

Getting Started Guide – Pathways Plugin for MedSavant

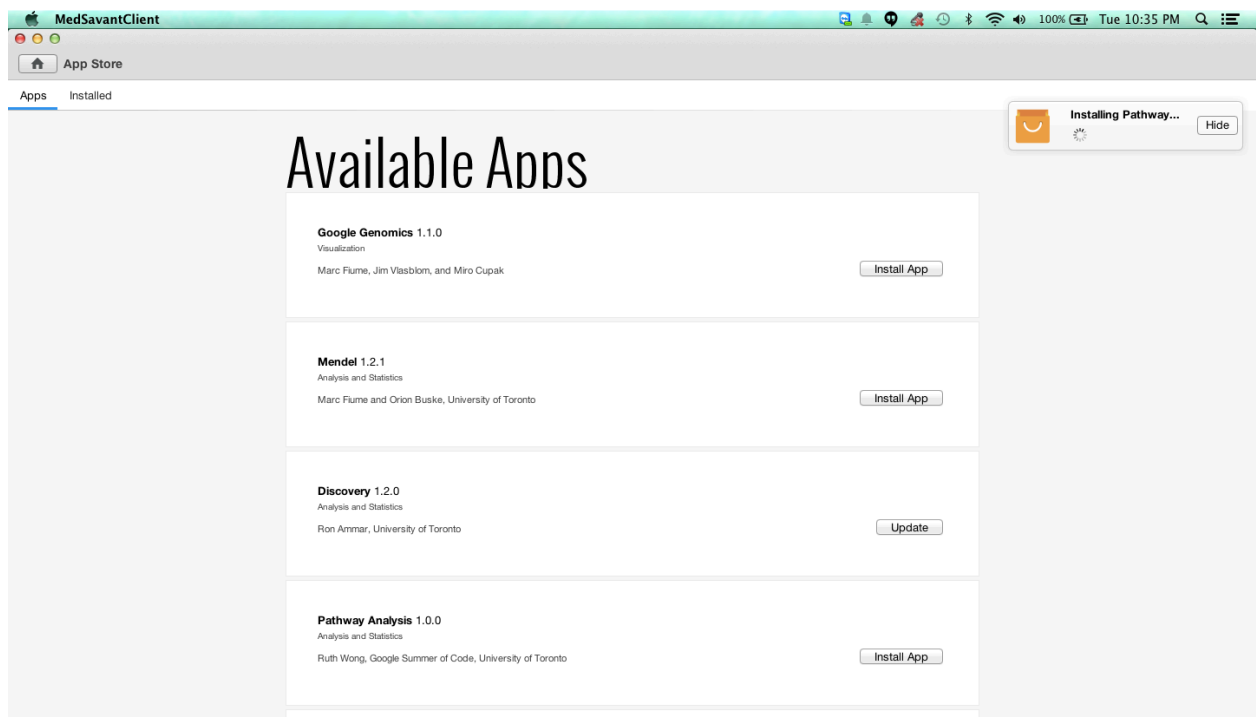
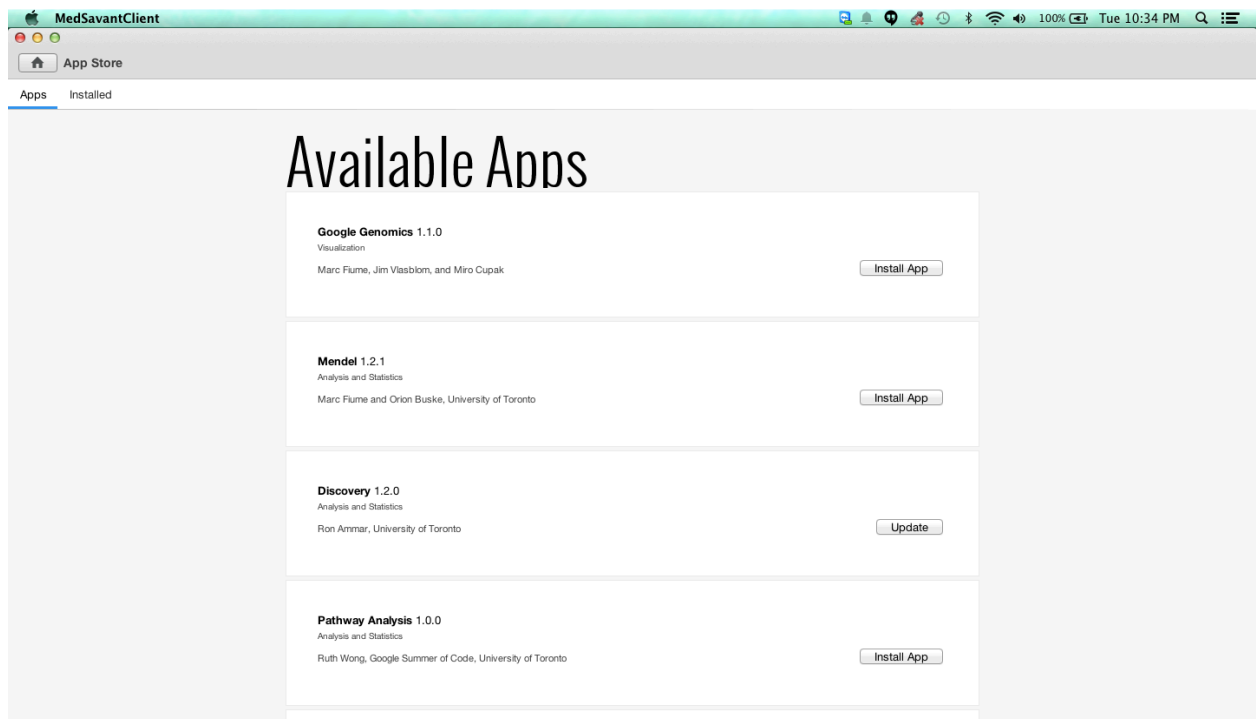
Installation

This guide assumes that you already have the MedSavant client running, and can connect to a MedSavant server. Medsavant can be [downloaded on the MedSavant website](#), instructions for [logging into a server can be found here](#), and if you are a system administrator and would like to set up a server, [instructions can be found here](#).

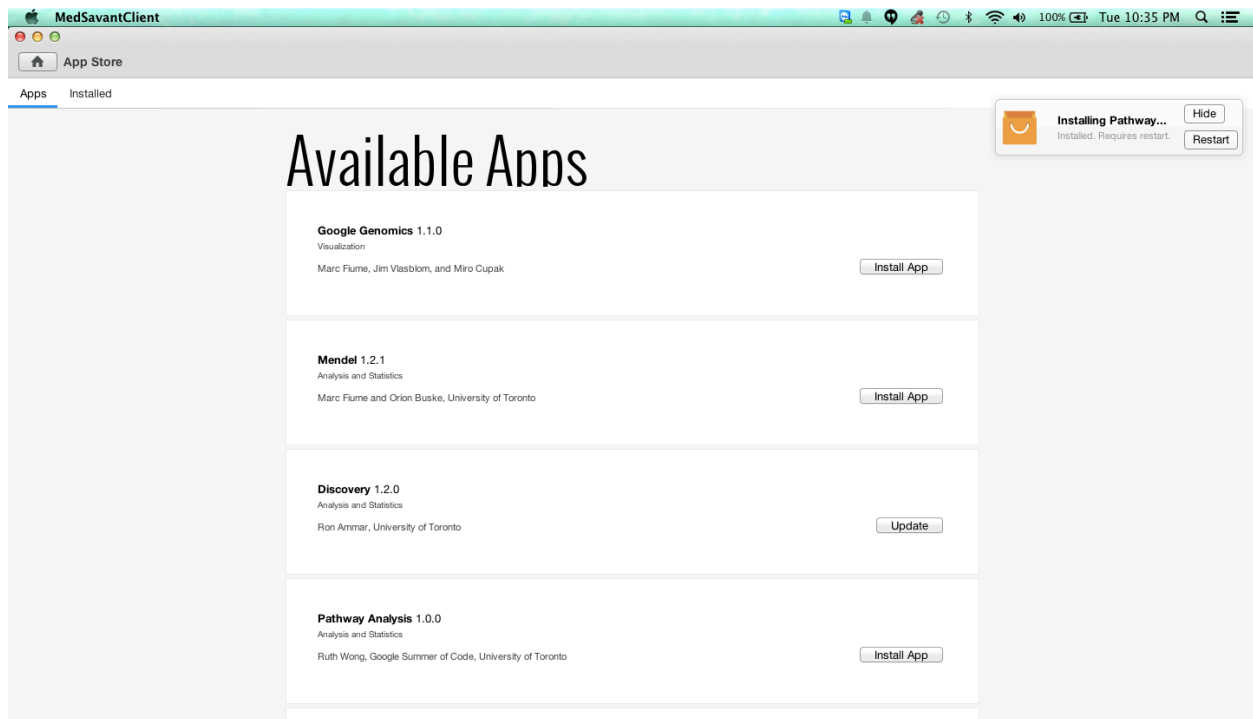
To get the plugin from the app store, log into MedSavant and click on the orange app store icon.



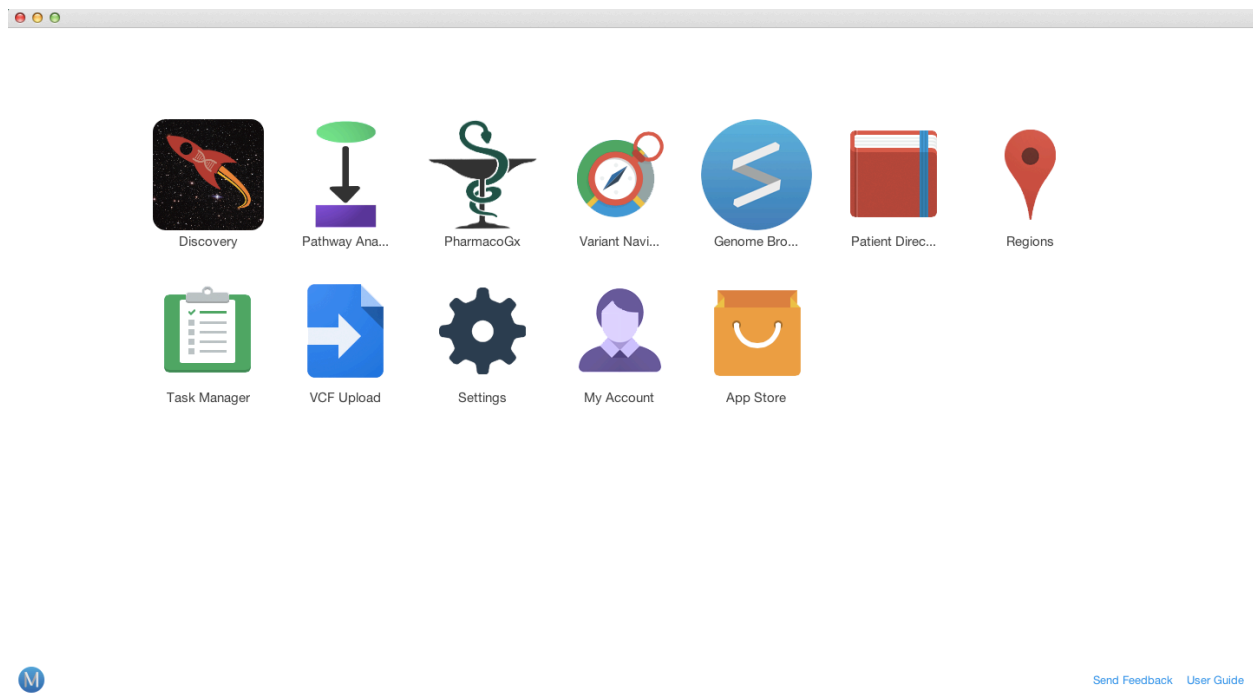
Find the app titled Pathway Analysis, and click Install.



After the app is done installing, you will be prompted to restart MedSavant.



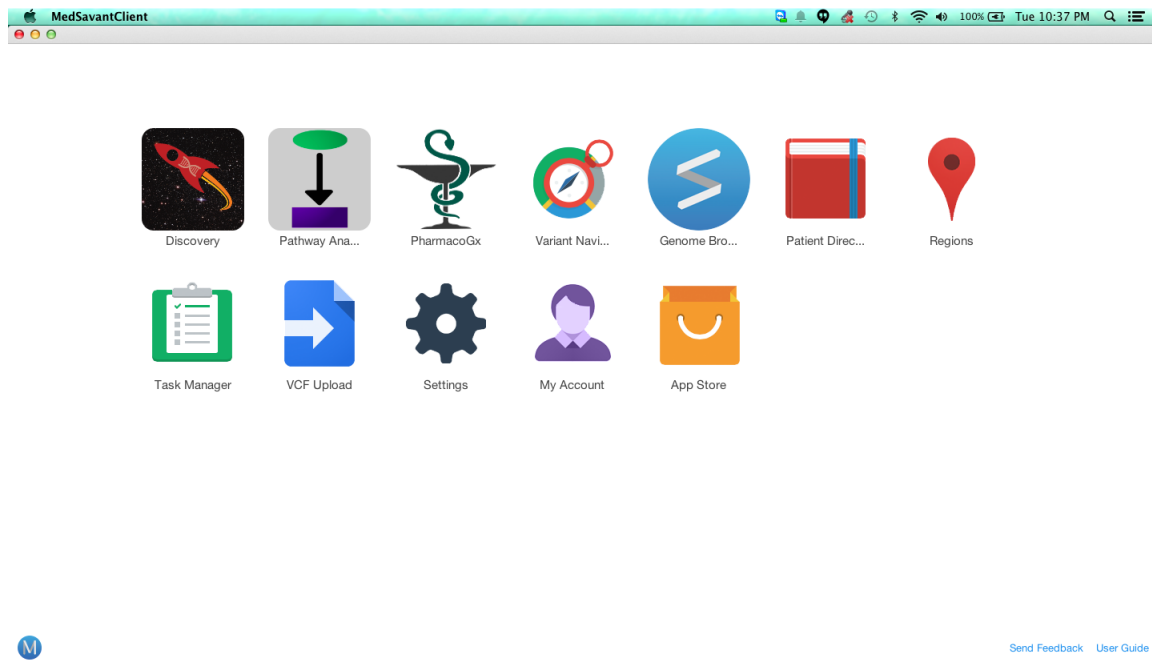
On restart, the app should appear in your home page.



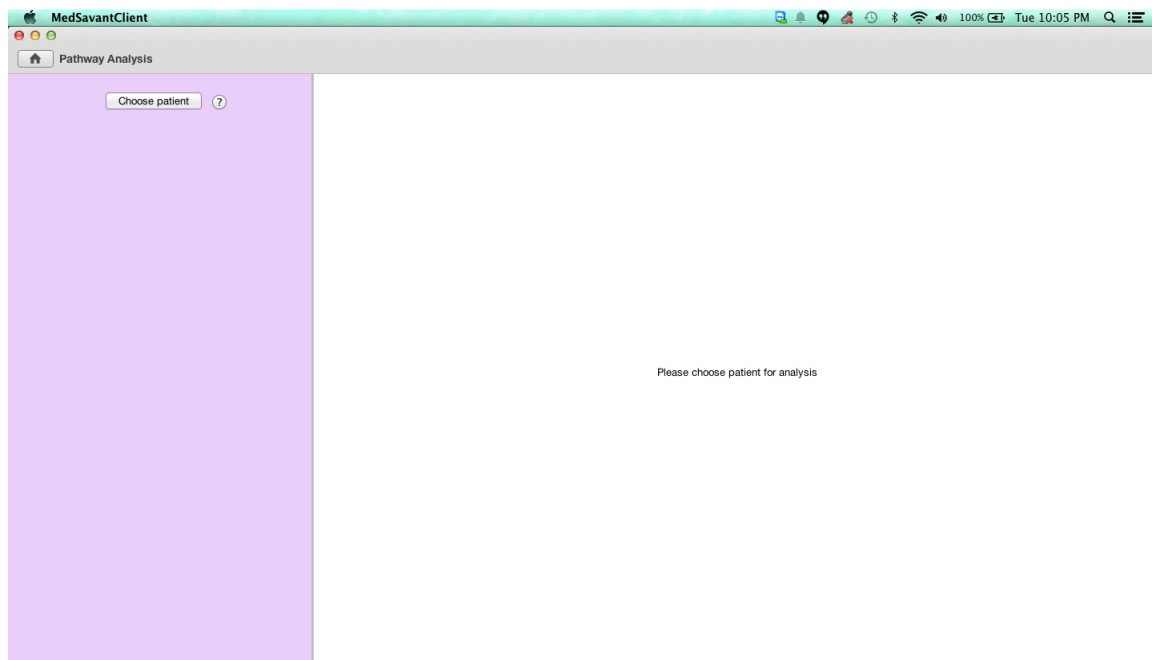
Alternatively, you can build the app from source. Build instructions can be found at <https://github.com/compbio-UofT/pathways-app>.

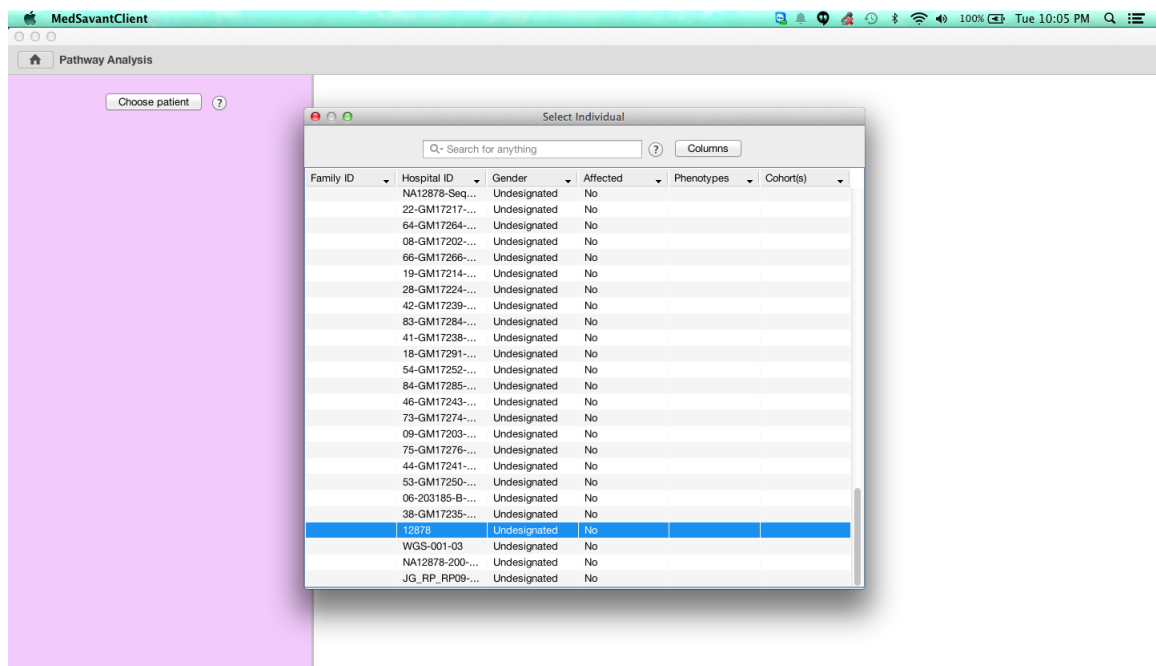
Quick Start

Click the Pathway App icon on the home page.

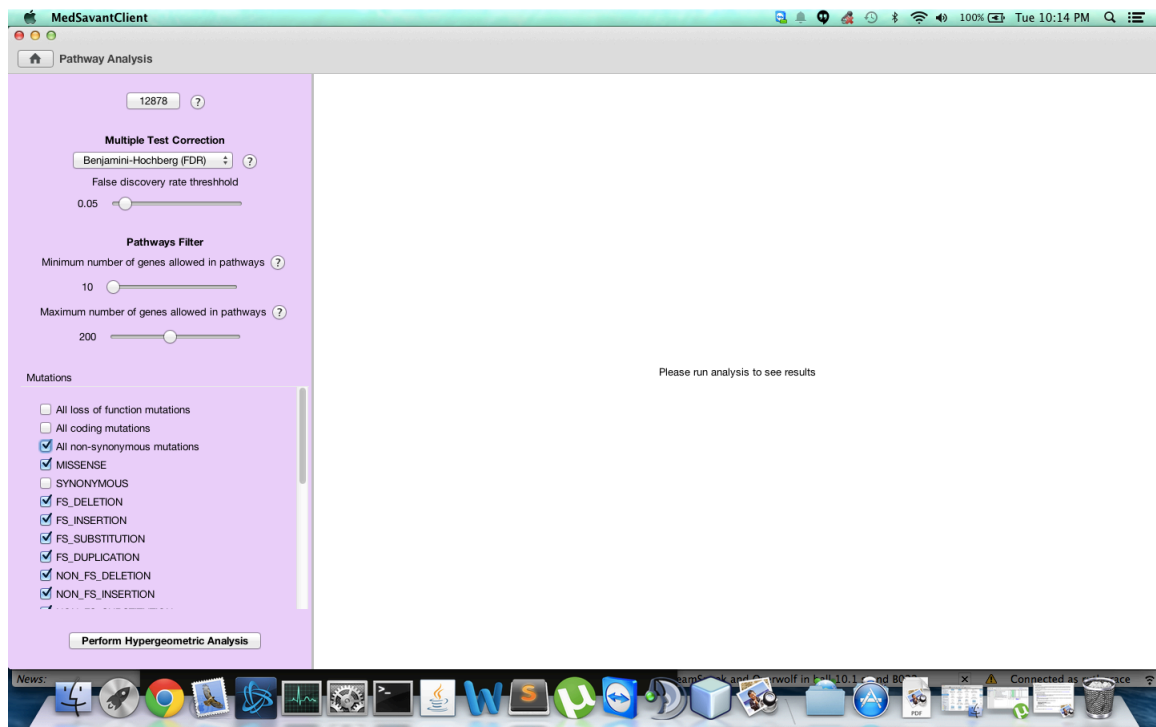


Choose a patient from the database. If you do not have any data in your database, you may upload Variant Call Files with the VCF Uploader app ([instructions can be found here](#)).





Adjust the analysis parameters and click Perform Hypergeometric Analysis.



MedSavantClient

100% Tue 10:06 PM

Pathway Analysis

12878

?

Multiple Test Correction

Benjamini-Hochberg (FDR)

False discovery rate threshold

0.05

Pathways Filter

Minimum number of genes allowed in pathways

10

Maximum number of genes allowed in pathways

200

Mutations

☐ All loss of function mutations

☐ All coding mutations

☐ All non-synonymous mutations

☒ MISSENSE

☐ SYNONYMOUS

☒ FS_DELETION

☒ FS_INSERTION

☒ FS_SUBSTITUTION

☒ FS_DUPLICATION

☒ NON_FS_DELETION

☒ NON_FS_INSERTION

Perform Hypergeometric Analysis

Analyzing pathway enrichment...

Your results will be shown.

MedSavantClient

100% Tue 10:16 PM

Pathway Analysis

12878

?

Multiple Test Correction

Benjamini-Hochberg (FDR)

False discovery rate threshold

0.05

Pathways Filter

Minimum number of genes allowed in pathways

10

Maximum number of genes allowed in pathways

200

Mutations

☐ All loss of function mutations

☐ All coding mutations

☒ All non-synonymous mutations

☒ MISSENSE

☐ SYNONYMOUS

☒ FS_DELETION

☒ FS_INSERTION

☒ FS_SUBSTITUTION

☒ FS_DUPLICATION

☒ NON_FS_DELETION

☒ NON_FS_INSERTION

Perform Hypergeometric Analysis

Results

About

Q-values

Pathway Name

Genes associated with patient variants in pathway

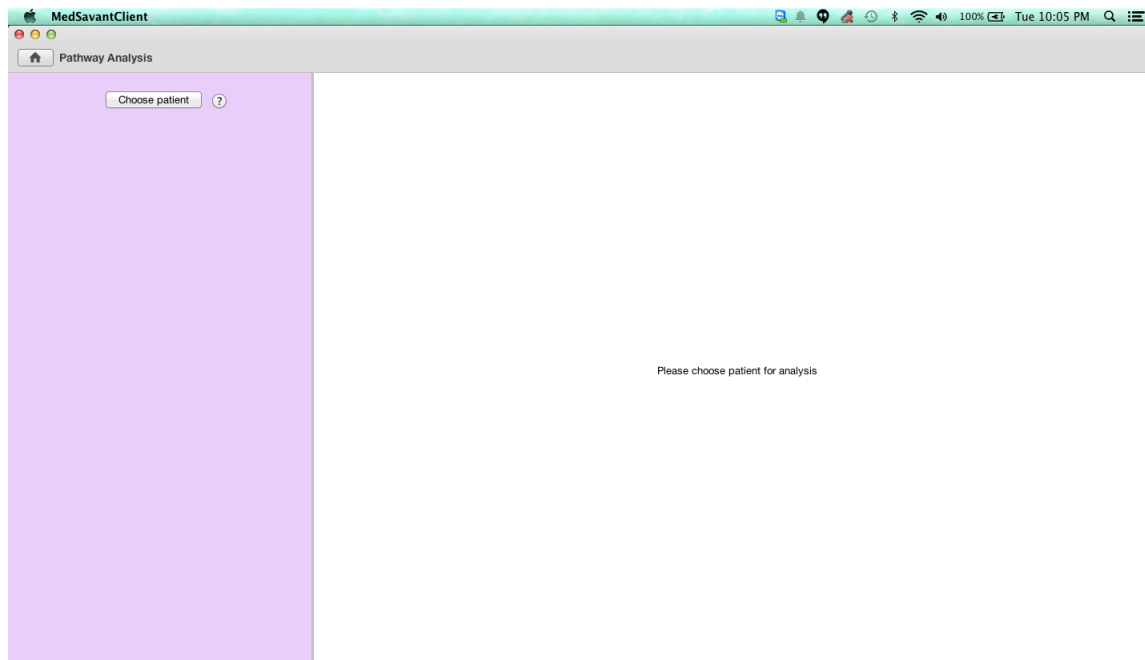
Genes in pathway

2.4848E-2	Allograft Rejection	HLA-DOB1, C7, C9, MICA, HLA-DRB1, C6, C5, IL13, HLA-DMB, HLA-DMA, TGFB...	86
1.3596E-2	Matrix Metalloproteinases	MMP9, MMP8, MMP27, MMP7, MMP17, MMP16, MMP28, MMP14, MMP3, MMP1...	31
1.6175E-2	Integrin-mediated Cell Adhesion	CAV2, ITGAL, TLN1, ITGAE, ITGB4, ITGA11, ITGB5, ITGB2, PXN, PTK2, DOCK1, ...	103
1.7345E-2	Statin Pathway	PDIA2, ABCA1, MTTP, FDFT1, ABCG8, APOA4, APOB, ABCG5, DGAT1, LRP1, AP...	30
1.5293E-2	Metapathway biotransformation	CYP3A7, KCNA82, CYP2C18, CYP11B2, CYP2D6, CYP2W1, GLYATL2, AKR1C3, ...	189
2.2373E-2	Complement and Coagulation Cascades	KNG1, C7, F12, CR1, A2M, CR2, C9, MASP2, C6, PLG, PLAUR, C8G, F13B, VWF...	61
2.5583E-2	Iron uptake and transport	TF, FXSD2, CLCN3, FXSD3, HTR3E, SLC11A2, UNC79, SLC9B1, SLC9B2, MOOL...	113
4.4825E-2	Fatty Acid Beta Oxidation	GPD2, GCDH, CPT1B, CPT2, ACADS, ECHS1, PNPLA2, CRAT, ACADL, HADHA, H...	34
4.8466E-2	Blood Clotting Cascade	F13B, VWF, F12, F5, FGB, KLKB1, SERPINE1, KLK1, PLAU, PLG	22

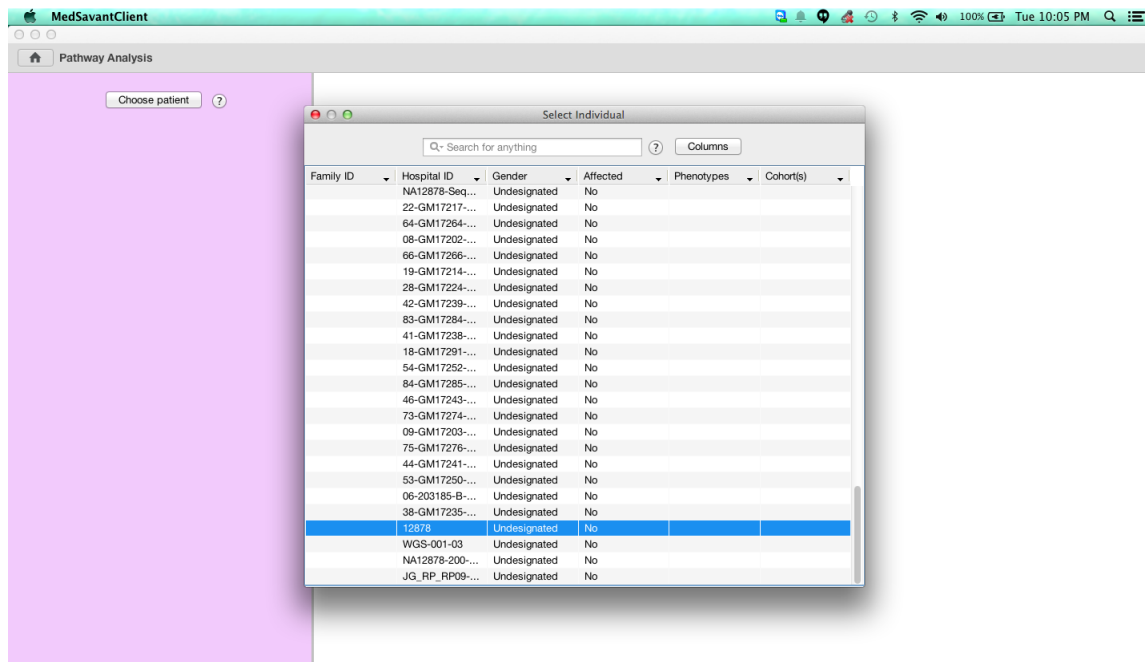
Please select pathway from table to see detailed pathway information.

Detailed Guide

Choose a patient from the database. If you do not have any data in your database, you may upload Variant Call Files with the VCF Uploader app ([instructions can be found here](#)).



The "choose patient" button opens a pop-up window. Select a patient. Here we have chosen patient 12878, from whom we have variants derived from whole exome data.



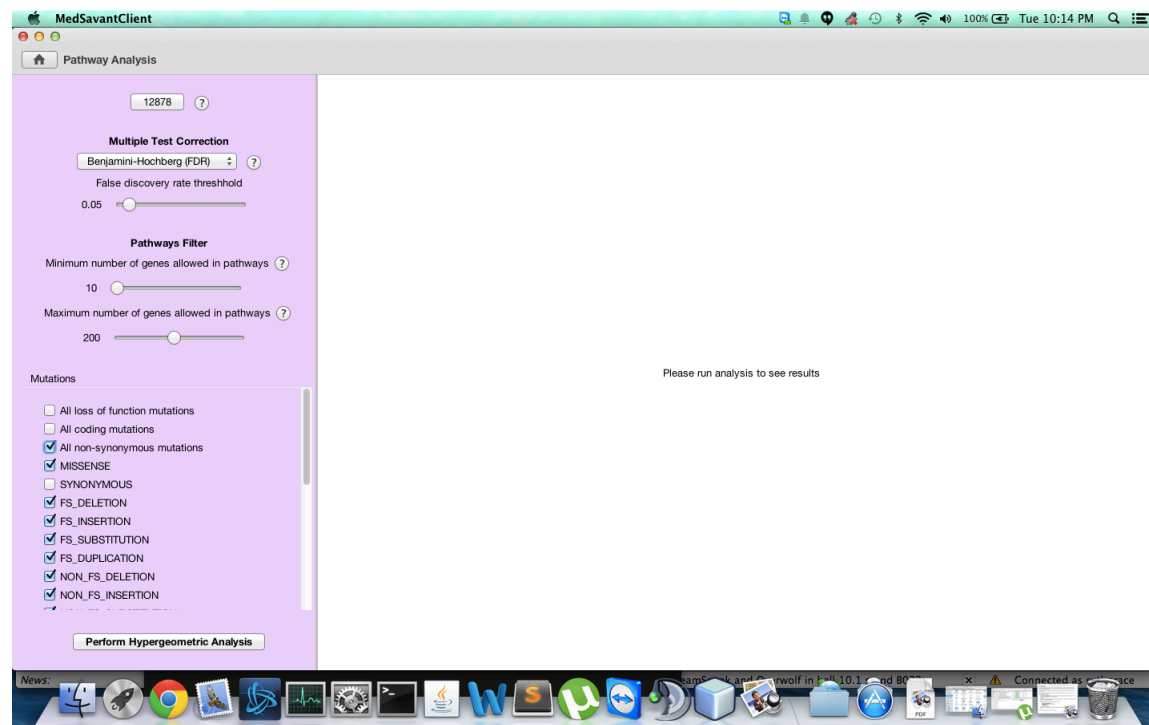
Once a patient is chosen, you have the option to change some analysis parameters: Multiple test correction type, the maximum and minimum number of genes allowed in pathways used for analysis, and the types of mutations you would like to analyze.

The multiple test corrections available are the Bonferroni correction, which is very conservative, and the Benjamini-Hochberg correction, which is less conservative, and based on a False Discovery Rate cutoff which you can adjust. The default FDR cutoff is 0.05.

There are also options for filtering which pathways should be used in the analysis. The standard cutoffs allow pathways with at least 10 genes and at most 200 genes. Pathways with too many genes are disproportionately likely to contain many genes from the sample, while pathways with too few genes are disproportionately likely to be called significant if they only contain one gene from the sample.

The mutation types available are from the Jannovar standard set. Variant call files are annotated by Jannovar after upload into the database.

Here we have selected all non-synonymous mutations, and left the other options at their default values.



The analysis may take up to a minute to run, depending on how many genes are associated with your variants.

The screenshot shows the MedSavantClient Pathway Analysis interface. On the left, there are configuration options: a sample size of 12878, a Multiple Test Correction method set to Benjamini-Hochberg (FDR) with a False discovery rate threshold of 0.05, and Pathways Filter settings for minimum (10) and maximum (200) genes allowed. Under Mutations, several options are checked, including All non-synonymous mutations, Missense, and various frameshift and non-synonymous mutation types. A 'Perform Hypergeometric Analysis' button is at the bottom. The main area on the right displays 'Analyzing pathway enrichment...' with a loading spinner.

Once the analysis is complete, a table of results will be displayed.

The screenshot shows the MedSavantClient Pathway Analysis interface after the analysis is complete. The left sidebar remains the same. The main area on the right now displays a table of results under the 'Results' tab. The table has four columns: Q-values, Pathway Name, Genes associated with patient variants in pathway, and Genes in pathway. The table lists several pathways, including Allograft Rejection, Matrix Metalloproteinases, Integrin-mediated Cell Adhesion, Statin Pathway, Metapathway biotransformation, Complement and Coagulation Cascades, Iron uptake and transport, Fatty Acid Beta Oxidation, and Blood Clotting Cascade. A message at the bottom states: 'Please select pathway from table to see detailed pathway information.'

Q-values	Pathway Name	Genes associated with patient variants in pathway	Genes in pathway
2.4848E-2	Allograft Rejection	HLA-DOB1, C7, C9, MICA, HLA-DRB1, C6, C5, IL13, HLA-DMB, HLA-DMA, TGFB...	86
1.3596E-2	Matrix Metalloproteinases	MMP9, MMP8, MMP27, MMP7, MMP17, MMP16, MMP28, MMP14, MMP3, MMP1...	31
1.6175E-2	Integrin-mediated Cell Adhesion	CAV2, ITGAL, TLN1, ITGAE, ITGB4, ITGA11, ITGB5, ITGB2, PXN, PTK2, DOCK1, ...	103
1.7345E-2	Statin Pathway	PDIA2, ABCA1, MTTP, FDFT1, ABCG8, APOA4, APOB, ABCG5, DGAT1, LRP1, AP...	30
1.5293E-2	Metapathway biotransformation	CYP3A7, KCNA82, CYP2C18, CYP11B2, CYP2D6, CYP2W1, GLYATL2, AKR1C3, ...	189
2.2373E-2	Complement and Coagulation Cascades	KNG1, C7, F12, CR1, A2M, CR2, C9, MASP2, C6, PLG, PLAUR, C8G, F13B, VWF...	61
2.5583E-2	Iron uptake and transport	TF, FXD2, CLCN3, FXD3, HTR3E, SLC11A2, UNC79, SLC9B1, SLC9B2, MOOL...	113
4.4825E-2	Fatty Acid Beta Oxidation	GPD2, GCDH, CPT1B, CPT2, ACADS, ECHS1, PNPLA2, CRAT, ACADL, HADHA, H...	34
4.8466E-2	Blood Clotting Cascade	F13B, VWF, F12, F5, FGB, KLKB1, SERPINE1, KLK1, PLAU, PLG	22

You can select a row of the table to see more details about the pathway.

MedSavantClient Pathway Analysis

12878 ?

Multiple Test Correction
Benjamini-Hochberg (FDR) ?
False discovery rate threshold: 0.05

Pathways Filter
Minimum number of genes allowed in pathways ? 10
Maximum number of genes allowed in pathways ? 200

Mutations

- ☐ All loss of function mutations
- ☐ All coding mutations
- ☒ All non-synonymous mutations
- ☒ MISSENSE
- ☐ SYNONYMOUS
- ☒ FS_DELETION
- ☒ FS_INSERTION
- ☒ FS_SUBSTITUTION
- ☒ FS_DUPLICATION
- ☒ NON_FS_DELETION
- ☒ NON_FS_INSERTION

Perform Hypergeometric Analysis

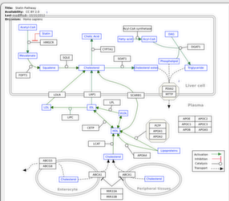
Results About

Search for anything Columns Export

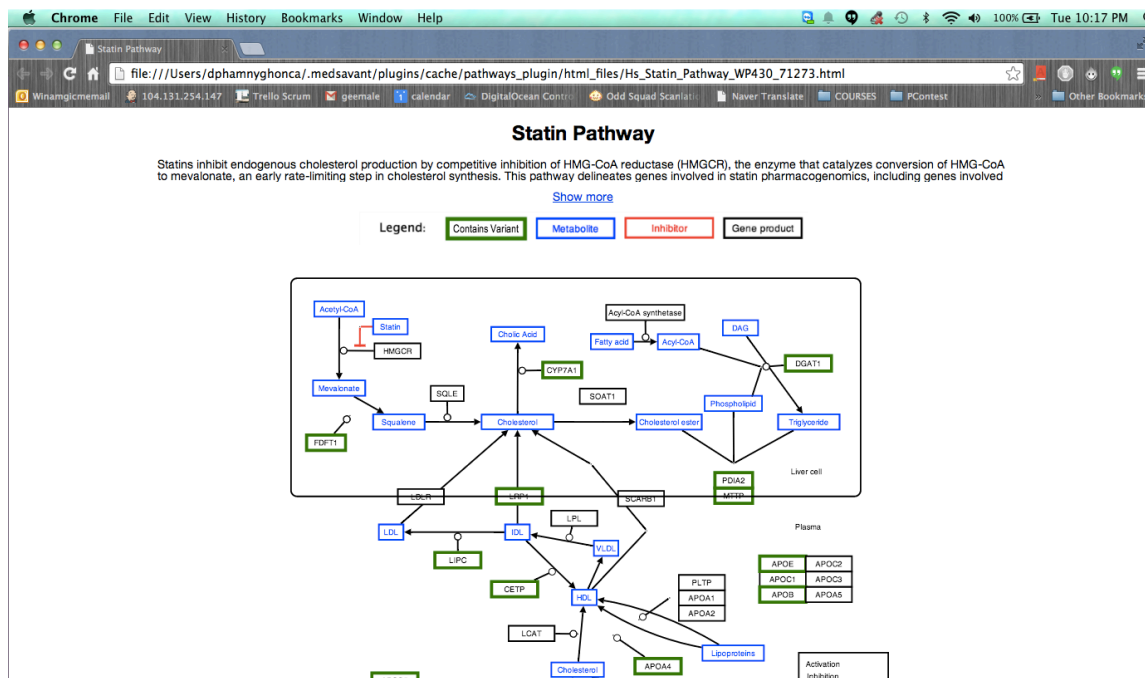
Q-values	Pathway Name	Genes associated with patient variants in pathway	Genes in pathway
2.4848E-2	Allograft Rejection	HLA-DQB1, C7, C9, MICA, HLA-DRB1, C6, C5, IL13, HLA-DMB, HLA-DMA, TGFB...	86
1.3596E-2	Matrix Metalloproteinases	MMP9, MMP8, MMP27, MMP7, MMP17, MMP16, MMP28, MMP14, MMP3, MMP...	31
1.6175E-2	Integrin-mediated Cell Adhesion	CAV2, ITGAL, TLN1, ITGAE, ITGB4, ITGA11, ITGB5, ITGB2, PXN, PTK2, DOCK1, ...	103
1.7345E-2	Statin Pathway	PDIA2, ABCA1, MTTIP, FDF11, ABCG8, APOA4, APOB, ABCG5, DGAT1, LRPI, AP...	30
1.5293E-2	Metapathway biotransformation	CYP3A7, KCNAB2, CYP2C18, CYP11B2, CYP2D6, CYP2W1, GLYATL2, AKR1C3, ...	189
2.2373E-2	Complement and Coagulation Cascades	KNG1, C7, F12, CR1, A2M, CR2, C9, MASP2, C6, PLG, PLAUR, C8G, F13B, VWF...	61
2.5583E-2	Iron uptake and transport	TF, FXVD2, CLCN3, FXVD3, HTR3E, SLC11A2, UNC79, SLC9B1, SLC9B2, MCOL...	113
4.4825E-2	Fatty Acid Beta Oxidation	GPD2, GCDH, CPT1B, CPT2, ACADS, ECHS1, PNPLA2, CRAT, ACADL, HADHA, H...	34
4.8466E-2	Blood Clotting Cascade	F13B, VWF, F12, F5, FGB, KLKB1, SERPINE1, KLK1, PLAUI, PLG	22

Statin Pathway
Statins inhibit endogenous cholesterol production by competitive inhibition of HMG-CoA

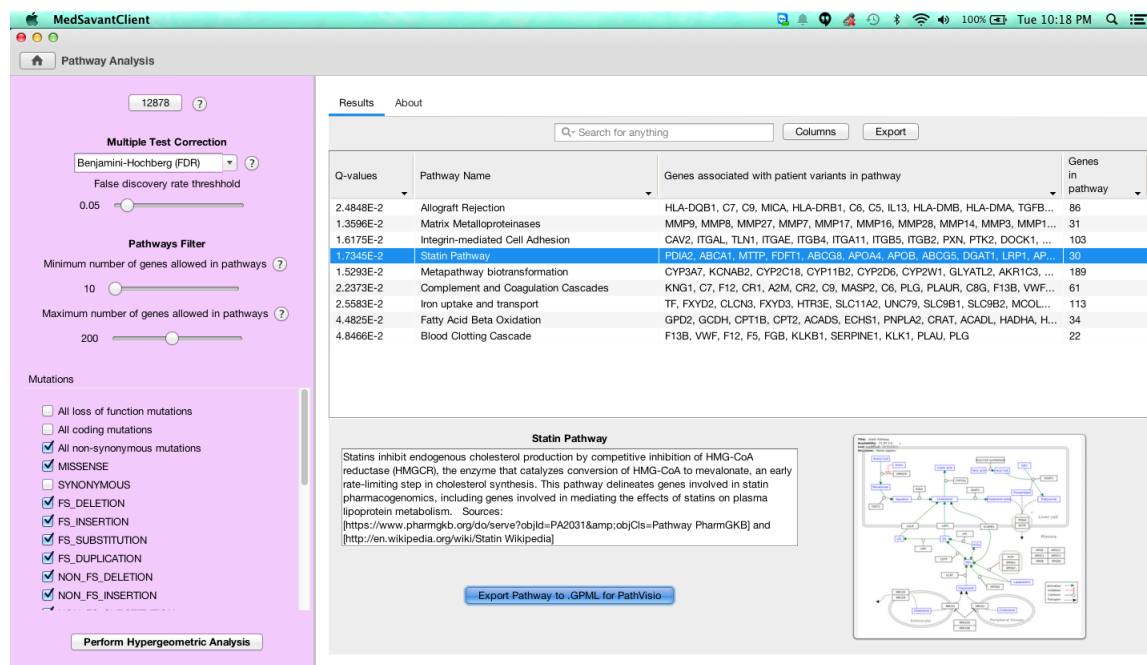
Export Pathway to .GPML for PathVisio



Clicking on the pathway thumbnail opens a Cytoscape.js view of the pathway in your default browser, which you can manipulate.



You can also export the pathway into GPML format for visualization in PathVisio.



The screenshot shows the MedSavantClient Pathway Analysis interface. On the left, the 'Multiple Test Correction' section is set to 'Benjamini-Hochberg (FDR)' with a 'False discovery rate threshold' of 0.05. The 'Pathways Filter' section shows a 'Minimum number of genes allowed in pathways' of 10 and a 'Maximum number of genes allowed in pathways' of 200. The 'Mutations' section has several checkboxes, including 'All non-synonymous mutations', 'MISSENSE', 'SYNONYMOUS', 'FS_DELETION', 'FS_INSERTION', 'FS_SUBSTITUTION', 'FS_DUPLICATION', 'NON_FS_DELETION', and 'NON_FS_INSERTION'. The 'Perform Hypergeometric Analysis' button is at the bottom of the left panel.

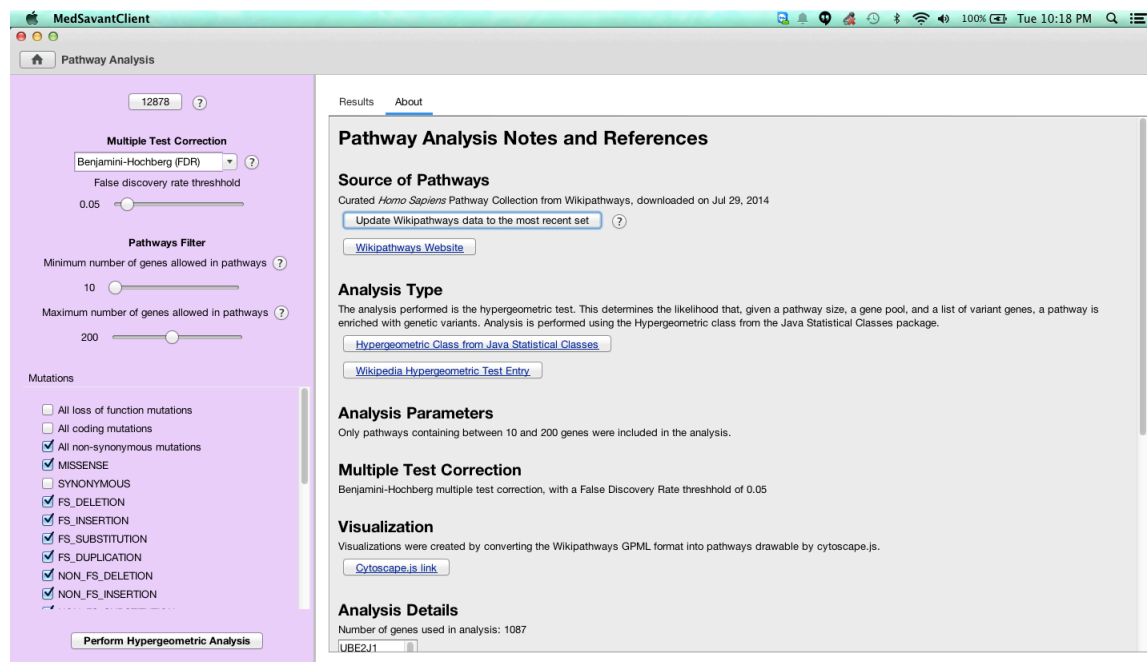
The 'Results' tab is active, displaying a table of pathways. The table has four columns: 'Q-values', 'Pathway Name', 'Genes associated with patient variants in pathway', and 'Genes in pathway'. The 'Statin Pathway' is highlighted in blue.

Q-values	Pathway Name	Genes associated with patient variants in pathway	Genes in pathway
2.4848E-2	Allograft Rejection	HLA-DQB1, C7, C9, MICA, HLA-DRB1, C6, C5, IL13, HLA-DMB, HLA-DMA, TGFB...	86
1.3596E-2	Matrix Metalloproteinases	MMP9, MMP8, MMP27, MMP7, MMP17, MMP16, MMP28, MMP14, MMP3, MMP1...	31
1.6175E-2	Integrin-mediated Cell Adhesion	CAV2, ITGAL, TLN1, ITGAE, ITGB4, ITGA11, ITGB5, ITGB2, PXN, PTK2, DOCK1, ...	103
1.7345E-2	Statin Pathway	PDIA2, ABCA1, MTP, FDF1, ABCG8, APOA4, APOB, ABCG5, DGAT1, LRP1, AP...	30
1.5293E-2	Metapathway biotransformation	CYP3A7, KCNA2, CYP2C18, CYP11B2, CYP2D6, CYP2W1, GLYATL2, AKR1C3, ...	189
2.2373E-2	Complement and Coagulation Cascades	KNG1, C7, F12, CR1, A2M, CR2, C9, MASP2, C6, PLG, PLAUR, C8, F13B, VWF...	61
2.5583E-2	Iron uptake and transport	TF, FXVD2, CLCN3, FXVD3, HTR3E, SLC11A2, UNC79, SLC9B1, SLC9B2, MCOL...	113
4.4825E-2	Fatty Acid Beta Oxidation	GPD2, GCDH, CPT1B, CPT2, ACADS, ECHS1, PNPLA2, CRAT, ACADL, HADHA, H...	34
4.8466E-2	Blood Clotting Cascade	F13B, VWF, F12, F5, FGB, KLKB1, SERPINE1, KLK1, PLAUI, PLG	22

Below the table, the 'Statin Pathway' is detailed. It states: 'Statins inhibit endogenous cholesterol production by competitive inhibition of HMG-CoA reductase (HMGCR), the enzyme that catalyzes conversion of HMG-CoA to mevalonate, an early rate-limiting step in cholesterol synthesis. This pathway delineates genes involved in statin pharmacogenomics, including genes involved in mediating the effects of statins on plasma lipoprotein metabolism. Sources: [https://www.pharmgkb.org/doserve?objid=PA2031&objCls=Pathway PharmGKB] and [http://en.wikipedia.org/wiki/Statin Wikipedia]'. A button 'Export Pathway to .GPML for PathVisio' is located below the text.

On the right, a diagram of the Statin Pathway is shown, illustrating the metabolic pathway from cholesterol to mevalonate and the subsequent steps in cholesterol synthesis.

Lastly, there is an About tab with information about the analysis.



The screenshot shows the MedSavantClient Pathway Analysis interface, specifically the 'About' tab. The left panel is identical to the 'Results' tab, showing the same settings for 'Multiple Test Correction', 'Pathways Filter', and 'Mutations'.

The 'About' tab is titled 'Pathway Analysis Notes and References'. It contains the following sections:

- Source of Pathways**: Curated *Homo Sapiens* Pathway Collection from Wikipathways, downloaded on Jul 29, 2014. A button 'Update Wikipathways data to the most recent set' is available, along with a link to the 'Wikipathways Website'.
- Analysis Type**: The analysis performed is the hypergeometric test. This determines the likelihood that, given a pathway size, a gene pool, and a list of variant genes, a pathway is enriched with genetic variants. Analysis is performed using the Hypergeometric class from the Java Statistical Classes package. Links to 'Hypergeometric Class from Java Statistical Classes' and 'Wikipedia Hypergeometric Test Entry' are provided.
- Analysis Parameters**: Only pathways containing between 10 and 200 genes were included in the analysis.
- Multiple Test Correction**: Benjamini-Hochberg multiple test correction, with a False Discovery Rate threshold of 0.05.
- Visualization**: Visualizations were created by converting the Wikipathways GPML format into pathways drawable by cytoscape.js. A link to 'Cytoscape.js link' is provided.
- Analysis Details**: Number of genes used in analysis: 1087. A button 'UBE2J1' is shown.

Here you can see which genes were not included in the analysis because they were not found in any pathways.

The screenshot shows the MedSavantClient Pathway Analysis interface. The left sidebar contains settings for the analysis, including a sample size of 12878, a Benjamini-Hochberg (FDR) multiple test correction method with a false discovery rate threshold of 0.05, and a pathways filter with a minimum of 10 and a maximum of 200 genes allowed. The mutations section is checked for all loss of function, all coding, all non-synonymous, missense, synonymous, frameshift deletion, frameshift insertion, frameshift substitution, frameshift duplication, non-frameshift deletion, and non-frameshift insertion. A 'Perform Hypergeometric Analysis' button is at the bottom of the sidebar. The main panel shows the results, including a visualization of pathways created by converting Wikipathways GPML format into pathways drawable by cytoscape.js, a link to the Cytoscape.js link, and analysis details. The analysis details section shows the number of genes used in the analysis (1087) and the number of genes not found in the Wikipathways gene set (4749). It also lists the number of pathways in the Wikipathways gene set (260) and the number of pathways in the Wikipathways gene set with associated genetic variants in this analysis (205). The pathways listed include Adipogenesis, Ovarian Infertility Genes, RANKL/RANK Signaling Pathway, Parkin-Ubiquitin Proteasomal System pathway, Endothelin Pathways, Focal Adhesion, Allograft Rejection, Matrix Metalloproteinases, Integrin-mediated Cell Adhesion, and Statin Pathway.

For more tutorials, see the [MedSavant website](#). For more information about the Pathways plugin, see the [GitHub repository](#).

This project was funded by Google Summer of Code, in conjunction with the Sick Kids Center for Computational Medicine, and the Bader Lab at the University of Toronto.