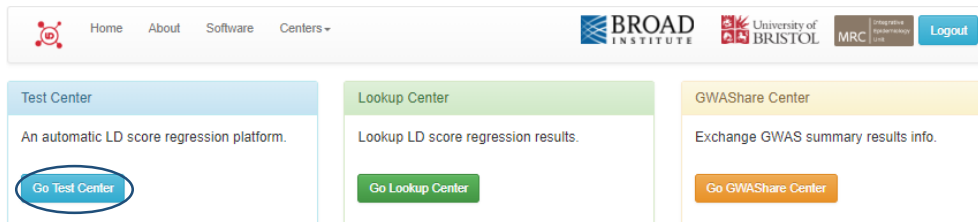
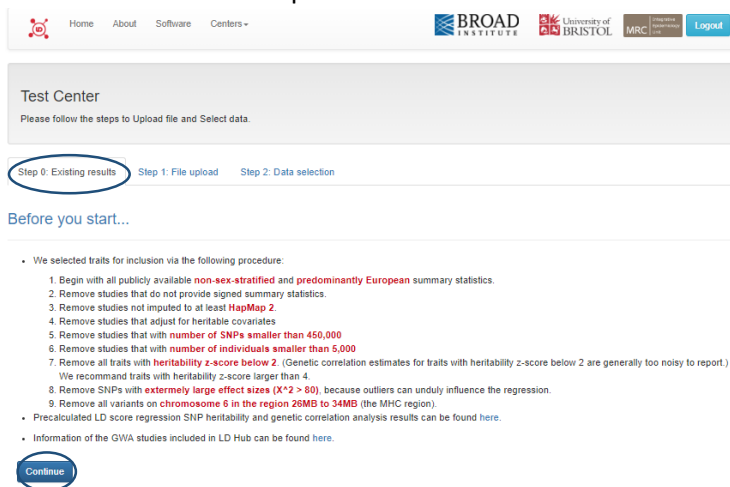


LDSC LD Score Regression

1. Go to the LD Hub website: <http://ldsc.broadinstitute.org/>. Click “Get Started with LD Hub”.
2. Click “Sign in with your Google account”. Use the following account and password:
Account jaxpolygenicity@gmail.com **password** jaxpolygenicity2018, or use your personal gmail account.
3. Go to the “Test Center” tab:



4. You are now on the “Step 0” tab. Click continue:



5. You are now on the “Step 1: File upload” tab. **Download the zip file** provided as an example (see picture below, or follow this link <http://ldsc.broadinstitute.org/static/example/pgc.cross.scz.withN.zip>; zip file name: pgc.cross.scz.withN.zip); this file contains the summary statistics for an early GWAS of schizophrenia. Next, **upload the zip file*** and name the trait (e.g. schizophrenia). Click continue:

Upload the trait of interest

To save your upload time, we highly recommend you to use the SNP list we used in LD Hub to reduce the number of SNPs in your uploaded file. Click [here](#) to download our SNP list (w_hm3.noMHC.snplist.zip).

Please upload the **zipped** file you just created. Click [here](#) to download an input example.

We notice that some users zip folder(s) together with the plain txt file (e.g. /myfolder/mydata.txt). By doing this, you will get an error: [Errno 2] No such file or directory

No file chosen
(required)

Name of your traits (e.g. scz) (required)

*If you are working on a Mac device, your file may get unzipped; make sure you upload a zip file.

6. You are now on the “Step 2: Data selection” tab. Select one or two traits that you may be interested in from the database. For example, you can select “smoking_behavior”, which will calculate genetic correlations between schizophrenia and several smoking-related phenotypes.

7. You have now been directed to a new tab: “Analysis in progress”. Be Careful! Do not close the window or click on any links before **both** the SNP heritability and genetic correlation analyses are done (see below). *This step should take less than two minutes, but will fluctuate depending on the internet connection and the number of traits selected.*

Result 1: QC and SNP heritability analysis

Data quality information

This log file shows some basic metadata about the summary statistics. If mean chi-square is below 1.02, the log file will warn you that the data probably are not suitable for LD Score regression.

SNP heritability

This file shows the heritabilities of your trait from single-trait LD Score regression. These estimates will be biased downwards by QC correction. Note that these heritability estimates are on the observed scale. Lambda GC is $\text{median}(\chi^2)/0.4549$. Mean χ^2 is the mean chi-square statistic. Intercept is the LD Score regression intercept. The intercept should be close to 1, unless the data have been QC corrected, in which case it will often be lower. Ratio is $(\text{intercept} - 1)/(\text{mean}(\chi^2) - 1)$, which measures the proportion of the inflation in the mean χ^2 that the LD Score regression intercept ascribes to causes other than polygenic heritability. The value of ratio should be close to zero, though in practice values of 10-20% are not uncommon, probably due to sample/reference LD Score mismatch or model misspecification (e.g., low LD variants have slightly higher h^2 per SNP).

Reminder: Please check your h^2 results before collect your rg results. 1) Is your h^2 out of bound (e.g. $h^2 < 0$ or $h^2 > 1$)? 2) Is the Z score of your $h^2 > 4$ ($h^2/SE > 4$)? If any of the two points did not applied for your trait, we will expect NA results in your rg results file.

Download QC log file

[Download QC information for the uploaded file](#)

Download SNP heritability results

[Download SNP heritability analysis results](#)

Result 2: Genetic correlation analysis

Genetic correlation

This csv file is a table summarizing all genetic correlation results. The columns are trait1 = trait you uploaded, trait2 = a trait in LD Hub, PMID = the PubMed ID of trait 2, Catlog = the catalog of trait 2, rg = genetic correlation, se = standard error of rg, p = p-value for rg, h^2_{obs} , $h^2_{\text{obs_se}}$ = observed scale h^2 for trait 2 and standard error, h^2_{int} , $h^2_{\text{int_se}}$ = single-trait LD Score regression intercept for trait 2 and standard error, $g_{\text{cov_int}}$, $g_{\text{cov_int_se}}$ = cross-trait LD Score regression intercept and standard error.

Reminder: the rg estimate is above 1.25 or below -1.25 means that one of the h^2 estimates was very close to zero. Since h^2 is in the denominator of rg, if h^2 is close to zero, rg estimates can blow up.

Download genetic correlation results

Your analysis is finished!

[Download genetic correlation analysis results](#)

You can save the following link and download the results within 24 hours.


<http://ftp.broadinstitute.org/statistics/pgc-coron/SCZ172013-05-withN.txt.d15a75ee-28a7-46c2-b23a-f7664754e0b1.rg.results.csv>




Reminder: The analysis of all traits may take about 2 hours.

8. Download or open in new tabs the SNP heritability and genetic correlation results. You can also download the QC (quality control) log file if you want to find out what filters have been applied to the file you uploaded.

9. What is the SNP heritability (“total observed scale h^2 ”) of schizophrenia? Does it show any significant genetic correlations (r_g) with, for example, any smoking related traits?

10. Now go to the “Lookup Center” tab. Here you can browse what is the SNP-heritability of your favorite trait from the database, and lookup pairs of genetic correlations.


[Home](#)
[About](#)
[Software](#)
[Centers](#)




[Logout](#)

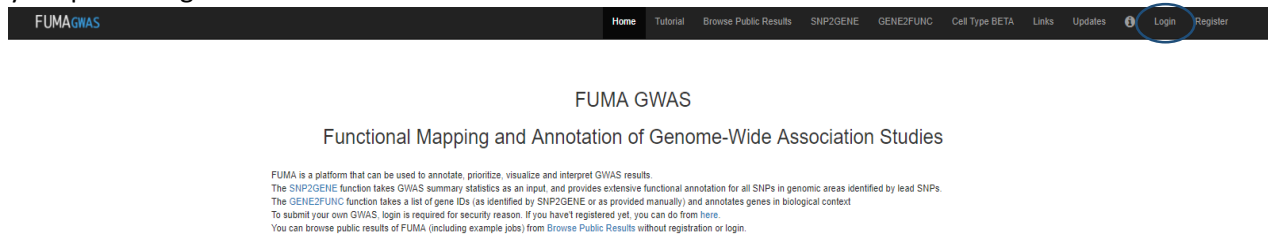
Test Center
 An automatic LD score regression platform.
[Go Test Center](#)

Lookup Center
 Lookup LD score regression results.
[Go Lookup Center](#)

GWAShare Center
 Exchange GWAS summary results info.
[Go GWAShare Center](#)

FUMA GWAS Functional Mapping and Annotation of Genome-Wide Association Studies

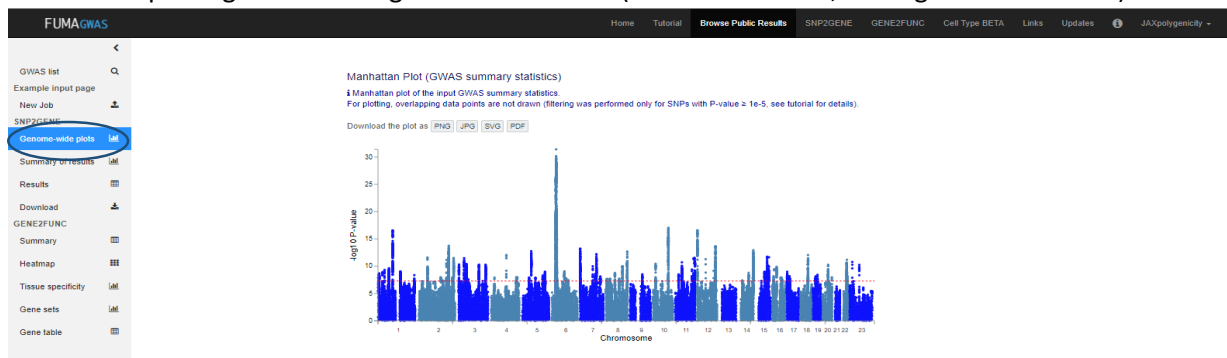
1. Go to the FUMA website: <http://fuma.ctglab.nl>
2. Log in: **Account** jaxpolygenicity@gmail.com **password** jaxpolygenicity2018. You can also register with your personal gmail account.



3. Go to the “Browse Public Results” tab and select “Example results of Schizophrenia”. Use the following questions as a guidance to learn more about the variants and genes associated with this trait.

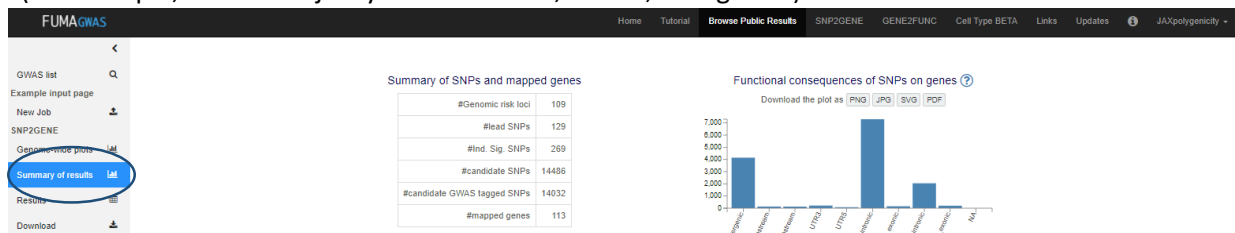
ID	title	author	email	phenotype	publication	sumstats	sumstats reference	notes	created date	last update
19	None	Jasmeen Roosenboom	jar261@pitt.edu	NA	NA	NA	NA	NA	2015-09-06	2015-09-06
15	None	Zalman Vaksman	vaksmanz@mail.chop.edu	NA	NA	NA	NA	NA	2015-08-09	2015-08-09
13	sev_only	JILL	ZCJT061@GMAIL.COM	NA	NA	NA	NA	NA	2015-07-26	2015-07-26
11	v4_vAmy	Alde Cordova	aldoc@stanford.edu	NA	NA	NA	NA	NA	2015-07-03	2015-07-03
10	Grove ASD V2	Oliver Pain	opain01@mail.bbk.ac.uk	NA	NA	NA	NA	NA	2015-06-19	2015-06-19
9	Grove ASD	Oliver Pain	opain01@mail.bbk.ac.uk	NA	NA	NA	NA	NA	2015-06-12	2015-06-12
8	None	xyfckd	xyfckd@mail.ustc.edu.cn	NA	NA	NA	NA	NA	2015-06-06	2015-06-06
6	trop I	Paul Welsh	paul.welsh@glasgow.ac.uk	NA	NA	NA	NA	NA	2015-05-24	2015-05-24
3	Example results of Schizophrenia	Yoko Watanabe	k.watanabe@vu.nl	Schizophrenia	PMID: 29184956	https://www.med.unc.edu/gcr/results-and-downloads/downloads	PMID: 25056061	Predefined lead SNPs in Supplementary Data 14 was used.	2017-05-19	2017-06-24

4. Go to the “Genome-wide plots” tab; you can find the Manhattan plot of the GWAS results. Are there any variants surpassing the GWAS significant threshold (P-value = 5×10^{-8} , or $-\log_{10}$ P-value = 7.3)?

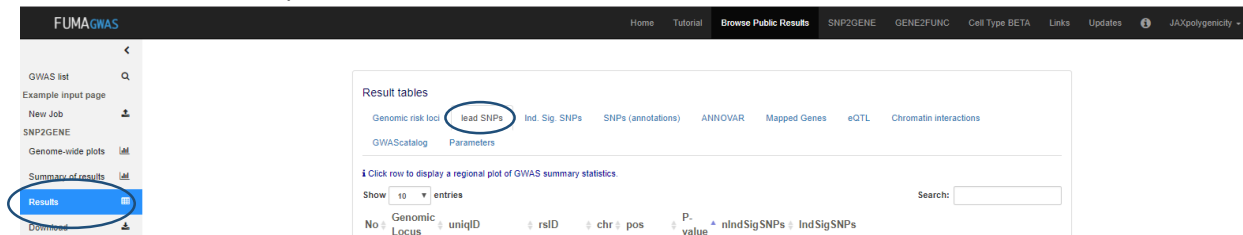


5. In the “Genome-wide plots” tab, scroll down to “MAGMA Gene-Set Analysis”. Can you identify some of the biological functions (pathways) associated with schizophrenia?
6. Next, scroll down to “MAGMA Tissue Expression Analysis”. Where are the majority of the genes expressed?

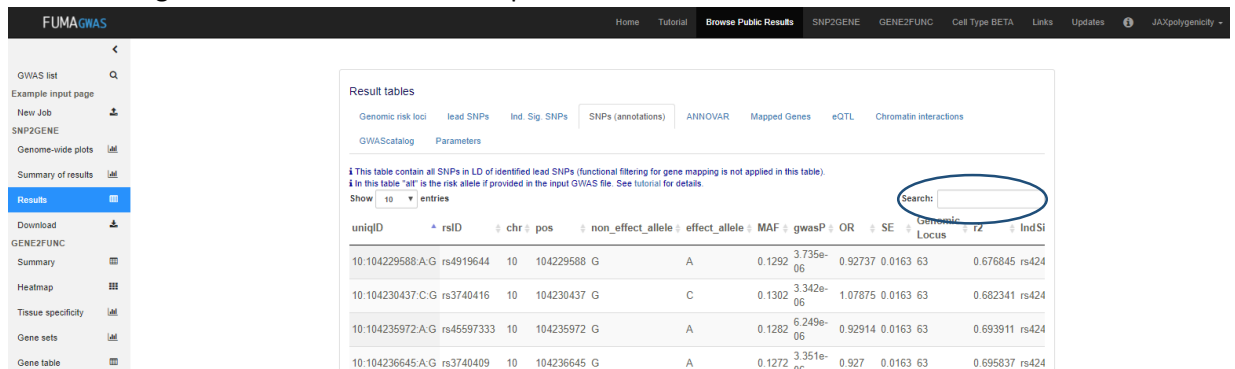
7. Go to the “Summary of results” tab. How many genomic risk loci have been associated with schizophrenia? What are the most common functional consequences of the annotated SNPs on genes (for example, are the majority of the intronic, exonic, intergenic?)



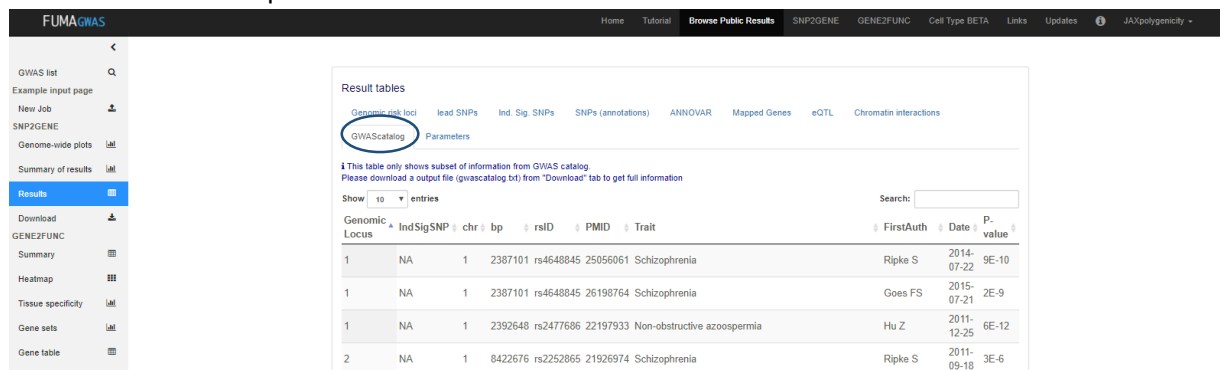
8. Go to the “Results” section, and select the “lead SNPs” tab. How many lead SNPs have been associated with schizophrenia?



9. Go to the “SNPs (annotations)” tab. You can search for your favorite gene (for example, *SLC39A8*) and check if this gene is associated with schizophrenia.



10. Lastly, go to the “GWAScatalog” tab. Are the SNPs associated with schizophrenia candidate SNPs for other traits from independent GWAS studies?



ANSWERS

LDSC

#9. Schizophrenia showed a SNP heritability estimate of 0.57 (SE=0.039). We can identify a positive genetic correlation between schizophrenia and age of smoking onset ($r_g=0.24$, SE=0.12, $P=0.038$), and a nominal positive genetic correlation between schizophrenia and smoking initiation ($r_g=0.11$, SE=0.06, $P=0.053$).

Log file showing the SNP-heritability results:

```
*****
* LD Score Regression (LDSC)
* Version 1.0.0
* (C) 2014-2015 Brendan Bulik-Sullivan and Hilary Finucane
* Broad Institute of MIT and Harvard / MIT Department of Mathematics
* GNU General Public License v3
*****
Call:
./ldsc.py \
--h2 /var/www/ldhub/ldhub_web/static/sumstats/pgc.cross.SCZ17.2013-05.withN.txt.394ab93c-1a72-4d17-9588-e755ca41d6a3.sumstats.gz \
--ref-ld-chr /root/ldsc/eur_w_ld_chr/ \
--out /var/www/ldhub/ldhub_web/static/sumstats/pgc.cross.SCZ17.2013-05.withN.txt.394ab93c-1a72-4d17-9588-e755ca41d6a3-h2 \
--w-ld-chr /root/ldsc/eur_w_ld_chr/

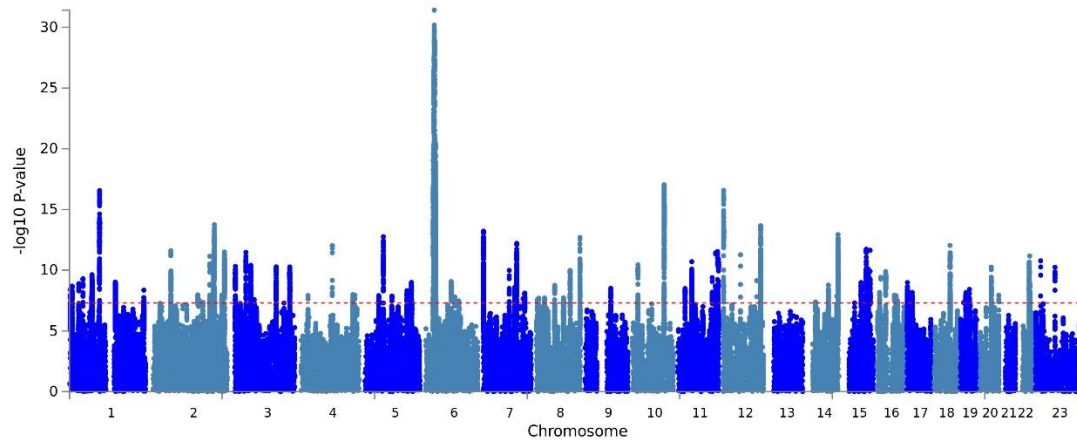
Beginning analysis at Sun Sep 9 20:09:30 2018
Reading summary statistics from
/var/www/ldhub/ldhub_web/static/sumstats/pgc.cross.SCZ17.2013-05.withN.txt.394ab93c-1a72-4d17-9588-e755ca41d6a3.sumstats.gz ...
Read summary statistics for 842651 SNPs.
Reading reference panel LD Score from /root/ldsc/eur_w_ld_chr/[1-22] ...
Read reference panel LD Scores for 1293150 SNPs.
Removing partitioned LD Scores with zero variance.
Reading regression weight LD Score from /root/ldsc/eur_w_ld_chr/[1-22] ...
Read regression weight LD Scores for 1293150 SNPs.
After merging with reference panel LD, 838647 SNPs remain.
After merging with regression SNP LD, 838647 SNPs remain.
Using two-step estimator with cutoff at 30.
Total Observed scale h2: 0.5665 (0.0387)
Lambda GC: 1.2005
Mean Chi^2: 1.2221
Intercept: 1.0074 (0.01)
Ratio: 0.0334 (0.0448)
Analysis finished at Sun Sep 9 20:09:43 2018
Total time elapsed: 12.42s
```

Log file showing the rg estimates along with the standard error and p-values (highlighted in grey):

trait1	trait2	PMID	Category	ethnicity	note	rg	se	z	p	h2_obs	h2_obs_se	h2_int	h2_int_se	gcov_int	gcov_int_se
scz	Age of smoking initiation	20418890	smoking	!European	SNPs fr	0.2428	0.1171	2.0728	0.0382	0.0596	0.0185	1.0004	0.0076	-0.0093	0.0063
scz	Cigarettes smoked per day	20418890	smoking	!European	SNPs fr	0.0171	0.0992	0.1729	0.8628	0.0549	0.0169	1.0101	0.0084	-0.0006	0.0064
scz	Former vs Current smoker	20418890	smoking	!European	SNPs fr	-0.0009	0.0976	-0.0094	0.9925	0.0531	0.0122	1.0117	0.0087	-0.0008	0.0065
scz	Ever vs never smoked	20418890	smoking	!European	SNPs fr	0.1129	0.0585	1.9303	0.0536	0.0714	0.0078	1.0038	0.0083	0.0199	0.0059

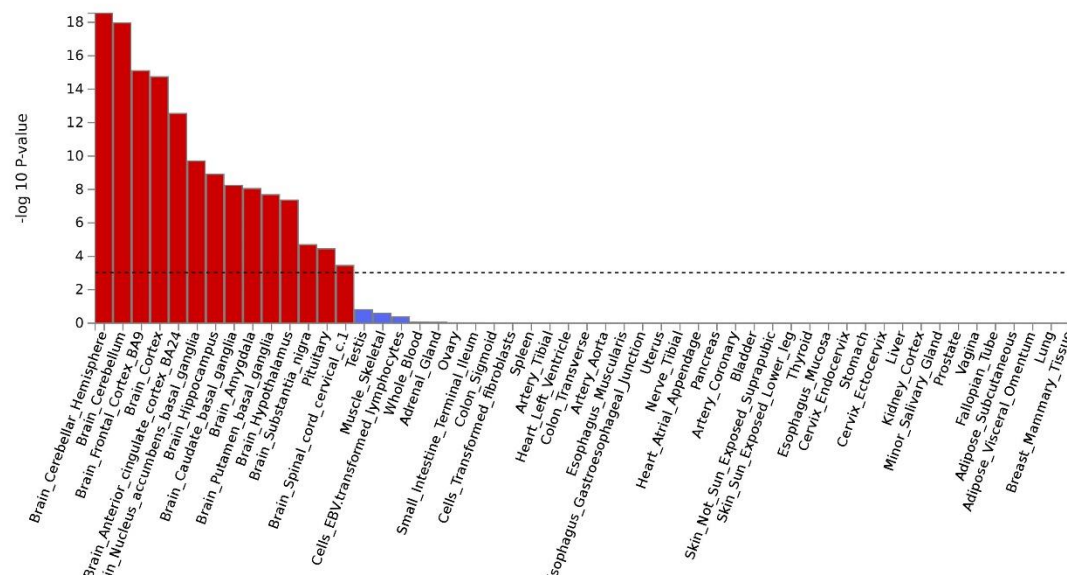
FUMA

#4. Find below the Manhattan plot showing the GWAS results of schizophrenia. There are many genetic variants surpassing the GWAS significant threshold.



#5. We can identify several pathways associated with schizophrenia; these belong to synapse organization and plasticity, and voltage gated calcium channel activity. All of these pathways survive bonferroni correction for multiple testing.

#6. The majority of the genes associated with schizophrenia are expressed in the brain.



#7 and #8. As we can see in the summary of results, there are 129 lead SNPs located in 109 loci. The majority of the annotated candidate SNPs (N=7,304) are intronic.

#10. SNPs associated with schizophrenia have also been associated with multiple other phenotypes, including bipolar disorder, educational attainment, nicotine dependence, among others.