 1 ue Nam	menu ill be an be associa escanMu omes G genomeN ne	Fig I <u>G</u> E S. E V ation Itil Hene Meta	gure 5.2 $\frac{1}{2}$ \rightarrow hown in iewed b 1 scan $\frac{1}{2}$ $\frac{1}{2}$ \frac	2.5.7 <u>Mult</u> n the by cli	Spe tivaria Runn cking	n matri	Associ Result node	ation Viewe "Show "Show "ishow value 1E- isplay WPs insid WP outsid -value [tion to see (F v De alues alues alues le genuite-10	of phe et the Figure etailed	enotype param 5.9.3) l Resul 6 Ti 8 He an plot Sl dth 600 i ght 400	eters The t" (F dth 12 ight 50 Ps outs	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
Ck th plo alue Gene-J a Namee alysis enome P Valu MI RP LU DL his ar thods Multix	he m bt W es ca based • gene s Geno Set [ue Nam nalysi	menu ill be an be associa escanMu omes G genomeM ne is geno te Exte	Fig I <u>G</u> E S. E V. ation Itil Hene Meta	gure 5.2 $iene \rightarrow$ hown in iewed b scan Groups bol Selea tAS a pheno d Simes tes	Mulin the by cli	Spe tivaria Runn cking	cify in the state of the state	Associ Result node	orrelation <u>ation</u> Viewe "Show plot disp s with p-v value 1E- isplay NPs insid NP outsid -value [2]	tion to see er (F v De slay values alues alues to le genu le genu le genu	of phe et the Figure etailed s <= 1E- < <= 5E- Manhantti es ☑ Wid	enotype param 5.9.3) l Resul 6 Ti 8 He an plot Sl dth 600 i ght 400	eters The t" (F dth 12 ight 50	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
Ck th plo alue Gene- a Name alysis enome P Valu BP LU DL his ar thods Multix	he m bt w based based based c gene s Geno Set [ue Nam nalysi	menu ill be an be associa escanMu omes G genomeN ne is geno te Exte	Fig I <u>G</u> S S S V ation	$\frac{\text{gure 5.2}}{\text{hown in}}$ $\frac{\text{hown in}}{\text{iewed b}}$ $\frac{\text{scan}}{\text{Groups}}$ $\frac{\text{bol}}{\text{fAS a pheno}}$ $\frac{\text{d Simes tes}}{\text{sotida pol}}$	2.5.7 <u>Mult</u> n the by cli t t t t t mornhi	Spe tivaria Runn cking	n matri	Associ Result node	ation Viewe "Show "Show "show "th p-v value 1E- isplay "Ps insid "Poutsid -value [tion to see (F v De alues alues alues to III IE-10	of phe et the Figure etailed s <= 1E- < <= 5E- Manhantti es 🗹 Wid es 🖉 Hei	enotype param 5.9.3) l Resul c Trian an plot Sl dth 600 i ght 400	eters The t'' (F dth 12 ight 50 Pr outs	alysi as F deta igure	is igure 5.9.2. ailed gene-b e 5.9.4).
Ck th plo alue Gene-J a Name alysis enome P Valu MI RP BP LU DL his ar thods Multiv [] Igg	he m bt W es ca based ; gene s Geno Set [ue Nam nalysi variat	menu ill be an be associa escanMu omes G genomeM ne is geno te Exte	Fig I <u>G</u> S S S V atior atior atior atior atior atior atior atior atior atior atior atior atior atior atior ation	gure 5.2 $iene \rightarrow$ hown in iewed b scan Groups bol Selection fAS a pheno d Simes tess leotide pol	2.5.7 <u>Mulin</u> n the by cli et tupe co t (MGAS ymorphi	Spe tivaria Runn cking	cify in the solution of the so	hout lind	orrelation <u>ation</u> Viewe "Show plot disp s with p-v value 1E- isplay MPs insid MP outsid -value [2 kage dise	tion to see er (F v De slay values alues alues to le genui le genui le genui	of phe et the Figure etailed s <= 1E- 	enotype param 5.9.3) l Resul 6 Ti 8 He an plot Sl dth 600 i ght 400	eters The t" (F dth 12 ight 50 IPs outs	alysi as F deta igure	is igure 5.9.2. niled gene-b e 5.9.4).
Ck th plo alue Gene-I a Name alysis enome P Valu MI RP BP P Valu MI thods Igr thods Igr Refet I. I gr thods S alue S alu	he m bt Wi es ca based based based based based based s Geno Set us lan lan lysi 'an do for 'for	nenu ill be an be associa escenMu omes G genomeN me is geno te Exte single , Gui ng ext er Slu multi	Fig Fig e S e V ation at	gure 5.2 iene_→ hown in iewed b scan Groups bol Selea HAS a pheno d Simes tes leotide pol Kwan JS, led Simes S, Dolan iate gene	2.5.1 <u>Mulin</u> the by cli et type co t (MGAS ymorphi Sham proced CV, Li -based	Spe tivaria Runn cking	n matri	Associ Associ Result node Ps Ps Ps six! I p d d s Si six! I p d d s si si si si si si si si si si si si s	ation Viewe "Show "Show with p- value IE- isplay mPs insid NP outsid -value [and pow : 2011 Posthu ation a	tion to so er (F v Do olay values alues 10 le genuitation le genui	of phe et the Figure etailed s <= 1E- <= 5E- Manhenttu es 🛛 Wid es 🖾 Hei brium (LD .1;88(3)) Li M. 1 is.	enotype param 5.9.3) l Resul dth 600 ight 400	eters The t" (F dth 12 ight 50 Ps outs	alysi as F deta igure	is igure 5.9.2. niled gene-b e 5.9.4).



5.2.6 Estimating driver-tissues by selective expression of genes associated with complex diseases or traits

This function is designed to estimate driver tissues by tissue-selective expression of phenotype-associated genes in GWAS. The tissues in which causal or susceptibility genes initiate the phenotypes are called driver or causal tissues.

For this function, KGG requires two types of inputs data, gene expression values of multiple tissues and GWAS summary statistics or association p-values at variants for a tested disease. The expression values at genes and transcripts or even exons can be used for the estimation. The GWAS p-values are used to detect susceptibility genes by a conditional gene-based association test (See 5.2.3).

Click the menu <u>Gene</u> \rightarrow <u>Driver Tissue (DESE)</u> to set the parameters and input specify an expression file, as Figure 5.2.6.1. The estimating driver-tissues analysis should use gene-based association analysis results by ECS, and input an expression file which contains the expression values of each genes in every tissue. Please download the expression file from the address: <u>http://grass.cgs.hku.hk/limx/rez/</u>. Then, click the button <u>Lord Genes</u> to load significant phenotype-associated genes according the threshold of multiple testing (Figure 5.2.6.1). Next, click the button <u>Run</u> to estimate driver-tissues. It will take you two or more hours.

The estimated driver-tissues will be prioritized according to their statistical significance. Four types selective-expression measures (robust-regression z-score, conventional z-score, MAD robust z-score, and ratio of vector-scalar projection) are used in the estimation analysis. A combined prioritization is generated by averaging the -log10(p) based on the four measures (as shown in Figure 5.2.6.2 and Figure 5.2.6.3).

ConsortEctation Consort Project Window Resource Window Consort C	Gene Associatio P Value Source: Method: Benjar Error Rate: 0.00	(DERSE) × n Set: genes_s pval sini & Hochberg 5 Group I	cz_ECS (1995) Distance: 5000000	v Load	▼ Clear d Genes	Expression Expression fr ects\Specifi Filter out er	ile See format cExpression\KGG\g cpression less the	Download tex.transcript. n 0.01 Unit	txt\getx.transcr as defined	ipt.nean.txt
Gene 5' extension : 5.0 Gene 3' extension : 5.0 SBF LD : 105 Halotynes WF Adjust P Vile : 10 Gene Senz genes gez ZGS Geneme : genomei Waltivariste Test : 10 F V subse sources : [gval] Test : ECS Show : Detailed Results	Genes IIssues	Gene	Chromosome	StartPosition.	Original?	#SHP	RankingScore	Salact	Cond P	SelectAll UnselectAll Run Export Remove

ct Window Resource Window 🖃	Driver Tissue (DERSE)	<								
Gscz	Gene Association Set:	anas ser FCS			Expres	sion				
P-value Files: PGC2_SCZ8_summary_txt	B W la Communication			1	Expre	ssion file See f	Eormat Downly	bad		
enome: genome1	r value Source. [pval				ects	SpecificExpression	\KGG\gtex.transcr	ipt.txt\getx.	transcrip	ot. mean. txt
Source : PGC2_SCZ8_summary.txt Version : bs19	Method: Benjamini & Ho	chberg (1995)		Load Genes	Filte	r out expression le	ss than 0.01 Un	nit as defined	1	
Gene database : RefGene	Error Rate: 0.05	Group Distance	: 5000000 bp							
Used columns : [pval]	Genes Tissues									_
Gene 3' extension : 5.0	Group Gen	e Chror	nosome StartPosit	on OriginalP	#SITF	RankingScor	e Select	Cond. P		SelectAll
SNP LD : 1KG Haplotypes VCF Adjust P Value : No	1 PRKCZ 1 PRKCZ-AS:	1	2036154	0.00016	121	3.81	<u>v</u>	?		UnselectAl
Gene Scan: genes_scz_ECS	1 FAAP20	1	2120987	0.00139	5	2.86	V	?		Run
Genome : genomel	1 INPRSF14 1 L0C100996	1 583 1	antornation			2.94 3.09	V V	?	-11	Export
P value sources : [pval]	1 MMEL1	1	Genes have 1	een loaded!		2.95	<u>v</u>	?		Romovo
Test : ECS	1 MIR4417	1	1 3			2.61		?		Itemove
Show . Detailed Results	1 TAS1R1 1 ZBTB48	1		10m		2.70		?		
	1 KLHL21	1	6650783	1.81E-5	24	4.74	V V	?		
	1 LINCO1672	1	6787586	0.00142	10	2.85		?		
	1 SLC45A1 1 RERE	1	8378144 8412463	1.47E-5 4.30E-10	28 439	4.83	V V	?		
	1 L0C102724	552 1	8484704	2.96E-8	21	7.53		?		
	2 CROCC	1	17248444	0.00295	24	2.53	V V	?		
	2 MFAP2 2 PADT1	1	17300998	0.00264	9	2.58	<u>v</u>	?		
	2 IGSF21	1	18434239	0.00185	575	2.73		?		
	2 CDC42 2 PIGV	1	22379119 27114678	0.00379	7	2.42	N N	?		
	2 MED18	1	28655512	0.00402	18	2.40	<u>v</u>	?		
	2 TRHAU1AP	1	28879528	0.00129	25	2.89		?		
	2 SING12 2 SNORD99	1	28905049 28905254	0.00132	3	2.88		?	-11	
(V4) - A systematic biological t Data Gene BioModule Power I 112/7440148 👩 🌾 🕽 😵 😡	Figure 5	.2.6.2	The disp	play afte	er loa	uding gel	nes	_		ſ
(V4) - A systematic biological t Data Gene BioNodule Power I D12/7401101 💦 📽 🗅 😵 👀 ctWindow Resource Window 🖛	Figure 5	ng system (The disp	olay afte	er loa	uding gen	nes			-
(V4) - A systematic biological t Data Gene BioModule Power] n2//(40100) 🕞 <table-cell> 🗎 😵 🔊 tWindow Resource Window Gesc</table-cell>	Figure 5	.2.6.2	The disp	play afte	er loa es (Tuly :	L. 2019)	nes			
(V4) - A systematic biological : Data Gene BioMedule Power] 12///40113 3 2 2 3 3 3 3 tWindow Resource Window E Gacz P-value Files: PGC2_SCZ8_summary.txt	Figure 5	ng system (RSE) × et: genes_sc val	The disp for Genome-wide	olay aft Genetic studi	er loa es (July : 	J Expression Expression	nes	rmat Dov	vnload	- -
(Y4) - A systematic biological Data Gane BioModule Power I 12/7401118 C C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 ng system (RSE) × iet: [genes_sc val i & Hochberg	The disp. For Genome-wide	olay aft Genetic studi	er loc es (July : Clear d Genes	Lang gen , 2019) Expression Expression Expression	nes file See fo s\SpecificExpre	rmat Doy	<u>vnload</u> ex. trans	cript.tx
(V4) - A systematic biological : Data Gane BioModule Power] 12//(0101) C C C C C C tWindow Resource Window C Gacz P-value Files: Poc2_SC28_summary.txt Genome: genome! Source : Poc2_SC28_summary.txt Version : hg19 Gane C Labora : Apfine	Figure 5	ng system : RSE) × i & Hochberg Group D	The disp. for Genome wide re_ECS (1995) istance:	Genetic studi	er loa es (Tuly :	L 2019) Expression Expression Filter out	nes file <u>See fo</u> s\SpecificExpre expression less	rmat Dov ssion\KGG\gt t than 0.01	vnload ex. trans Vnit as	cript.tx
(V1) - A systematic biological : Data Gene BioModule Power [12//(UIII) C C C C C C C C C C C C C C C C C C	Figure 5	RSE) × i & Hochberg Group D	The disp. For Genome wide re_BCS (1995) setunce: 5000000	Genetic studi	er loa es (Tuly)	L 2019) Expression Expression Filter out	nes file <u>See fo</u> s\SpecificExpre expression less	rmat Dov ssion\KGG\gt i than 0.01	vnload ex. trans	cript.tx
(V1) - A systematic biological : Data Gene BioModule Power [12//(CUIN) C C C C C C C C C C C C C C C C C C C	Figure 5	RSE) × i & Hochberg Group D	The disp. For Genome wide re_ECS (1995) istance: 5000000	Genetic studi	er loca es (July : Clear d Genes	L 2019) Expression Expression Filter out Retigner	nes	rmat Doy ssion/KGC/gt : than [0.01	vnload ex. trans Unit as	cript. tr
(V4) - A systematic biological : Data Gene BioModule Power] 12//(40100) C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 .ng system 	The disp. For Genome wide :z_ECS (1995) istance: 5000000 'essionZ Convert 4.22-20	Genetic studi	es (July : Clear d d Genes MADZ :-16	Lapression Expression Expression Filter out Ratio0ffreoj 1.22-25	nes file See fo >SpecificExpre expression less iection Aver 19.9665	rmat Doy ssion\XGG\gt t than 0.01 agedLog(p) 40983412553	vnload ex. trans Unit as	cript.tx cript.tx s defined SelectAll UnselectAll
(V4) - A systematic biological : Data Gene BioModule Power I ://(UID) C C C C C ://(UID) C C C C C ://(UID) C C C C C ://(UID) C C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 	The disp. For Genome-wide sz_ECS (1995) istance: 5000000 ession2 Convert 4.22-20 6.22-18 5.22-18	Genetic studi	er loca es (Taly) Clear d d Genes MADZ	Lang gen Lang gen Expression Expression Filter out Ratio0fFreg 1.22-25 3.32-24 4.14-21	file See fo s\SpecificExpre expression less is critical and aver 19.9005 18.8700 19.7700	rmat Doy sion\XGG\gt t than 0.01 agedLog(p) 40983412553 40150248355	vnload ex. trans Unit as	cript.tx
 (V4) - A systematic biological Data Gene BioHodule Power I 12/740118 C - C - C - C - C - C - C - C - C - C	Figure 5	.2.6.2 	The disp. For Genome-wide re_ECS (1995) istance: 5000000 #25-20 6.22-18 1.62-16 1.12-16	Genetic studi	er loca es (Taly) (Lear d Genes MADZ HADZ HADZ HADZ HADZ HADZ HADZ HADZ	Expression Expression Expression E:\Projecti Filter out Ratio0fProj 1.22-25 3.32-24 4.1E-21 1.38-21	file See fo s)SpecificExpre expression less is Gree 19.9665 18.8766 17.2722 17.0722 17.0722	rmat Dow sion\XGG\gt. than 0.01 40983412553 40150240355 8448202584 0851300817	Vnload ex. trans Vnit as	eript.tx
V4) - A systematic biological : Data Gene BioNodule Power I t2//c0110 C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 ng system RSE) × set: genes_so val i & Mochberg Group D Group D 1.52-21 1.52	The disp. For Genome-wide cc_ECS (1995) istance: 5000000 6.22-10 1.02-16 1.12-16	Genetic studi	er loca es (Tuly : Clear d d Genes MADZ -16 -15 -17	Expression Expression Expression Filter out Ratio0fFroj 1.22-25 3.32-24 4.18-21 2.55-13	file See fo s\SpecificExpre expression less isection Aver 19.9665 18.8766 17.2722 17.0388 16.4852	rmat Doy ssion/XGG/gt/ ssion/XGG/gt/ ssion/XGG/gt/ ssion/XGG/gt/ ssion/XGG/ ssion/XGG/ ssion/XGG/ ssion/XGG/ ssion/XGG/ ssion/XGG/ ssion/XGG/ ssion/XGG/ gt/ ssion/XGG/gt/	vnload ex. trans Unit ar	eript.tx
(V1) - A systematic biological : Data Gene BioModule Power I 12///C01101 (a) (b) (b) (b) (b) (b) (b) (b) (b) (b) (b	Figure 5	.2.6.2 Ag system Ag	The disp. For Genome-wide cc_ECS (1995) istance: 5000000 6.22-10 6.62-10 1.12-16 6.07-16 3.82-16 7.98-17	Genetic studi Genetic studi y Load bp ttionalZ 5.55 5.52	er loca es (Tuly) clear d Genes MADZ 	Expression Expression E:\Project: Filter out RatioOfFroj 1.2E-25 3.3E-24 4.1E-21 1.3E-21 1.3E-21 1.3E-21 1.5E-13 3.0E-20 1.1E-15	file See for s\SpecificExpre expression iscion Aver 19.9665 18.8766 18.722 17.0388 16.4858 16.4858 16.4964 16.4988	rmat Doy ssion X60° gt. t than 0.01 40983412553 40150240355 46150240355 46150240355 46150240355 46150240355	Vnload ex. trans Unit ar	cript.tx
V1) - A systematic biological Data Gene BioModule Power I 12//(ULUS) VWindow Resource Window VWindow Resource Window Cenze P-value Files: M PGC2_SC28_summary.txt Gences: Powrsion: hg19 Gene K: Surce : POC2_SC28_summary.txt Genes Kathesse : Reforme Used columns : [pval] Gene S'extension : 5.0 Gene Ser: SBP LD : 1KG Haplotypes WCF Adjust P Value : Ho Genese : genese: Maltiveriste Test : Ho Pervine sources : [pval] Test : ECS	Figure 5	.2.6.2 ing system ing system	The disp. For Genome wide reECS (1996) istance: 5000000 4.22-20 6.22-18 1.6	Genetic studi Genetic studi v Load bp tionalZ 5.33 2.22 3.44 6.43 1.22 5.55 5.22 5.52 5.52 5.52 5.52 5.52	er loca es (Tuly) clear d Genes MADZ t-16 t-15 t-15 t-15 t-16 t-15 t-16	Image: state	file See for s\SpecificExpre expression isetion Aver 19.965 18.8766 17.7328 16.4864 16.2200 16.2200 16.4464 16.2200	rmat Doy ssion\KGG\gt t than 0.01 agedLog(p) 40963412553 44150248355 44150248355 44150248355 44150248355 445820584 4959170407 4954741902 434545540214	vnload ex. trans	cript.tx cs defined SelectAll UnselectAll Run Export Remove
 V4) - A systematic biological Data Gene BioModule Power [22/7C0110] Window Resource Window E- Secz P-value Files: M FGC2_SC28_summary.txt Gene Genomal Seurce : FGC2_SC28_summary.txt Version : hg19 Gene S' extension : 5.0 Gene S' extension : 5.0 SHP LD : HG Hapletypes VCF Adjust P Value : Ho Gene Scan: gene_scz_ECS Geneme : genomal Nultiveriate Test : Ho P value sources : [pval] Test : ECS Show : Detailed Results 	Figure 5	.2.6.2 .1.g system 	The disp. for Genome-wide cr_ECS (1995) istance: 5000000 ression2 Conver 4.22-20 6.22-18 1.02-16 1.12-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 1.02-16 5.382-18 5	Genetic studi Genetic studi v Loav bp tionalZ 5.33 2.22 3.44 6.33 1.25 5.52 2.355 5.52 2.355 6.11	er loca es (Taly) clear d d Genes MADZ HIG HIG HIG HIG HIG HIG HIG HIG HIG HIG	Laboration (Constraint) Expression Expression Expression Filter out Filter out Ratioloffred 1.22-25 3.38-24 4.18-21 2.58-13 3.59-21 2.59-13 5.22-16 5.22-16 5.22-16	file See for file See for s\SpecificExpre expression extion Aver 18.8766 18.8766 17.722 17.0388 16.4852 16.4864 16.2290 15.2906 15.2906 15.2121	rmat Dov sion\X66\gt than 0.01 40983412553 44150248355 44150248355 44150248355 44150248355 448020584 40051300817 169317407 493248949 948476031627	vnload ex. trans	cript.tx
 Y4) - A systematic biological Data Gene BioModule Power I (2/740110) C C C C C C C C C C C C C C C C C C C	Figure 5 Knowledge-based wins A (1) (2) (2) (2) (2) (2) (2) (2) (2) (2) (2	.2.6.2 	The disp. For Genome-wide reg_ECS (1995) istance: 5000000 6.22-18 1.62-16 6.22-18 1.62-16 6.302-16 6.302-16 6.302-16 6.302-17 4.32-14 4.32-14 4.35-13 4.62-14 4.35-13 4.62-14	Constic studi Genetic studi	er loca es (Taly) Clear d d Genes MADZ 1-16 1-16 1-16 1-16 1-15 1-16 1-16 1-16	Expression Expression Expression Expression Filter out Ratio0fProj 1.28-25 3.38-24 4.18-21 1.38-21 2.58-13 3.04-20 1.18-15 5.22-16 3.22-11 8.22-15 1.38-11	file See fo file See fo s\SpecificExpre expression less iection Aver 19.9665 18.8766 17.7222 17.0382 16.4624 16.0240 15.2906 15.215 15.1651 15.451	rmat Doy ssion/KGC/gt than 0.01 40983412553 46150248355 8448020584 0851300817 14931741907 34545540214 4932248949 084741907 25095330377 81213452770	vnload ex. trans	eript.tx
 V4) - A systematic biological Data Gene BioModule Power I (2/74010) C C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 Ag system RSE) × set: genes_sc val i & Nochberg Group D 0.37-18 7.52-12 0.52-12 1.62-18 1.72-18 3.72-18	The disp. For Genome wide cr_ECS (1995) istance: 5000000 6.22-16 0.62-16 0.62-16 1.12-16 0.768-17 4.38-14 4.38-13 4.65-14 3.55-13 8.72-13	Conctic studi Conctic studi Concti	er loca es (Tuly) es (Tuly) Clear d d Genes MADZ 1-16 1-16 1-16 1-16 1-16 1-16 1-16 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18 1-16 1-18	Expression Expression Expression E:\Projecti Filter out RatioOfFroj 1.2E-25 3.3E-24 4.1E-21 1.3E-21 2.5E-13 3.0E-20 1.1E-15 5.2E-16 3.2E-11 2.5E-13 3.0E-20 1.1E-15 5.2E-16 3.2E-11 2.2E-15 1.3E-11 2.0E-18	file See for file See for s\SpecificExpre expression iscion Aver iscion Iscion iscin Iscin	rmat Doy ssion \KGO\gt (s than 0 01 40983412553 46150248555 46150248555 46150248555 46150248555 46150248555 46150248554 4085170407 435454519021 435454519021 435454519021 435454519021 435454519021 4554558779 51213452779	vnload ex.trans Unit ar	eript.tx
V(1) - A systematic biological Data Gene BioModule Fower I 12//(CUID) C C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 .1g system ESE) × 	The disp. For Genome wide cr_ECS (1995) istance: 5000000 cessionZ Convex 4.22-20 6.22-18 1.62-	Constic studi Constic studi Consti	er loca es (Tuly) es (Tuly) Clear d Genes MADZ H-16 H-16 H-15 H-16 H-17 H-16 H-17 H-16 H-17 H-16 H-17 H-16 H-17 H-16 H-17 H-17 H-17	Image: state	Average file See for s\SpecificExpre expression isetion Average 19.965 18.8766 17.0382 16.4852 16.4852 16.4852 16.4852 16.2200 15.2121 15.1457 15.0511 14.4991	rmat Doy asion \KGG\gt than 0.01 40963412553 46150243355 46150243355 46150243355 46150243355 46150243355 4615024355 46150245 4615111635 25945380737 81213452779 5294324726	vnload ex. trans Unit ar	cript.tx s defined SelectAll UnselectAll Run Export Remove
 (Y4) - A systematic biological Data Gene BioModule Power [12///COLUME] (X) (X) (X) (X) (X) (X) (X) (X) (X) (X)	Figure 5	.2.6.2 .xg system KSE) × .i.g system KSE) × .i.g system Group D 	The disp. For Genome wide for Genome wide reg.ECS (1995) istance: 5000000 ressionZ Conver 4.22-20 6.22-18 1.62-18 1.62-18 6.02-18 1.62-19 1.62-18 1.62-	Constic studi Genetic studi	es (Tuly) es (Tuly) Clear d Genes MADZ HIG HIG HIG HIG HIG HIG HIG HIG	Expression Expression Expression Expression Filter out Ratio0fFreq 1.22-25 3.32-24 1.12-25 3.32-24 1.32-21 2.50-20 1.12-15 3.22-16 3.22-16 3.22-15 1.32-11 2.02-18 2.22-13 1.42-17	file See for file See for s\SpecificExpre expression extion Aver is See 18.8766 See 18.8766 See 18.8766 See 18.8766 See 16.4852 See 16.2200 See 15.2201 See 15.1657 See 14.4991 See 14.2005 See	reat Dov si on \K66\gt than 0.01 4098341255 40490244255 4049029440355 4049029440 40851300817 1693170407 43454540214 404902084 00476031627 25095380737 82943224894 00476031627 25095380737 82943227726	vnload ex. trans	cript.tx
 V4) - A systematic biological Data Gene BioHodule Power I 12/740110 C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 	The disp. for Genome-wide for Genome-wide (1995) (Genetic studi Genetic studi	Er loca es (Tuly) Clear d Genes MADZ -16 -16 -16 -15 -15 -15 -16 -16 -16 -16 -16 -16 -16 -16	Expression Expression Expression Expression Expression Filter out Ratio0fFroj 1.2E-25 3.3E-24 4.1E-21 1.3E-21 2.5E-130 1.1E-15 5.2E-16 2.2E-13 1.4E-17 1.9E-17 1.9E-16	file See fo file See fo >SpecificExpre expression 19.9665 18.8766 17.7222 16.4852 16.4525 16.4525 16.4521 15.1221 15.1511 14.7071 14.3255 14.2927 13.0322	rmst Doy ssion/XGG/gt/ ssion/X	vnload ex. trans	eript.tx
 (Y4) - A systematic biological Data Gene BioModule Power I (27/40110) C C C C C C C C C C C C C C C C C C C	Figure 5	R.g. system RSE) × set: genes_sc val i & Hochberg Group D RobustRege 3.72-18 7.52-18 3.72-18 3.72-18 3.72-18 3.72-18 3.72-18 3.72-18 3.72-16 3.82-16 7.72-16 3.72-16 3.72-16 1.92-15	The disp. For Genome wide cc_ECS (1995) is stance: 5000000 essionZ Conver 4.22-20 6.22-18 1.62-18 1.62-18 1.62-19 1.12-16 6.02-18 3.82-16 3.82-16 3.82-14 4.32-13 4.62-14 4.32-13 8.72-14 4.32-13 8.72-14 1.62-11 5.44-12 5.45-15 5.45	Conctic studi Conctic studi Concti	er loca es (Tuly) es (Tuly) Clear d Genes MADZ 1-16 1-17 1-18	Expression Expression Expression E: \Project: Filter out RatioOfFroj 1.2E-25 3.3E-24 4.1E-21 1.3E-21 1.3E-21 1.3E-21 1.3E-21 1.3E-21 1.3E-21 1.3E-15 2.2E-16 3.2E-11 2.2E-15 1.3E-11 2.2E-13 4.1E-21 1.5E-16	Aver file See for s\SpecificExpre extion Aver 19.9665 18.8766 16.4852 16.4852 16.4852 16.4967 15.221 15.4909 15.215 15.111 14.7071 14.2997 13.0322 12.8675	rmat Doy szion XGO'gt. i than 0.01 4099342553 46150248355 46150248355 46150248355 46150248355 46150248355 46150248355 4615024855 461502485 46150248649 064786031627 25095380137 81213452779 52095380137 81213452779 52095380137 81213452779 5005380137 46150248 46150246	Vnload ex. trans Unit at	eript.tx
 (Y4) - A systematic biological t Data Gene BioModule Power I 12//(40110) C C C C C C C C C C C C C C C C C C C	Figure 5	Ag system ist genes_sc ist genes_sc ist genes_sc ist genes_sc ist genes_sc ist sc ist genes_sc ist sc ist<	The disp. For Genome-wide cc_ECS (1995) tistance: 5000000 cessionZ Conver 4.22-20 6.22-18 1.62	Conctic studi Genetic studi	es (Tuly) es (Tuly) clear d d Genes MADZ d Genes MADZ d Genes into into into into into into into into	Expression Expression Expression E: \Project: Filter out Retio0EProj 1.2E-25 3.3E-24 4.1E-21 1.3E-21 3.0E-20 1.1E-15 5.2E-16 3.2E-11 2.2E-15 1.3E-11 2.2E-15 1.3E-11 2.4E-13 4.4E-17 1.5E-16 3.2E-15 3.3E-11 2.2E-13 3.4E-17 3.4E-17 1.5E-16 4.1E-12 1.3E-12	Aver See for file See for s\SpecificExpre expression ection Aver 19.9665 18.8766 17.0328 16.4852 16.4852 16.4852 16.4454 15.200 15.121 15.1677 15.071 14.4991 14.2297 13.0322 12.8796 12.5796	rmat Doy asion \X56 \gt. than 0.01 40963412553 46150243355 46150243355 46150243355 46150243051 4054741902 345454741902 345454741902 345454741902 345454741902 345454741902 345454741902 345454741903 3455454741903 345454741903 345454741903 345454741903 345454741903 345454741903 345454741903 345454749 3455454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 345454749 34545474749 34545474749 345454747493454547474747474747474747474747474747474	vnload ex. trans Unit ar	cript.tx
(Y4) - A systematic biological Data Cene BioModule Power I (Y/ACUIN) C C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 .1.g system 	The disp. For Genome wide For Genome wide reg.ECS (1995) tistance: 5000000 4.22-20 6.22-18 1.02-18 1.02-16 1.12-16 6.02-16 1.12-16 5.02-16 1.12-16 5.02-14 4.33-13 4.02-14 4.33-13 8.72-14 4.33-13 8.72-14 4.33-13 8.72-14 4.33-13 8.72-14 1.0	Constic studi Constic studi Consti	es (Tuly) es (Tuly) (Clear) d Genes MADZ -16 -15 -16 -15 -16 -15 -16 -15 -16 -15 -16 -17 -16 -16 -17 -16 -17 -16 -17 -16 -17 -16 -17 -16 -17 -16 -17 -16 -17 -17 -17 -17 -17 -17 -17 -17	Lapression Lapression Expression Fxpression Expression E:\Project: Filter out Ratio02Froj 1.22-25 3.37-24 4.18-21 1.38-21 3.52-16 3.22-16 3.22-16 3.22-16 3.22-16 3.22-16 3.22-17 1.32-11 2.21-35 1.22-13 1.42-17 1.52-16 3.22-11 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52-17 1.52-16 1.52	file See fo file See fo s\SpecificExpre expression extion Aver iection Aver iection Aver is SpecificExpre iection Aver is SpecificExpre is	rmat Doy asion\KGG\gt than 0.01 agedLog(p) 40963412553 46150248355 46150248355 464802054 40851300817 1693170407 34545540214 404802031827 34545540214 98470631827 32945387728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 52943287728 81213452779 81213452779 52943287728 81213452779 52943287728 81213452779 8121345779 812134779 812134779 812134779 812134779 812134779 812134779 812134779 812134779 812134779 812134779 8121347779 8121347779 81213477777777777777777777777777777777777	vnload ex. trans	cript.tx s defined SelectAll UnselectAll Run Export Remove
(V4) - A systematic biological buts Gene BioModule Power I 12//COLUME Compared Compared Compared HWindow Resource Window Compared Genes Fryulue Files: MCC2_SC2_scummary.txt Genese: genomel Source : PCC2_SC2_scummary.t Version : hg19 Gene database : Reform Ward columns : [pval] Gene S' extension : 5.0 Gene database : [pval] Gene S' extension : 5.0 Gene 3' extension : 5.0 Gene 6 canse : [pval] Haltiveriste Test : Ho F value sources : [pval] F value sources : [pval] F value sources : [pval] F value sources : [pval] F value sources : [pval]	Figure 5	.2.6.2 .3.g system ISED × ISED ×	The disp. For Genome wide for Genome wide reg.ECS (1995) ressionZ Convex 4. 22-20 6. 22-18 1. 62-18 1. 62-19 1. 62-18 1. 72-14 1. 72-11 1. 72-11 1. 72-11 1. 72-11 1. 72-18 1. 72-1	Genetic studi Genetic studi	Clear MADZ Clear d Genes MADZ 1-16 1-15 1-16 1-15 1-16 1-16 1-16 1-16 1-16 1-16 1-16 1-16 1-17 1-18 1-16 1-18 1-16 1-18 1-18 1-19 1-11 1-11 1-12 1-12 1-12	Laboration Content of	file See for file See for s\SpecificExpre expression extion Aver iection Aver iection 19.9665 18.8766 18.7722 17.7222 17.0388 16.4562 15.2121 15.1457 15.0511 14.2997 13.0322 12.8559 12.5128 12.5122 11.0552 11.9772 11.0565	Prest Dow ssion\X66G\gt stion\X66G\gt stion\X66G\gt stion\X66G\gt 4089241255 34052424355 448020544 346551300817 1693174072 34545540214 492248940 3451300817 2509432248940 34545240214 904352248940 345529697 904352297726 8943313556 904335299773 30781674304 90213224167 30791874304 07006289186 6180644072 6180644072 8180644072	vnload ex. trans	cript.tx
(Y4) - A systematic biological buts Gene BioModule Power I to//courne Compared Co	Figure 5	RGE 2.6.2 RGE system RSE x set: genes_sc val i i & Hochberg Group D Group D i & Hochberg 0.372-18 0.372-18 7.52-18 0.322-12 2.62-17 1.22-16 1.22-16 3.62-16 7.372-16 6.32-12 2.62-17 1.22-16 1.22-16 3.62-16 7.72-16 6.32-15 2.52-14 6.92-15 1.92-15 1.92-15 1.12-14 4.52-11 1.12-14 1.92-16	The disp. for Genome-wide for Genome-wide (1995) istance: 5000000 6.22-18 1.62-18 1.62-18 1.62-18 1.62-18 1.62-18 3.62-18 3.62-14 4.52-12 3.62-14 4.52-14 4.52-13 5.62-14 4.52-14 1.62-11 5.42-12 1.62-11 1.62-12 1.72-12 1.72	Constic studi Genetic studi Constic studi Consti	er loca es (fuly) es (fuly) (Clear) d Genes MADZ (Clear) d Genes MADZ (Clear) d Genes (Clear) (Clear) (Cl	Expression Expression Expression Expression Filter out Filter out 1.2E-25 3.3E-24 4.1E-21 1.3E-25 3.5E-10 1.1E-15 S.2E-16 3.2E-18 1.2E-25 1.3E-21 2.5E-16 1.2E-15 1.2E-15 1.3E-11 2.0E-18 2.2E-13 1.4E-17 5.5E-16 4.1E-12 5.5E-12 4.6E-9 2.7E-0	file See fo file See fo >SpecificExpre expression less expression less iection Aver iection 19.9665 18.8766 17.2722 17.0388 16.4852 16.4852 16.2221 15.222 17.0388 16.4352 16.2009 15.212 15.4651 14.2097 12.8979 12.5596 12.25596 12.5599 12.5599 12.5599 12.5599 10.5695 9.42749 0.42749 0.42749	rmat Doy ssion/XGO/gt/ ssion/XGO/gt/ ssion/XGO/gt/ d0851242553 d44080242553 d4405024355 844802084 d0551300617 1495124005 1300617 1495130407 3454554021 25095300737 31213452779 25095300737 31213452779 2509530737 31213452779 2509530737 31213452779 2509530737 31213452779 2509530737 3121345279 2509530737 3121345279 2509530737 3121345279 2509530737 3121345279 2509530737 3121345279 2509530737 25095312 2	vnioad ex. trans	eript.tx s defined SelectAll UnselectAll Run Export Remove
 V4) - A systematic biological Data Gene BioNodule Power I V2/740010 C C C C C C C C C C C C C C C C C C	Figure 5	R. 5. 2. R. 5. 2. RG 575108 S. 2. RSE) × S. 2. set: genes_cov S. 2. set: genes_cov S. 2. set: Group D S. 2. S. 2. S. 2.	The disp. For Genome wide cr_ECS (1995) is stance: 5000000 6 .22-16 0 .62-16 0 .62-16 1 .18-16 0 .62-16 1 .18-16 0 .62-18 1 .62-18 1 .62-13 3 .82-10 3 .82-10 3 .82-13 3 .82-13 3 .82-13 3 .82-13 3 .82-13 3 .82-13 3 .82-13 3 .82-14 1 .82-13 9 .82-14 1 .82-13 9 .82-14 1 .82-13 9 .82-14 1 .82-11 1 .18-11 1	Conctic studi Genetic studi Conctic studi Concti	er loca es (Tuly) es (Tuly) Clear d Genes MADZ 1-16 1-17 1-16 1-16 1-16 1-16 1-16 1-16 1-16 1-16 1-16 1-16 1-17 1-16 1-16 1-16 1-17 1-16 1-16 1-16 1-17 1-16 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-16 1-17 1-17 1-16 1-17 1-17 1-16 1-17 1-17 1-16 1-17 1-17 1-16 1-17 1-17 1-16 1-17	Loging generation Expression Expression E: \Project: Filter out Batio0EProj 1.28-25 3.392-24 4.18-21 1.38-21 2.58-13 3.06-20 1.18-21 1.38-21 2.28-16 3.22+11 3.78-21 1.18-21 1.98-17 1.98-17 1.98-17 1.98-17 1.98-16 2.78-12 4.67-9 2.78-9 1.18-8	Aver file See for file See for s\SpecificExpre expression ection Aver 19.9065 18.8766 18.8766 17.722 17.0388 16.4852 16.4452 15.209 15.212 15.1457 15.457 15.051 14.2997 14.3250 14.2997 13.0322 12.5596 12.5526 12.5596 12.5526 13.0322 11.9075 14.2997 11.0566 9.24380 9.24380	rmat Doy ssion \KGO\gt. stian \GO\gt. 40983412553 46150248355 46150248355 46150248355 46150248355 46150248355 46150248355 46150248355 46150248355 46150248355 46150248355 4615024405 4615024949 4615024949 4615024949 4615024949 4615024949 461502494 461502494 461502494 461502494 461502494 461502494 461502494 461502494 4615024 46	Vnload	eript.tx
 (Y4) - A systematic biological Data Gene BioModule Power I (Y4/COLUD) C C C C C C C C C C C C C C C C C C C	Figure 5	.2.6.2 .ac system ESE) × .ac system ESE) × .ac genes_sc val .ac Kochberg .ac Strong D .ac	The disp. For Genome-wide for Genome-wide cr_ECS (1995) tistance: 5000000 tistance: 5000000 4.22-20 6.22-18 1.6	Cenetic studi Genetic studi	er loca es (Tuly) es (Tuly) Clear d Genes MADZ -16 -16 -17 -15 -16 -17 -15 -16 -17 -15 -16 -17 -15 -16 -17 -15 -16 -17 -17 -16 -17 -17 -17 -18 -17 -19 -18 -17 -19 -18 -17 -19 -18 -17 -19 -18 -17 -19 -19 -19 -19 -19 -19 -19 -19	Laboration Content of	Average file See for s\SpecificExpre expression extion Average is SpecificExpre extion Average is SpecificExpre extion Average is SpecificExpre is SpecificExpr is	rmat Doy asi on \KGG\gt gt than 0.01 409603412553 dt150248355 441502484355 dt150248355 44150248435 dt150248355 845020584.0714 dt93170407 40551300517 dt952248940 94515248745 dt952248940 945154111035 2075331355 90733529697138 dt9325248940 90748597138 dt933524897138 90748597138 dt935248940 902824187 3079187404 90708289188 d1980844072 9283990281 5103308768 91301817709 7002514022	vnload ex. trans	cript.tx

6. Power estimation of set-based tests by SPS

STEP 1: Power estimation, Power \rightarrow calculator. The interface is divided into two parts. Set the basic parameters on the left, then you can get the results on the right.

🜉 SPS: a simulation tool for calculating power of set-based genetic association tests	- o x
All variants Self-define Real data-plink Real data-vef Total Variants: 20 Repeat Region: 1	Results Displaying Power at sets ID Odds Freq Humbe LD GATES ScaChi HIST SKAT
Minor Allele Frequency (MAF): (from) 0.02 (to) 0.07 (step) 0.05 SWP Dependence: O Independent Dependent Linkage Discoulibrium (LD, 0.6 (to) 0.6 (step) 0.1	Ĵ
Reference Li et al., SFS: a simulation tool for calculating power of set-based genetic association tests. Genet Epidemiol. 2015;39(5):395-7	Power at variants ID Odds Ratio Frequence Number LD Powers
Shal Load SKAT (The PC should install R language [>=2.13.0] first!) Download R Bisk variants	v
Risk SNFs: (from) 1 (to) 3 (step) 1 Odds Ratio: (from) 1.8 (to) 2.2 (step) 0.05 Disease Prevalence: 0.05 Genetic Model: Additive Model ✓ Position of Risk Veriants:	Chart Options Change MAF: Change LD:
Population & Sample Population Size: 500000 Humber of Case: 500 Humber of Control: 500	
Simulation & Test Sampling Times: 1000 P Value Threshold: 0.05 Parallel Running Humber 3 Meta-analysis: No ~ Start Stop 0%	

Figure 6.1. The Main interface of SPS.

STEP 2: Set the parameters of all variants, including the number of SNPs, the minor allele frequency (MAF) and LD information. If these SNP markers are divided into several LD blocks, the markers in the same LD block have the same LD with each other. But the LD is set to 0 when the markers belong to different blocks. All of these markers and their LD pattern can be replicated to make up of a larger marker set. Some of these parameters can vary in a certain region, such as MAF and LD, so that the users can investigate how powerful will be affected by changing the critical parameters conveniently. In addition, these parameters can also read from the real data (Plink binary genotype files and vcf file). In this case, the LD information will be calculated from the input genotypes.

Self-define Real data-plink Real data-vcf	
Yotal Variants: 20 LD Block: 2 Repeat Region: 1	
Minor Allele Frequency (MAF): (from) 0.02 (to) 0.07 (step) 0.05	
NP Dependence: 🔿 Independent 💿 Dependent	
inkage Disequilibrium (LD, 0.6 (to) 0.6 (step) 0.1	
Reference	^
Li et al., SPS: a simulation tool for calculating power of set-based genetic association tests.	
Genet Epidemiol. 2015;39(5):395-7	

Figure 6.2.1 Set parameters by users.

All variants	
Self-define Real data-plink Real data-vcf	
Family File: E:\KGG\plink\test.fam	
Map File: E:\KGG\plink\test.bim	
BED File: E:\KGG\plink\test.bed	
Consider the first 10 variants; Repeat Region 1	
SNP Dependence: Independent Dependent 	

Figure 6.2.2 Set parameters by plink file.

All variants Self-define Real data-plink Real data-vcf			
VCF File: E:\KGG\vcf_example.vcf			
SNP Dependence: Independent Dependent			

Figure 6.2.3 Set parameters by vcf file.

Parameter	Description
Total Variants	The total number of SNDs tested in a set
	The number of LD blocks. Variants in the same block are in LD and
LD Block	The number of LD blocks. Variants in the same block are in LD and
	that in different blocks have no LD.
Repeat Region	The number of copies of SNPs. The SNP will be copied for several
Ttepeut Ttegion	times to form a larger set and so does the LD pattern of the.
	The frequency of the least common allele occurs in the population.
Minor Allele Frequency	The MAF can increase from a initial value to a terminal value
	according to a step value that set from the GUI.
	The relationship between SNPs. If the SNPs are dependent, the user
	should set the LD value (r), otherwise 0 is set as default. The LD
SNP Dependence	information can also be read from the real data, where it will be
	calculated based on the allele frequency.
	The r score used to represent LD information. The SNPs in the same
	block are dependent and keep the same r value, while SNPs in the
Linkage Disequilibrium	different blocks are independent with each other and the r value is set
(LD, r)	as 0. The r value can also increase from an initial value to a final
	value hy a sten value
	The path of the Plink files. The valid file path can be input by the
Family File	button on the right. If the three files have the same file prefix and are
Map File	stored in the same directory the other file nother will be filled
BED File	stored in the same directory, the other the paths will be fined
	automatically when one file is set.
Consider the first	The number of SNP that input from the real data. The real data
several SNPs	usually include large size of SNPs, which is unnecessary for our
	simulation. Hence, we just consider the first several SNPs as our

Table to list parameters:

	study objects.
VCF File	The path of a VCF file.

STEP 3: Set parameters of risk variants.

Risk variants			
Risk SNPs: (from)	1 (to) 3	(step) 1	
Odds Ratio: (from)	1.8 (to) 2.2	(step) 0.05	
Disease Prevalence:	0.05	Genetic Model:	Additive Model 🛛 👻
Position of Risk Varian	nts:		Random
	(Start from 1; Sep	parated by space o	or comma.)

Figure 6.3 Set parameters about risk variants.

Parameter	Description
Diale SNDa	The number of risk SNPs. This parameter can increase from a smaller
KISK SINPS	to a larger value step by step.
	The value used to quantify the association between risk SNPs and
Odds Ratio	disease. This parameter can increase from a smaller to a larger value
	step by step.
Disaasa Provolongo	The proportion of a population found to suffer the disease. This will
Disease Flevalence	be used in the genetic model.
Constis Madal	The genetic model of risk loci. The additive model and multiplicative
Genetic Model	model are candidates in SPS.
Desition of Disk	The location information of risk variants within the total variants. The
Position of KISK	users can click the random button for automatic setting or set by
variants	themselves.

Table to list parameters:

STEP 4: Set population and sample. The larger population size and number of case and control are recommended, because they make the result more accurate and stable, but it will take more time correspondingly. So the user should keep balance between them.

Population & Sample			
Population Size:	500000		
Number of Case:	500	Number of Control:	500

Figure 6.4. Set population and sample.

Table to list parameters:	
Parameter	Description
Population Size	The number of individuals in a population generated by simulation according to the certain genotype and phenotype.
Number of Case	The number of individuals that suffer the disease.
Number of Control	The number of individuals that do not suffer the disease.

STEP 5: Set simulation and meta-analysis parameters. A number of case-control samples will be randomly drawn with replacement from the population. And they are subject to calculate the p value of the set-based test. The number of p values passing the threshold will be counted to calculate the power. In order to speed up the simulation process, you can set several parallel

threads, but more memory resource is needed.

The meta-analysis can be carried out at the variant level or set level. At variant level, the p values of variants in different studies will be combined according Fisher's Combination Test and these meta-p values will be treated by GATES, ScaChi and HYTS. Alternatively, at set level, the p value of variants in a set should be conducted by GATES, ScaChi and HYTS, and then the set-based p values in different studies are aggregated. SPS can also mimic locus heterogeneity by randomizing risk loci of each study in meta-analysis.

Simulation & Test		
Sampling Times: 1000 P V	alue Threshold:	0.05
Parallel Running Number 3		
Meta-analysis: No 👻		

Figure 6.5.1 Set simulation without meta-analysis.

Simulation & Test
Sampling Times: 1000 P Value Threshold: 0.05
Parallel Running Number 3
Meta-analysis: At variants Number of Studies: 3
Randomize risk loci of each study (mimic genetic locus heterogeneity)

Figure 6.5.2 Set simulation with meta-analysis.

Parameter	Description				
Sampling Times	The number samples randomly drawn from the case and				
Sampling Times	control group. For each time, a case-control study is achieved.				
	The threshold of type I error that used in the case-control study.				
P Value Threshold	For SNP-based test, the bonferroni correction is conducted as				
	default.				
	The number of threads that running concurrently. The multiple				
Parallel Running Number	threads mechanism is used here to speed up the running of				
	program. However, this may cost a large volume of memory.				
Moto analysis	Whether to perform meta-analysis. If performed, the users				
Wieta-allarysis	should choose the meta-analysis at variants level or at set level.				
Number of Studies	The number of studies considered in the meta-analysis.				
Dandamina risk lasi of each	Whether to consider the genetic heterogeneity. If considered,				
study	the position of risk loci of each study will be set randomly to				
study	mimic the heterogeneity.				

Table to list parameters:

STEP 6: Run the program. Click the Start button and run the program. The results from tables are shown in the right part immediately. The progress bar provides the real time information of running. If you want to stop the running program, just click the "stop" button.

🔤 SPS: a simulation tool for calculating power of set-based genetic association tests								-	- 🗆 X
All variants	Results Dis	playing							
Self-define Real data-plink Real data-vcf									Power at sets
Total Variants: 20 Repeat Region: 1	ID	0dds	Freq	Numbe	LD	GATES	ScaChi	HYST	SKAT
	15	1.9	0.07	1	U. 6	0.9310	0.8431	0.9140	. NA
Minor Allele Frequency (MAF): (from) 0.02 (to) 0.07 (step) 0.05	16	1.9	0.07	2	0.6	0.9820	0.9630	0.9900	. NA
SNP Dependence: O Independent 💿 Dependent	17	1.9	0.07	3	0.6	1.0	1.0	1.0	NA
	18	1.95	0.02	1	0.6	0.4915	0.3586	0.4445	. NA
Linkage Disequilibrium (LD, 0.6 (To) 0.6 (Step) 0.1	19	1.95	0.02	2	0.6	0.6853	0.6723	0. 7292	. NA 🗸
Reference Li et al., SPS: a simulation tool for calculating								Powe	r at variants
power of set-based genetic association tests.	ID	Odds R	atio Fr	equence	Number .	LD	P-1		P-2
Genet Epidemiol. 2015;39(5):395-7	10	1.9	U. (JI	2	U. 6	0.20	97902	0.248/512
	17	1.9	0.0	07	3	0.6	0.83	21678	0.8681318
0717	18	1.95	0.0	02	1	0.6	0.09	69030	0.0659340
ShRi	19	1.95	0.0	02	2	0.6	0.07	39260	0.0909090 🗸
Load SKAT (The PC should install R language [>=2.13.0] first!) <u>Download</u> R	<								>
Risk varients Risk SUFs: (from) 1 (to) 3 (step) 1 Odds Ratio: (from) 1 8 (to) 2.2 (step) 0.05 Disease Prevalence: 0.05 Genetic Model: Additive Model > Position of Risk Variants: 8 12 10 Population & Sample Population Size: 500000 Humber of Case: 500 Humber of Control: 500 Simulation & Test Sampling Times: 1000 P value Threshold: 0.05 Parallel Running Humber 3	Chart Opt Change M. 0.02 Change Li 0.6	ions kF: v v v							odds — power

Figure 6.6 Run the program.

STEP 7: Save the result. Users can review the power from tables at the SNP level and set level. A line chart is drawn to show the variation of power within different odd ratios with the given MAF and LD information. You can also change the MAF and LD values to update the chart, and right-click on the tables to save the results as excel files or txt files. The chart can be saved by right-click as well.

🚎 SPS: a simulation tool for calculating power of set-based genetic association tests	- 🗆 X
All variants Self-define Real data-plink Real data-vef	Results Displaying Power at sats
Total Variants: 20 LD Block: 2 Repeat Region: 1 Minor Allele Frequency (MAF): (from) 0.02 (to) 0.07 (step) 0.05 SNF Dependence: O Independent © Dependent Linkage Disequilibrium (LD, 0.6 (to) 0.6 (step) 0.1 Reference Li et al., SPS: a simulation tool for calculating power of set-based genetic association tests. Genet Epidemiol. 2015;39(5):395-7	ID Odds Freq Humbe LD OATES SeaChi HUST Stat 49 2.2 0.02 2 0.6 0.9800 0.9990 9.990 BA 50 2.2 0.02 3 0.6 0.99910 0.9990 9.990 BA 51 2.2 0.07 1 0.6 0.99910 0.9930 HA 52 2.2 0.07 2 0.6 1.0 1.0 I.0 HA 53 2.2 0.07 3 0.6 1.0 1.0 HA × Forest at variants D Ddds Ratio Prequence Humber LD P-1 P-2# 0.0 2.2 0.07 1 0.6 0.5313655 0.960049 × 51 2.2 0.07 1 0.6 0.5313655 0.9600494 × 52 2.2 0.07 1 0.6
SKAT	53 2.2 0.07 3 0.6 0.9700299 1.0 > > > > >
Disease Prevalence: 0.05 Genetic Model: Additive Model \checkmark Position of Risk Variants: 2 4 13 Population & Sample	$\begin{array}{c} 0.02 \\ 0.02 \\ 0.8 \\ 0.6 \\ 0.5$
Population Size: 500000 Number of Case: 500 Number of Control: 500 Simulation & Test Sampling Times: 1000 P Value Threshold: 0.05	0.4 0.3 0.2 0.1 0.0 1.80 1.85 1.90 1.95 2.00 2.05 2.10 2.15 2.20
Parallel Running Number 3 Meta-analysis: No v	Odds GATES-1 → ScaCh-1 → HYTS-1 → GATES-2 → ScaCh-2 → HYTS-2 → GATES-3 → ScaCh-3 → HYTS-3

Figure 6.7.1 The output of SPS.

ID	Odds Ratio	Frequence	er of Risk Al	LD	GATES	ScaChi	HYST
0	1	0.1	1	0.5	0.055	0.025	0.048
1	1	0.1	1	0.б	0.04	0.016	0.033
2	1	0.1	2	0.5	0.042	0.025	0.041
3	1	0.1	2	0.б	0.047	0.021	0.047
4	1	0.1	3	0.5	0.049	0.017	0.048
5	1	0.1	3	0.б	0.049	0.029	0.044
б	1	0.1	4	0.5	0.044	0.028	0.049
7	1	0.1	4	0.6	0.062	0.027	0.055

Figure 6.7.2 The saved table of set-based power.

ID	Odds Ratio	Frequence	er of Risk Al	LD	P-1	P-2	P-3*	P-4	P-5	P-6	P-7*
0	1	0.1	1	0.5	0.002	0.004	0.005	0.001	0.001	0.002	0.001
1	1	0.1	1	0.6	0.003	0.001	0	0.001	0	0.003	0.001
2	1	0.1	2	0.5	0.003	0.003	0.001	0.002	0.001	0.001	0.002
3	1	0.1	2	0.6	0.003	0.006	0.005	0.002	0.002	0.001	0.004
4	1	0.1	3	0.5	0.004	0.002	0.001	0.003	0.003	0	0.001
5	1	0.1	3	0.6	0.003	0.002	0.003	0.001	0.002	0.001	0.004
6	1	0.1	4	0.5	0.003	0.004	0.003	0.004	0.003	0.004	0.004
7	1	0.1	4	0.6	0.001	0.002	0.001	0.003	0.004	0.002	0.001
8	1	0.2	1	0.5	0.004	0.004	0.005	0.002	0.004	0.002	0.005
9	1	0.2	1	0.6	0.008	0.004	0.003	0.001	0.003	0	0.005
			< - 0								

Figure 6.7.3 The saved table of variant-based power.

STEP 8: The SKAT tool is integrated into SPS to detect the significant SNPs, especially for rare variants. In order to running SKAT, you should install R software (version >=2.13.0) firstly. Please click the tag 'Download R' to download R, and several R packages are needed too. These packages are "RServe", "SKAT" and "snow". "RServe" is a java-r interface. "Snow" is used to run SKAT in parallel. If you don't know how to install these packages, just paste the prompt message in your R platform. When SKAT is ticked, the power calculated by SKAT will be added to the table and chart. Although SPS can run SKAT in parallel, this is also a time-consuming part and the sampling time should not be set too large.



Figure 6.8.1 SKAT option.



Figure 6.8.2 The prompt message.



Figure 6.8.3 The output when SKAT added.

7. Updates from KGG3.5 to KGG4.0

Much progress was made from KGG 3.0 to KGG 3.5, mainly including:

- 1) A more powerful gene-based association analysis by effective Chi-squared test;
- 2) A powerful conditional gene-based association analysis by effective Chi-squared test;
- 3) A robust geneset based association analysis by LD-attenuating rank-sum test;

🕼 KGG(V4) - A systematic biological Knowledge-based mining system for Genome-wide Genetic studies (July 1, 2019)											
Project Data (sene BioWodule Yower 1月(1) 韵口(W) 辞即(H)											
1571/7401HB 🔞 🗲 🗋 🔻 🕹 🐣											
	🚘 Driver Tissue (DER	SE) ×									
KGGscz	Gene Association Se	et: genes_scz_ECS		T	Expression file	See format Downloa	a				
PGC2_SC	P Value Source: pv	al	•	Clear							
Sourc	Method: Benjamini	& Hochberg (1995)	•	Load Genes	ects SpecificExpre	ession\AGG\gtex.transcrip	t. txt \getx.				
Versi	Error Rate: 0.05	Group Distance:	5000000 bp		Filter out express	ion less than 0.01 Unit	as defined				
JIred	,		,								
- 🖌 Gene	Genes Tissues						SelectAll				
🛛 🖍 Gene	TissueName	RobustRegressionZ	ConventionalZ	MADZ	RatioOfProjection	AveragedLog(p)	Jerectari				
STP I	Brain-FrontalCor	4.8E-21	4.2E-20	5.3E-16	1.2E-25	19.966540983412553	UnselectAll				
🚽 🖍 🖌 🖌 🖌 🖌 🖌	Brain-Anteriorci	7.0E-20	6.2E-18	2.2E-16	3.3E-24	18.876646150248355					
🖻 🎆 Gene Sca	Brain-Hippocampus	3.7E-18	1.6E-16	3.4E-16	4.1E-21	17.27228448620584	Run				
Genor	Brain-Cortex	7.5E-18	1.1E-16	6.3E-15	1.3E-21	17.03880851300817					
2 H-142	Esophagus-Muscul	6.3E-22	6.0E-16	1.2E-17	2.5E-13	16.48521693170407	Export				
	Brain-Nucleusacc	2.6E-17	3.8E-16	5.5E-15	3.0E-20	16.44644954741902					
	Artery-Tibial	1.9E-18	7.8E-17	5.2E-16	1.1E-15	16.024034545540214	Remove				
Test	Brain-Cerebellar	8.8E-18	1			15.29094932248949					
Show	Esophagus-Gastro	1.7E-21				15.212108476031627					
	Brain-Spinalcord	1.2E-18				15.145725095380737					
	Colon-Sigmoid	1.2E-16	The conditional gene-	based test has been	finished!	15.051181213452779					
	Brain-Caudate (ba	3.8E-16				14.707152943287726					
	Artery-Aorta	7.4E-16		ie 1		14. 499184514111635					
	Brain-Amygdala	3.7E-16	X	10 <u>-</u>		14. 325020753313556					
	Brain-Hypothalamus	7.7E-16	9.0E 14	9.92 13	1.52 1/	14.299790433529697					
	Brain-Putamen (ba	6.3E-15	1.8E-11	4.4E-12	1.5E-16	13.032296960879796					
	Cells-Transforme	1.9E-15	5.4E-12	4.4E-14	4.1E-12	12.68757446397138					
	Brain-Cerebellum	2.5E-14	1.6E-11	1.1E-14	1.3E-12	12.559600218324167					
	Brain-Substantia	8.9E-15	1.1E-11	1.4E-11	6.7E-15	12.512630791674304					
	Heart-AtrialAppe	1.1E-14	7.4E-11	3.8E-13	7.5E-12	11.907207006289186					
	Artery-Coronary	4.5E-11	7.5E-12	3.2E-12	5.6E-12	11.056961980844072					
	Vagina	1.2E-10	1.3E-9	2.8E-11	4.6E-9	9. 427492863990281					
	Adipose-Visceral	1.8E-10	1.9E-9	1.1E-10	2.7E-9	9.248805103380686					
	Muscle-Skeletal	2.2E-12	1.3E-8	5.6E-10	1.1E-8	9.191183018177709					
	Heart-LeftVentricle	1.6E-13	4.3E-7	5.4E-9	7.9E-9	8.881897802514022					