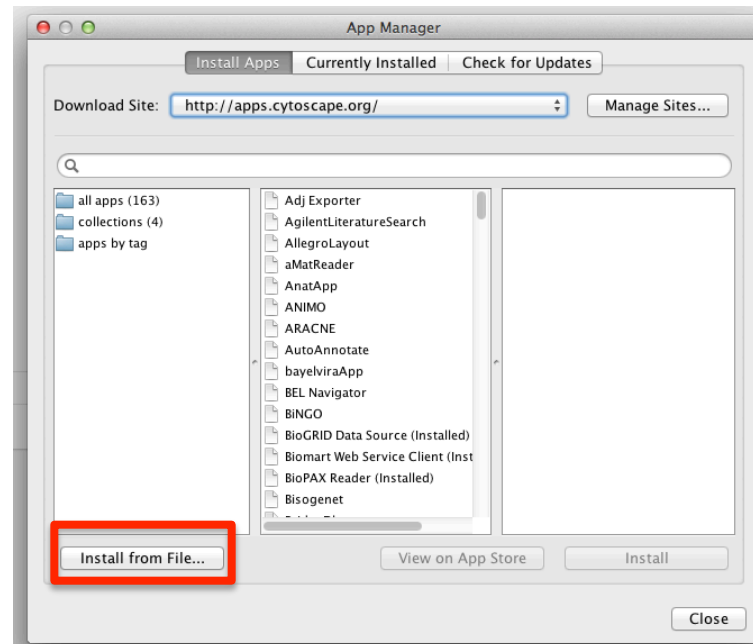
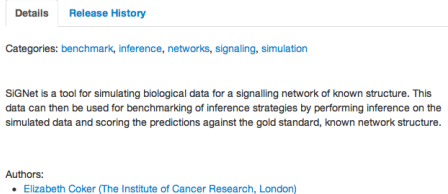
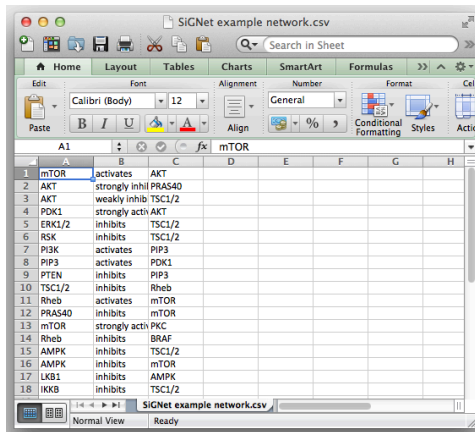


SiGNet Benchmarking Example

SiGNet is a tool for simulating biological data for a signalling network of known structure. This data can then be used for benchmarking of inference strategies by performing inference on the simulated data and scoring the predictions against the gold standard, known network structure. Here we illustrate how this can be achieved.

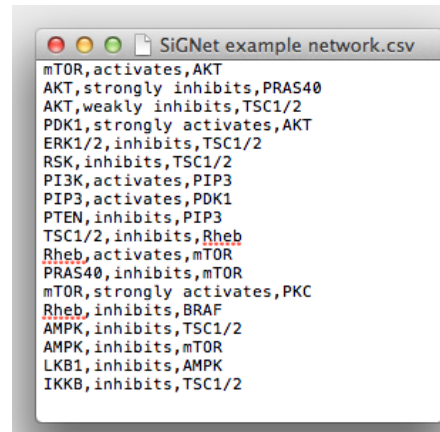


1. Download the SiGNet plugin from either the Cytoscape App store or from <http://signet.icr.ac.uk>. Open Cytoscape version 3.2+ and select 'Apps' then 'App Manager' from the Cytoscape toolbar. In the App Manager click 'Install from File' and select the file you have just downloaded.



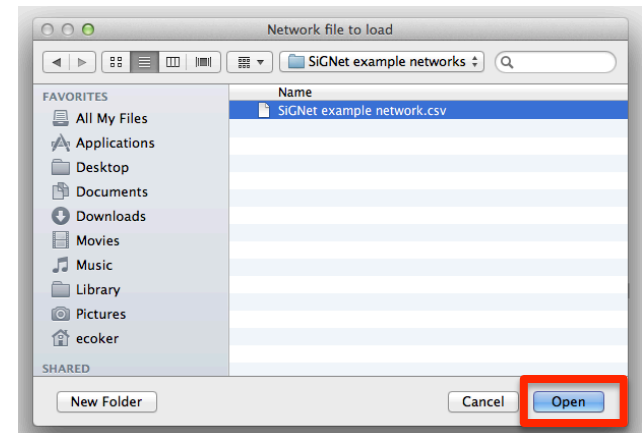
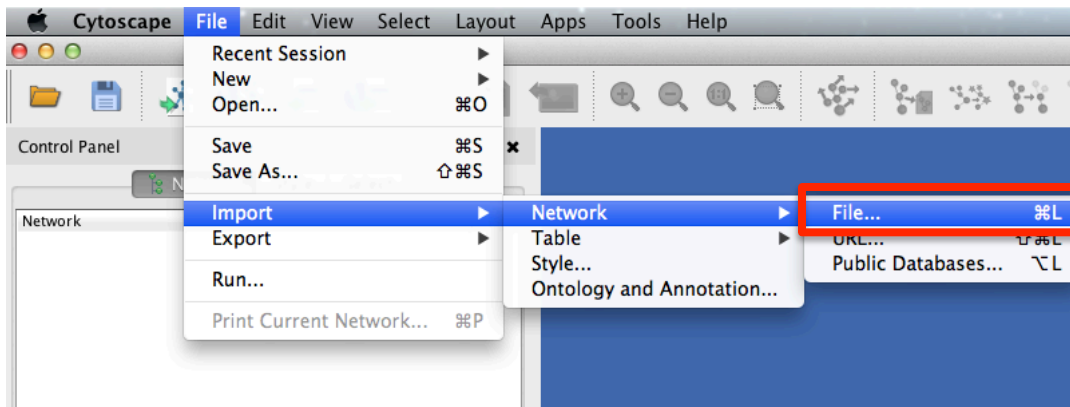
SiGNet example network.csv

A1	B	C	D	E	F	G	H
1	mTOR	activates	AKT				
2	AKT	strongly inhibits	PRAS40				
3	AKT	weakly inhibits	TSC1/2				
4	PI3K	strongly activates	AKT				
5	ERK1/2	inhibits	TSC1/2				
6	RSK	inhibits	TSC1/2				
7	PI3K	activates	PIP3				
8	PIP3	activates	PDK1				
9	PTEN	inhibits	PIP3				
10	TSC1/2	inhibits	Rheb				
11	Rheb	activates	mTOR				
12	PRAS40	inhibits	mTOR				
13	mTOR	strongly activates	PKC				
14	Rheb	inhibits	BRAF				
15	AMPK	inhibits	TSC1/2				
16	AMPK	inhibits	mTOR				
17	LKB1	inhibits	AMPK				
18	IKKB	inhibits	TSC1/2				



SiGNet example network.csv

```
mTOR,activates,AKT
AKT,strongly inhibits,PRAS40
AKT,weakly inhibits,TSC1/2
PI3K,strongly activates,AKT
ERK1/2,inhibits,TSC1/2
RSK,inhibits,TSC1/2
PI3K,activates,PIP3
PIP3,activates,PDK1
PTEN,inhibits,PIP3
TSC1/2,inhibits,Rheb
Rheb,activates,mTOR
PRAS40,inhibits,mTOR
mTOR,strongly activates,PKC
Rheb,inhibits,BRAF
AMPK,inhibits,TSC1/2
AMPK,inhibits,mTOR
LKB1,inhibits,AMPK
IKKB,inhibits,TSC1/2
```



2. Create a file of interactions to be present in the network SiGNet will use for simulation. Edges should be described using one of the following terms: 'activates', 'strongly activates', 'weakly activates', 'inhibits', 'strongly inhibits', 'weakly inhibits' or 'binds' (for interactions that do not affect the activity of the target node). If the interaction is left blank or does not match this controlled vocabulary, SiGNet will later replace the interaction with 'activates' and warn the user that this has occurred. In Cytoscape, select 'File', 'Import', 'Network' and then 'File', and select the file you have just created.

Import Network From Table

Select a Network Collection

Network Collection: Create new network collection

Node Identifier Mapping Column: shared name

Interaction Definition

Source Interaction: Column 1

Interaction Type: Column 2

Target Interaction: Column 3

Columns in BLUE will be loaded as EDGE ATTRIBUTES.

Advanced

☒ Show Text File Import Options

Text File Import Options

Delimiter: ☐ Tab ☒ Comma ☐ Semicolon ☐ Space ☐ Other

Preview Options: ☐ Show all entries in the file ☒ Show first 100 entries.

Column Names: ☐ Transfer first line as column names Start Import Row: 1 Comment Line: ☐

Network Import Options: Default Interaction: pp

Refresh Preview

Preview

Text File

Left Click: Enable/Disable Column, Right Click: Edit Column

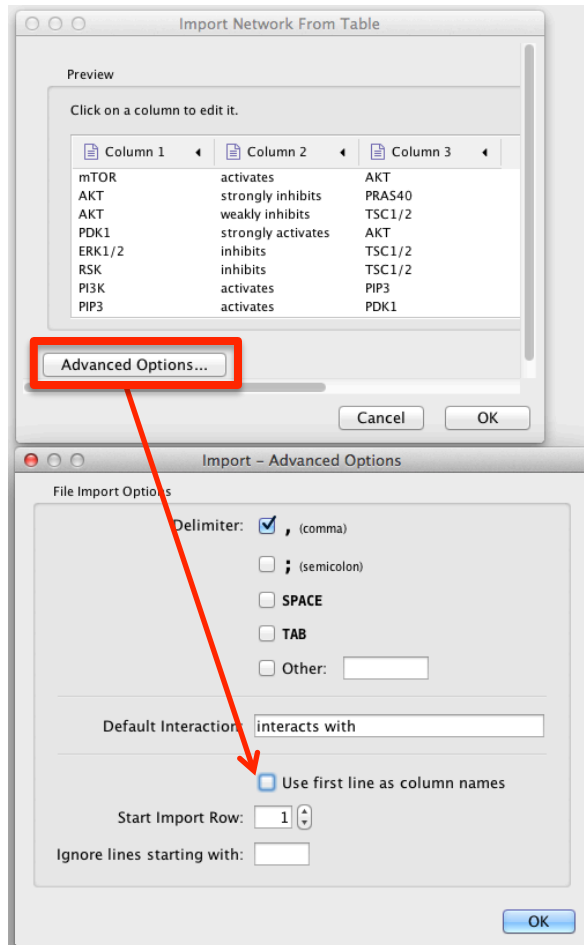
Column 1	Column 2	Column 3
mTOR	activates	AKT
AKT	strongly inhibi...	PRAS40
AKT	weakly inhibits	TSC1/2
PDK1	strongly activa...	AKT
ERK1/2	inhibits	TSC1/2
RSK	inhibits	TSC1/2
PI3K	activates	PIP3
PIP3	activates	PDK1

OK Cancel

For Cytoscape 3.2.x and 3.3.x:

Click 'Show Text File Import Options' and ensure 'Space' is **NOT** selected.

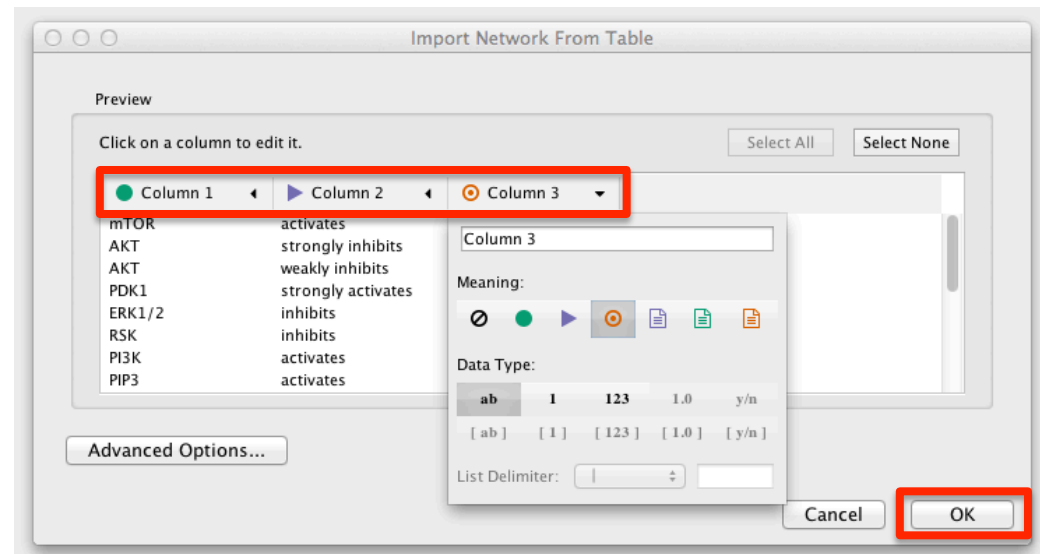
For 'Source Interaction', select 'Column 1'. For 'Interaction Type', select 'Column 2'. For 'Target Interaction', select 'Column 3'. Click 'OK'.

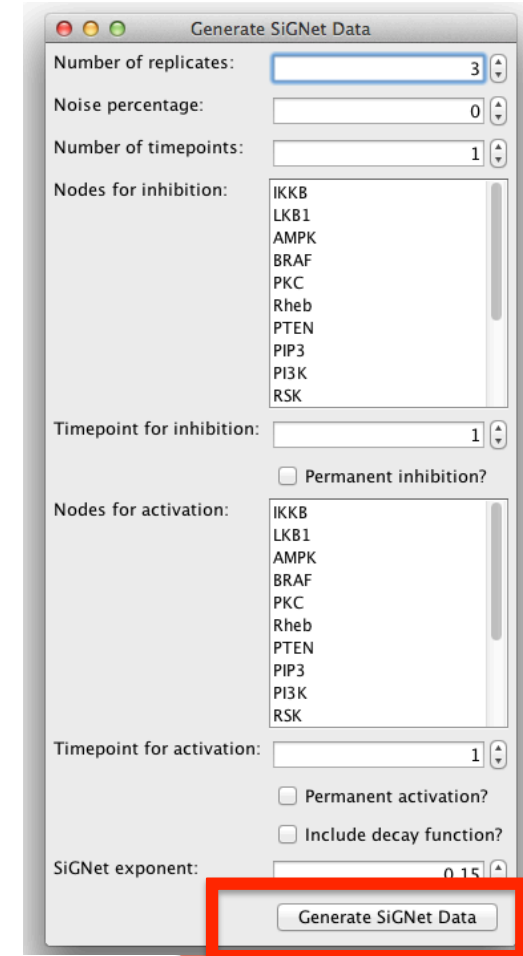
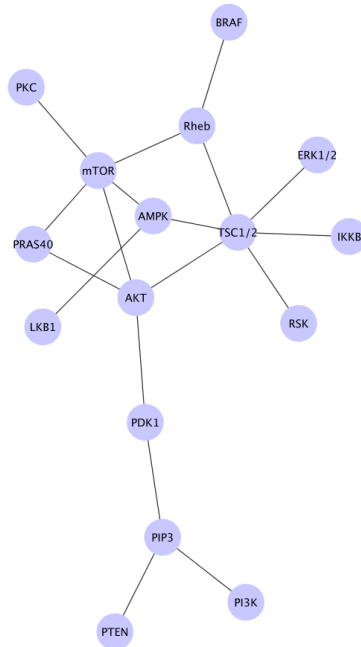
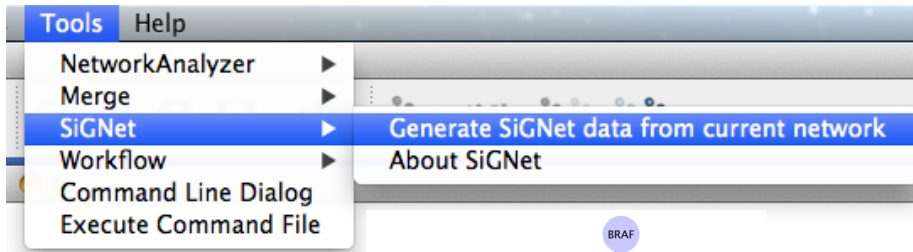


For Cytoscape 3.4.0+:

Select 'Advanced Options' and ensure 'Use first line as column names' is **NOT** selected.

Click on 'Column 1' and select 'Source node' as its meaning. Click on 'Column 2' and select 'Interaction type' as its meaning. Click on 'Column 3' and select 'Target node' as its meaning. Click 'OK'.

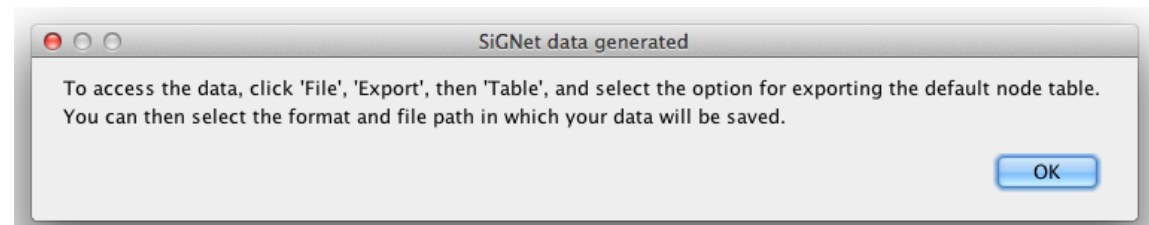
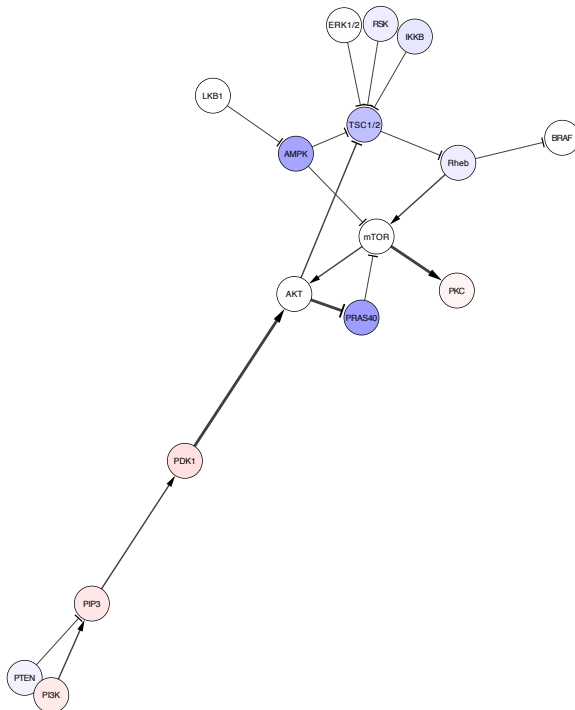




3. Click 'Tools' in the Cytoscape toolbar, then 'SiGNet' and 'Generate SiGNet data from current network'. In the popup menu, select the number of 'experimental replicates' required, the level of noise in the system (if required), the number of timepoints and whether any nodes should be inhibited or activated, and when. When you are happy with the parameters displayed, click 'Generate SiGNet Data'.

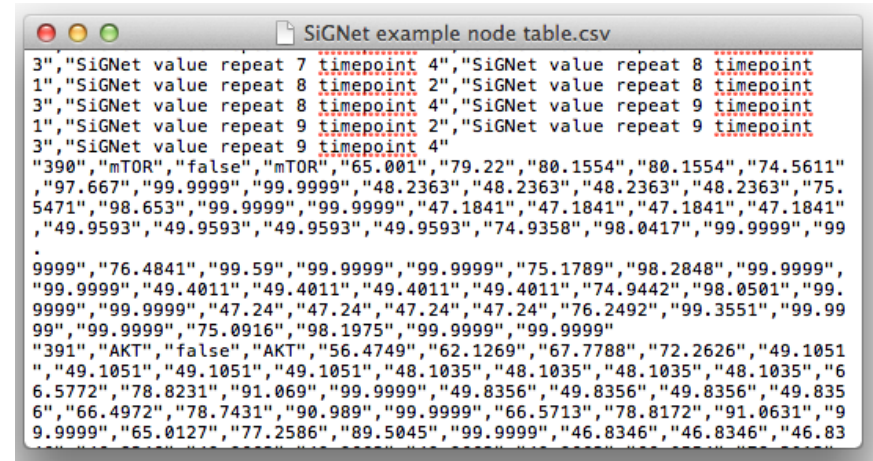
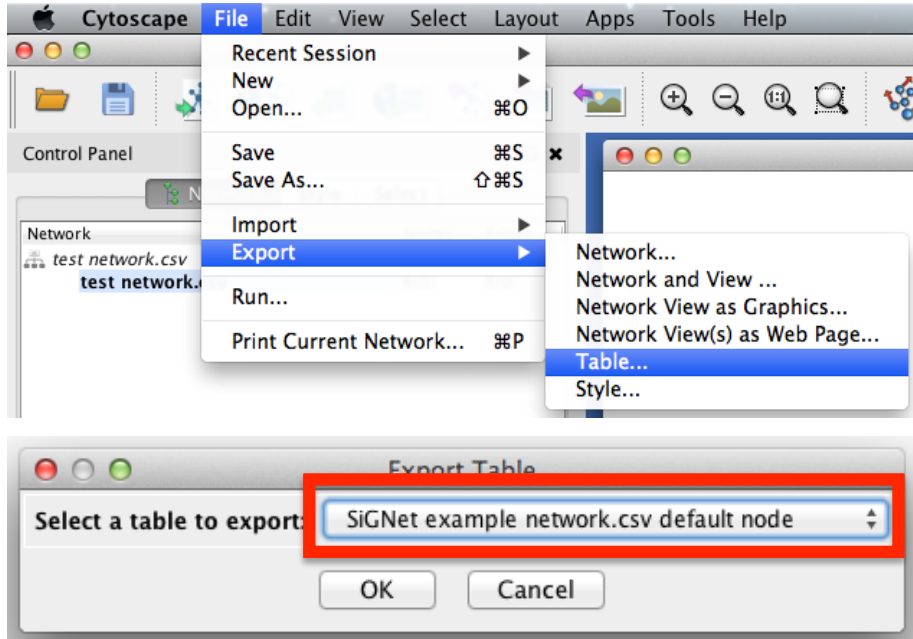
shared...	name	△ SiGNet value average timepoint 1	SiGNet value repeat 1 timepoint 1	SiGNet value repeat 2 timepoint 1	SiGNet value repeat 3 timepoint 1	SiGNet value repeat 4 timepoint 1
mTOR	mTOR	75.8557	77.1068	75.6155	77.5359	76.2225
PKC	PKC	68.1572	77.384	75.1258	75.7901	76.0154
PDK1	PDK1	66.6257	85.5357	45.7638	82.4639	83.114
PIP3	PIP3	57.6546	48.9556	63.2422	49.8048	47.1651
AKT	AKT	56.5174	64.4207	49.1723	62.6577	64.2085
PI3K	PI3K	53.7134	64.144	47.5145	64.4513	45.1646
PTEN	PTEN	45.0414	39.3388	49.6009	40.2854	46.8341
BRAF	BRAF	44.9246	47.6155	42.7305	41.8907	42.0304
ERK1/2	ERK1/2	44.3313	48.3133	38.6268	45.5328	38.5318

Node Table

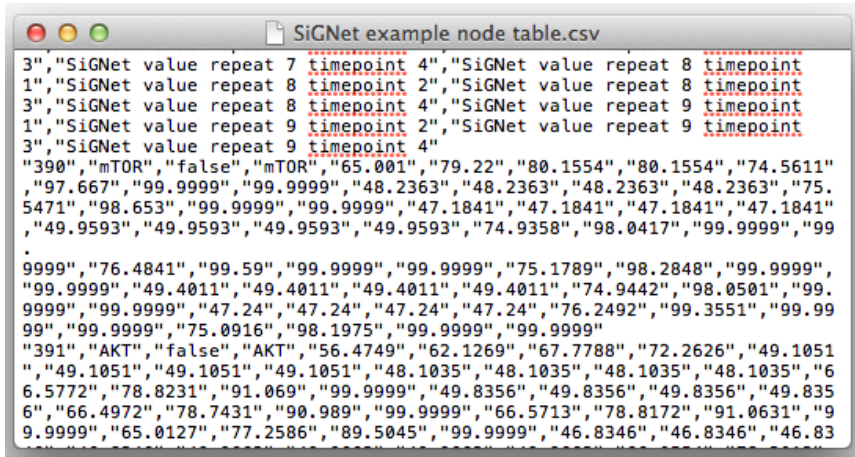


4. The SiGNet algorithm now runs and populates the node table with data for each node at each replicate and at each time point, along with node averages for each time point. A pop up message will appear explaining how to export the data.

In the network shown here, nodes have been coloured according to their SiGNet average value for timepoint 1 – this can be achieved using the ‘Style – Node’ tab under the Cytoscape Control Panel.



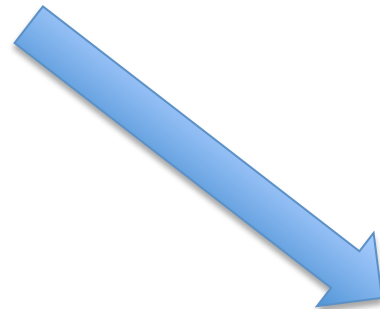
5. By clicking 'File', 'Export', 'Table' and then selecting the default node table to export, the SiGNet data you have generated will be exported to the location you specify.



```

3","SiGNet value repeat 7 timepoint 4","SiGNet value repeat 8 timepoint
1","SiGNet value repeat 8 timepoint 2","SiGNet value repeat 8 timepoint
3","SiGNet value repeat 8 timepoint 4","SiGNet value repeat 9 timepoint
1","SiGNet value repeat 9 timepoint 2","SiGNet value repeat 9 timepoint
3","SiGNet value repeat 9 timepoint 4"
"390","mTOR","false","mTOR","65.001","79.22","80.1554","80.1554","74.5611"
,"97.667","99.9999","99.9999","48.2363","48.2363","48.2363","48.2363","75.
5471","98.653","99.9999","99.9999","47.1841","47.1841","47.1841","47.1841"
,"49.9593","49.9593","49.9593","49.9593","74.9358","98.0417","99.9999","99
9999","76.4841","99.59","99.9999","99.9999","75.1789","98.2848","99.9999",
"99.9999","49.4011","49.4011","49.4011","49.4011","74.9442","98.0501","99.
9999","99.9999","47.24","47.24","47.24","47.24","76.2492","99.3551","99.99
99","99.9999","75.0916","98.1975","99.9999","99.9999"
"391","AKT","false","AKT","56.4749","62.1269","67.7788","72.2626","49.1051
","49.1051","49.1051","49.1051","48.1035","48.1035","48.1035","48.1035","6
6.5772","78.8231","91.069","99.9999","49.8356","49.8356","49.8356","49.835
6","66.4972","78.7431","90.989","99.9999","66.5713","78.8172","91.0631","9
9.9999","65.0127","77.2586","89.5045","99.9999","46.8346","46.8346","46.83

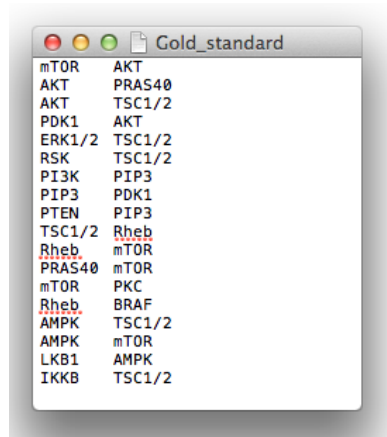
```



Input file into inference algorithm
e.g. Parmigene CLR algorithm, Parmigene MRNET
algorithm

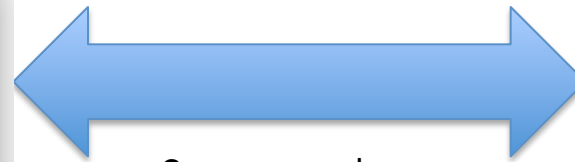
6. The files of simulated data generated by SiGNet can now be used for inference of network structure. Many inference techniques are available, including those based on correlations, Bayesian inference, mutual information and hybrid approaches. Here we are using two algorithms from the Parmigene R Package¹ to generate inferred interactions between nodes for timepoint 3.

1. Sales G and Romualdi C. *parmigene*—a parallel R package for mutual information estimation and gene network reconstruction. *Bioinformatics* (2011) 27 (13): 1876-1877 doi:10.1093/bioinformatics/btr274



Gold_standard	
mTOR	AKT
AKT	PRAS40
AKT	TSC1/2
PDK1	AKT
ERK1/2	TSC1/2
RSK	TSC1/2
PI3K	PIP3
PIP3	PDK1
PTEN	PIP3
TSC1/2	Rheb
Rheb	mTOR
PRAS40	mTOR
mTOR	PKC
Rheb	BRAF
AMPK	TSC1/2
AMPK	mTOR
LKB1	AMPK
IKKB	TSC1/2

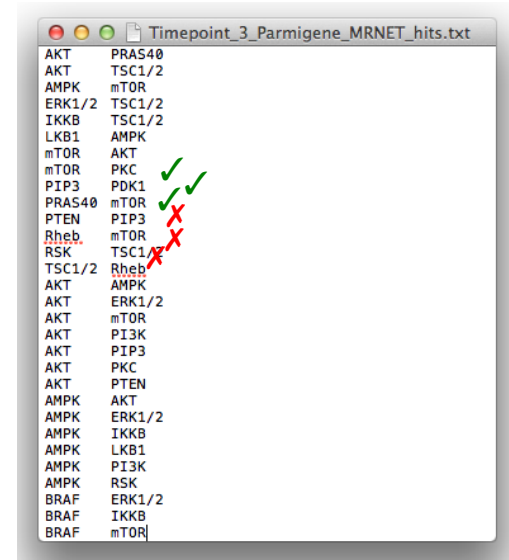
Gold standard interactions



Compare and score

$$\text{Sensitivity} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$



Timepoint_3_Parmigene_MRNET_hits.txt	
AKT	PRAS40
AKT	TSC1/2
AMPK	mTOR
ERK1/2	TSC1/2
IKKB	TSC1/2
LKB1	AMPK
mTOR	AKT
mTOR	PKC
PIP3	PDK1
PRAS40	mTOR
PTEN	PIP3
Rheb	mTOR
RSK	TSC1/2
TSC1/2	Rheb
AKT	AMPK
AKT	ERK1/2
AKT	mTOR
AKT	PI3K
AKT	PIP3
AKT	PKC
AKT	PTEN
AMPK	AKT
AMPK	ERK1/2
AMPK	IKKB
AMPK	LKB1
AMPK	PI3K
AMPK	RSK
BRAF	ERK1/2
BRAF	IKKB
BRAF	mTOR

Edges inferred by Parmigene MRNET algorithm

Technique	Number in SiGNet Gold Standard	Number in inferred file	True positives	False positives	False negatives	Sensitivity	Precision	F1
Parmigene MRNET	18	170	14	156	4	0.778	0.082	0.149
Parmigene ARACNE A	18	130	10	120	8	0.556	0.077	0.135
Parmigene CLR	18	158	11	147	7	0.611	0.070	0.125

7. Inferred edges can now be checked against the gold standard, “correct” network structure we used in Step 1. For example, the user can score the results of inference in terms of true positives (i.e. interactions correctly identified by the inference approach), true negatives, false positives and false negatives. These scores can then be used to calculate the accuracy, precision, sensitivity, false discovery rate and so on for the inference techniques used. In this example we can see that the Parmigene MRNET algorithm has a higher sensitivity and precision than other Parmigene algorithms for the network we simulated in SiGNet.