

Sage Bionetworks Mission

Sage Bionetworks is a non-profit organization with a vision to create a “commons” where integrative bionetworks are evolved by contributor scientists with a shared vision to accelerate the elimination of human disease

Building Disease Maps

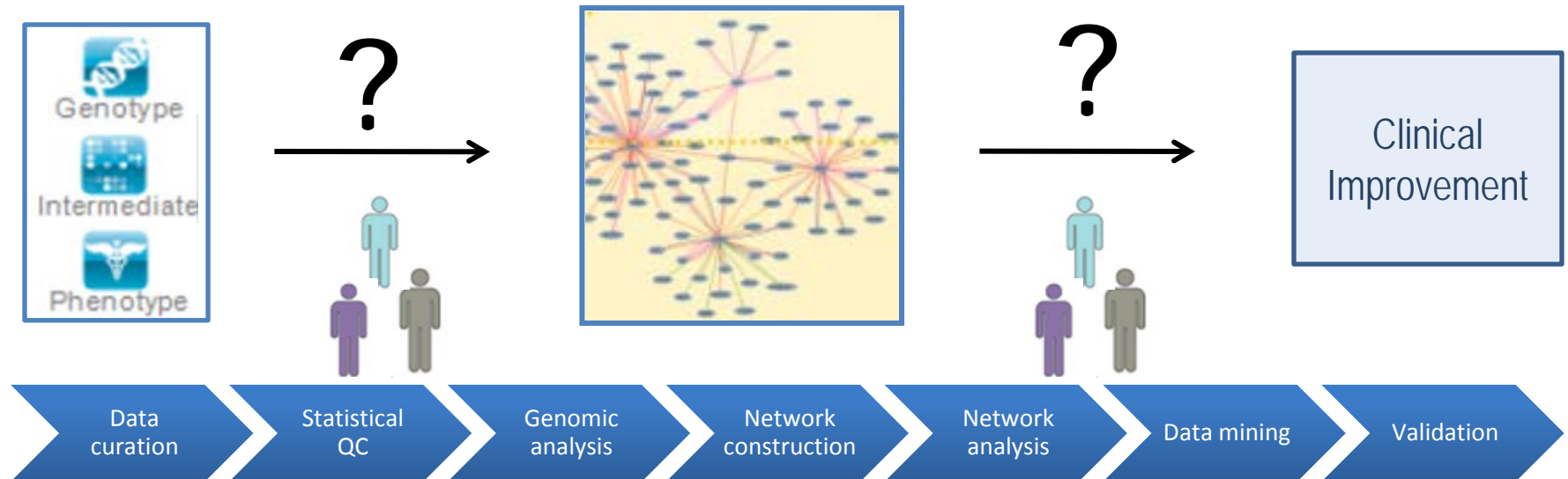
Data Repository



Commons Pilots

Discovery Platform

Need: Community-based analysis is required to build accurate models of disease



Within the analytical community:

- Quality assessment of data, tools and models
- Reproducibility of models
- Reusability of data
- Validation of models across multiple disease models and patient subtypes

Across the field of biology:

- Validation of model predictions in experimental/clinical setting
- Accelerated pace towards improved therapeutic development

Watch What I Do, Not What I Say

 **Synapse** | CONTRIBUTE to the CURE

Welcome Michael K.
[My profile](#) | [Logout](#)

DatasetsToolsNetworksPeopleProjects

Home > [All Projects](#) > [Federation Warburg Project](#) > Network Generation Analysis

Network Generation Analysis

Privacy Setting: [Public](#) » [Learn more...](#)

Overview

In this analysis, we generate a list of key drivers for the network model governing prostate cancer. The methodology is based on the Rosetta Bayesian network approach

» [Read more...](#)

Created by: Xudong Dai on 01-Dec-2011
Last modified by: Xudong Dai on 12-Apr-2011
Version: 2.0
Public datasets used: [MSKCC Prostate Cancer - exon expression](#)
[MSKCC Prostate Cancer - QC'd phenotypes](#)
[MSKCC Prostate Cancer - CNV](#)
Public tools used: [Coexpression Module](#)
[Bayesian Network Generation](#)
Public networks generated: [Warburg Cancer Network](#)

[Follow this analysis](#)

Notes for this analysis

[View All](#)
[Add a note](#)

Expression Normalization rerun [Xudong Dai](#) 12-Apr-2011

Gene List Selection run [Xudong Dai](#) 16-Feb-2011

Network Module run [Xudong Dai](#) 15-Feb-2011

OK, ready to go.
Xudong Dai 20-Jan-2011

We need to build the priors for sample matching before .
Jun Zhu 20-Jan-2011

Analysis

Select a step to view its details and resources

[Load in Taverna](#)
[Add a step](#)

Step types

- Dataset
- Normalization
- Correlation
- Network
- Gene List



```
graph TD; A[MSKCC Prostate Cancer exon expression] --> B[Expression Normalization v2 by Xudong D. 12-Apr-2011]; B --> C[MSKCC Prostate Cancer Phenotypes]; B --> D[MSKCC Prostate Cancer.CNV]; C --> E[Survival Analysis v1 by Xudong D. 16-Feb-2011]; B --> F[Coexpression Network Generation v1 by Zhang B. 02-Feb-2011]; D --> G[Sample Matching v1 by Xudong D. 20-Jan-2011];
```

Real-time tools to track data analysis projects

Reduce, Reuse, Recycle

 **Synapse**

CONTRIBUTE to the CURE

Welcome Michael K.
[My profile](#) | [Logout](#)

DatasetsToolsNetworksPeopleProjects

[Home](#) > All Datasets

All Datasets

Sage Bionetworks has established a catalog of datasets for use in integrative genomics analysis and building predictive computational disease models. The goal is to collate, curate, and host these datasets for use by the entire research community.

[» Read more...](#)

Invitation to Participate

Do you know of or have you generated a dataset that you would like to have curated and quality controlled to Sage Bionetworks standards?

[Contact Us](#)

FILTER LIST: By Keywords(s)  Any Sample Size  All Species  All Tissues  All Diseases   Add / remove columns

Dataset Name	Layers	Number of Samples	Status	Species	Tumor / Tissue Type	Disease	Investigator	Created On
MSKCC Prostate Cancer	  	261	Current	Human	Prostate	Cancer	Charles Sawyers	11-Apr-2011
Mouse Model of Diet-Induced Atherosclerosis	  	111	Current	Mouse	Liver	Metabolic Disease	Jake Lusis	11-Apr-2011
Mouse Model of Sexually Dimorphic Atherosclerotic Traits	  	334	Current	Mouse	Muscle, Liver, Brain, Adipose	CVD	Jake Lusis	11-Apr-2011
Human Liver Cohort	 	517	Current	Human	Liver	CVD	Fred Guengrich/Steve Strom/ Erin Schuetz	11-Apr-2011
Heterogeneous Stock Mice	  	1900	Current	Mouse	Hippocampus, Lung, Liver	Multiple	Flint/Mott	11-Apr-2011
Cancer Cell line Panel	  	649	Current	Human	Cell Line	Cancer	Richard Wooster	11-Apr-2011
Harvard Brain Tissue Resource Center	  	803	Current	Human	Brain visual cortex, Brain cerebellum, Brain prefrontal cortex	Neurological Disease	Francine Benes/ Eric Schadt	11-Apr-2011
Mouse Model of Blood Pressure	  	350	Current	Mouse	Kidney cortex, Liver, Kidney medulla, Adipose	CVD	Eric Schadt	11-Apr-2011
LFN-Kronos-PHASE I	  	364	Current	Human	Brain Cerebellum, Brain Parietal Cortex, Brain Temporal Cortex, Brain Frontal Cortex	Alzheimers	Amanda Myers	11-Apr-2011

Page 1 of 12

Displaying 1 - 10 of 115

A Repository of Data, Tools, and Models

Most of the People You Need to Work with Don't Work with You

 **Synapse** | CONTRIBUTE to the CURE

Welcome Michael K.
[My profile](#) | [Logout](#)

DatasetsToolsNetworks**People**Projects

[Home](#) > [All People](#) >

Charles Sawyers, Howard Hughes Medical Institute

 **Profile**

Dr. Sawyers is investigating the signaling pathways that drive the growth of cancer cells, with an eye toward designing new treatment options for patients with chronic myeloid leukemia, prostate cancer, and glioblastoma.

Our laboratory is focused on characterizing signal transduction pathway abnormalities in various cancers with an eye toward translational implications. One example is chronic myeloid leukemia (CML) which is caused by the BCR-ABL fusion gene, expressed as a consequence of the Philadelphia chromosome translocation.

Profession/Title: Chair, Human Oncology and Pathogenesis Program Investigator

Affiliation: Howard Hughes Medical Institute

Contact Info: martinb@mskcc.org

Website(s): www.mskcc.org/mskcc/html/72308.cfm
<http://www.mskcc.org/prg/prg/bios/994.cfm>
[Federation Warburg](#)

Followed Projects:

Followed Datasets: [MSKCC Prostate Cancer](#)

[Privacy Policy](#) | [Platform support](#) | [Creative Commons License](#)


[FAQ](#) | [Submit feedback](#) | [Visit the Sage website](#)

 **Life Sciences**
DISCOVERY FUND

 **Sage Bionetworks**
1100 Fairview Ave N
Seattle WA 98109
sage@sagebase.org

An online community to connect people and projects

My Other Computer is Amazon

 **Synapse** | CONTRIBUTE to the CURE

Welcome Michael K.
[My profile](#) | [Logout](#)

[Datasets](#) [Tools](#) [Networks](#) [People](#) [Projects](#)

[Advanced Search](#)

Welcome to SYNAPSE

Synapse is an online community built for a new biology... more open, more collaborative, more data-driven, and more reproducible.

On Synapse you can:

- Discover and explore new [Datasets](#)
- Learn new [Tools](#) and run powerful analysis in the cloud
- Explore cutting edge [Network Models](#)
- Find [People](#) to collaborate with
- Start new research [Projects](#), with anyone working anywhere

Synapse is currently hosting over 20 curated global coherent datasets in which genomic variation and molecular traits are combined with clinically-relevant phenotypes.

We are actively working to build an open repository of high value studies.

Call for DATA

Do you know of a dataset that you would like to see made available?

[Contact Us](#)

Synapse STATS

- » [134 People](#)
- » [45 Projects](#)
- » [78 Networks](#)
- » [14 Tools](#)


[Register](#)

Featured PROJECT


The Federation Aging project is exploring the large effect of age on human gene expression by exploring a set of genes whose expression levels are tightly associated with age.

» [Learn more...](#)


Top Synapse CONTRIBUTORS



Charles Sawyers et al, MSKCC
DNA copy number, mRNA expression profiles, candidate gene resequencing, and clinical traits from 218 prostate tumors.



Amanda Myers, University of Miami
Genotypes, mRNA expression profiles, and clinical traits from 176 Alzheimer disease and 188 non-diseased brain samples.



Beatriz Carvalho & Gerrit Meijer, VUMC
DNA copy number, mRNA expression, and tumor traits from 36 colorectal adenocarcinoma and 33 adenoma tumors.

» [See all contributors...](#)

Synapse NEWS

12-Apr-2011: A new analysis method for normalizing data was published by Brig Mecham. » [Learn more...](#)

08-Apr-2011: Eric Schadt and colleagues have published an article on a new model for diabetes. » [Learn more...](#)

Example of an EXPERIENCE IN SYNAPSE:

75 x 75

75 x 75

75 x 75

75 x 75



75 x 75


75 x 75

75 x 75

[Browse datasets & projects](#) [Download datasets](#) [Upload your analyses](#) [Upload your own dataset](#) [Create a new project](#) [Collaborate with other users](#) [Discover new insights](#)

[Privacy Policy](#) [Platform support](#) [Creative Commons License](#)

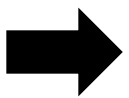
 

 **Sage Bionetworks**
1100 Fairview Ave N

Cloud-based data stores and compute resources

Do: What will the Synapse Platform and Sage Commons Enable?

- Development of **Robust, Reproducible, and Reusable** analytical methods
- **Integration of Data, Tools and Methods** from across community
- Development of a **Disease Model Repository**
- Forum for **New Collaborations** between technically and geographically distinct scientific groups
- Access to **Cloud-Compute** resources co-located with large-scale data
- See also the [Platform Vision Document](#) posted on our wiki



Translation of Data to Knowledge to Improved Therapeutic

Approaches to Disease: An improved molecular understanding of the causes of disease leading to quicker, more rational and more successful development of new drugs

Welcome to SYNAPSE

Synapse is an online community built for a new biology... more open, more collaborative, more data-driven, and more reproducible.

On Synapse you can:

- Discover and explore new [Datasets](#)
- Learn new [Tools](#) and run powerful analysis in the cloud
- Explore cutting edge [Network Models](#)
- Find [People](#) to collaborate with
- Start new research [Projects](#), with anyone working anywhere

Synapse is currently hosting over 20 curated global coherent datasets in which genomic variation and molecular traits are combined with clinically-relevant phenotypes.

We are actively working to build an open repository of high value studies.

Call for DATA

Do you know of a dataset that you would like to see made available?

[Contact Us](#)

Synapse STATS

- » [134 People](#)
- » [45 Projects](#)
- » [78 Networks](#)
- » [14 Tools](#)

[Register](#)

Example of an EXPERIENCE IN SYNAPSE:

75 x 75

75 x 75

75 x 75

75 x 75

75 x 75

75 x 75

75 x 75

Browse datasets & projects

Download datasets

Upload your analyses

Upload your own dataset

Create a new project

Collaborate with other users

Discover new insights

Featured PROJECT

The Federation Aging project is exploring the large effect of age on human gene expression by exploring a set of genes whose expression levels are tightly associated with age.

» [Learn more...](#)

Top Synapse CONTRIBUTORS



Charles Sawyers et al, MSKCC
 DNA copy number, mRNA expression profiles, candidate gene resequencing, and clinical traits from 218 prostate tumors.



Amanda Myers, University of Miami
 Genotypes, mRNA expression profiles, and clinical traits from 176 Alzheimer disease and 188 non-diseased brain samples.



Beatriz Carvalho & Gerrit Meijer, VUMC
 DNA copy number, mRNA expression, and tumor traits from 36 colorectal adenocarcinoma and 33 adenoma tumors.

» [See all contributors...](#)

Synapse NEWS

12-Apr-2011: A new analysis method for normalizing data was published by Brig Mecham. » [Learn more...](#)

08-Apr-2011: Eric Schadt and colleagues have published an article on a new model for diabetes. » [Learn more...](#)

[Home](#) > All Datasets

All Datasets



Sage Bionetworks has established a catalog of datasets for use in integrative genomics analysis and building predictive computational disease models. The goal is to collate, curate, and host these datasets for use by the entire research community.




» [Read more...](#)

Invitation to Participate

Do you know of or have you generated a dataset that you would like to have curated and quality controlled to Sage Bionetworks standards?

[Contact Us](#)

FILTER LIST: By Keywords(s)  Any Sample Size All Species All Tissues All Diseases  [Add / remove columns](#)


Dataset Name	Layers	Number of Samples	Status	Species	Tumor / Tissue Type	Disease	Investigator	Created On
MSKCC Prostate Cancer	  	261	Current	Human	Prostate	Cancer	Charles Sawyers	11-Apr-2011
Mouse Model of Diet-Induced Atherosclerosis	  	111	Current	Mouse	Liver	Metabolic Disease	Jake Lusis	11-Apr-2011
Mouse Model of Sexually Dimorphic Atherosclerotic Traits	  	334	Current	Mouse	Muscle, Liver, Brain, Adipose	CVD	Jake Lusis	11-Apr-2011
Human Liver Cohort	 	517	Current	Human	Liver	CVD	Fred Guengrich/Steve Strom/ Erin Schuetz	11-Apr-2011
Heterogeneous Stock Mice	  	1900	Current	Mouse	Hippocampus, Lung, Liver	Multiple	Flint/Mott	11-Apr-2011
Cancer Cell line Panel	  	649	Current	Human	Cell Line	Cancer	Richard Wooster	11-Apr-2011
Harvard Brain Tissue Resource Center	  	803	Current	Human	Brain visual cortex, Brain cerebellum, Brain prefrontal cortex	Neurological Disease	Francine Benes/ Eric Schadt	11-Apr-2011
Mouse Model of Blood Pressure	  	350	Current	Mouse	Kidney cortex, Liver, Kidney medulla, Adipose	CVD	Eric Schadt	11-Apr-2011
LFN-Kronos-PHASE I	  	364	Current	Human	Brain Cerebellum, Brain Parietal Cortex, Brain Temporal Cortex, Brain Frontal Cortex	Alzheimers	Amanda Myers	11-Apr-2011


[Home](#) > [All Datasets](#) > MSKCC Prostate Cancer

MSKCC Prostate Cancer

Privacy Setting: [Public](#) » [Learn more...](#)

Overview

 Follow this Dataset

 See Terms of Use


Expand

Genetic and epigenetic alterations have been identified that lead to transcriptional Annotation of prostate cancer genomes provides a foundation for discoveries that can impact disease understanding and treatment. Concordant assessment of DNA copy number, mRNA expression, and focused exon resequencing in the 218 prostate cancer tumors represented in this dataset have identified the nuclear receptor coactivator NCOA2 as an oncogene in approximately...







Disease(s): Cancer
Species: Human
Tissue Type(s): GCD
Tissue/Tumor: Prostate
Reference Publication: [Integrative genomic profiling of human prostate cancer. Taylor BS et al., Cancer Cell. 2010 Jul 13;18\(1\):11-22.](#)
Other Publications: 0 [view](#)
Status: Current
Version: 1.0.0

Posted: 17-Apr-2011
Creator: Charles Sawyers
Curated On: 13-Sep-2008
Contributor(s)/Institution: Memorial Sloan Kettering Cancer Center
Followers: 7 [view](#)
Number of Samples: 261
Number of Downloads: 32
Download Availability: unspecified
Release Notes: [view](#)



Dataset Layers

 Download Dataset

Select a name to view sample data

Name	Type	Version	Number of S	Download
Curated phenotypes		1.0.0	261	download
sequencing		1.0.0	50	download
CNV		1.0.0	230	download
mRNA expression		1.0.0	142	download
exon expression		1.0.0	184	download
QCed phenotypes		1.0.0	261	download

Conversations and Events for this Dataset

 View All
 Add a note

New analysis using this dataset in the Federation Warburg Project [Xudong D.](#) 21-Nov-2010

The paper references 149 matched normal samples. Do you have this data?
Sam S. 22-Dec-2010



We have only able to obtain 29 matched normal samples.
Matt F. 22-Nov-2010

New analysis using this dataset in the MetaGEO Project [Brig M.](#) 01-Apr-2011

Look out, there is a large batch effect in this data. Have been working on a way to remove it.
Brig M. 02-Apr-2011

Appreciate the help. We can post new version of the

Analyses using this Dataset

 View All
 Suggest an analysis

Analysis	Project	Last Modified
Network Generation	Federation Warburg	12-Apr-2011
Network Survey	Federation Warburg	23-Mar-2011
Interactome	Federation Warburg	04-Mar-2011
GSE 21034	MetaGEO	2-Apr-2011

Charles Sawyers, Howard Hughes Medical Institute



Profile

Dr. Sawyers is investigating the signaling pathways that drive the growth of cancer cells, with an eye toward designing new treatment options for patients with chronic myeloid leukemia, prostate cancer, and glioblastoma.

Our laboratory is focused on characterizing signal transduction pathway abnormalities in various cancers with an eye toward translational implications. One example is chronic myeloid leukemia (CML) which is caused by the BCR-ABL fusion gene, expressed as a consequence of the Philadelphia chromosome translocation.

Profession/Title: Chair, Human Oncology and Pathogenesis
Program Investigator
Howard Hughes Medical Institute

Affiliation:

Contact Info: martinb@mskcc.org


Website(s): www.mskcc.org/mskcc/html/72308.cfm
<http://www.mskcc.org/prg/prg/bios/994.cfm>
 Federation Warburg


Followed Projects:
Followed Datasets: [MSKCC Prostate Cancer](#)

MSKCC Prostate Cancer

Privacy Setting: [Public](#) » [Learn more...](#)

Overview

 Follow this Dataset

 See Terms of Use

Expand

Genetic and epigenetic alterations have been identified that lead to transcriptional Annotation of prostate cancer genomes provides a foundation for discoveries that can impact disease understanding and treatment. Concordant assessment of DNA copy number, mRNA expression, and focused exon resequencing in the 218 prostate cancer tumors represented in this dataset have identified the nuclear receptor coactivator NCOA2 as an oncogene in approximately...

Disease(s): Cancer
Species: Human
Tissue Type(s): GCD
Tissue/Tumor: Prostate
Reference Publication: [Integrative genomic profiling of human prostate cancer. Taylor BS et al., Cancer Cell. 2010 Jul 13;18\(1\):11-22.](#)
Other Publications: 0 [view](#)
Status: Current
Version: 1.0.0

Posted: 11-Apr-2011
Creator: Charles Sawyers
Curated On: 13-Sep-2008
Contributor(s)/Institution: Memorial Sloan Kettering Cancer Center
Followers: 7 [view](#)
Number of Samples: 261
Number of Downloads: 32
Download Availability: unspecified
Release Notes: [view](#)



Dataset Layers

 Download Dataset

Select a name to view sample data

Name	Type	Version	Number of S	Download
Curated phenotypes		1.0.0	261	download
sequencing		1.0.0	50	download
CNV		1.0.0	230	download
mRNA expression		1.0.0	142	download
exon expression		1.0.0	184	download
QCed phenotypes		1.0.0	261	download

Conversations and Events for this Dataset

 View All
 Add a note

New analysis using this dataset in the Federation Warburg Project [Xudong D.](#) 21-Nov-2010

The paper references 149 matched normal samples. Do you have this data?
Sam S. 22-Dec-2010



We have only able to obtain 29 matched normal samples.
Matt F. 22-Nov-2010

New analysis using this dataset in the MetaGEO Project [Brig M.](#) 01-Apr-2011

Look out, there is a large batch effect in this data. Have been working on a way to remove it.
Brig M. 02-Apr-2011

Appreciate the help. We can post new version of the

Analyses using this Dataset

 View All
 Suggest an analysis

Analysis	Project	Last Modified
Network Generation	Federation Warburg	2-Apr-2011
Network Survey	Federation Warburg	23-Mar-2011
Interactome	Federation Warburg	04-Mar-2011
GSE 21034	MetaGEO	2-Apr-2011

Federation Warburg Project

Privacy Setting: [Public](#) » [Learn more...](#)

Overview

[Follow this project](#)

[Send leaders a message](#)

In oncology, the Warburg effect is the observation that most cancer cells predominantly produce energy by a high rate of glycolysis followed by lactic acid fermentation in the cytosol, rather than by a comparatively low rate of glycolysis followed by oxidation of pyruvate in mitochondria like most normal cells. The latter process is aerobic (uses oxygen). Malignant rapidly-growing tumor cells typically have glycolytic rates that are up to 200 times higher than those of their normal tissues of origin; this occurs even if oxygen is plentiful.

The aim of this project is to improve understanding of the mechanism by which cancer cells undergo a transition in their metabolic state associated with immortalization.

» [Read more...](#)

Project formed: 01-Nov-2010

Leaders: [Xudong Dai](#), [Mariano Alvarez](#), [Andrew Beck](#)

Members: 8 [view](#)

Publications: 5 [view](#)

Status: Active

Project Web Site: [Federation Google Site](#)

Analyses

Select a name to view the analysis details and resources

[Add analysis](#)

Analysis Name	Latest Activity	Modified by / Date
Network generation	Expression Normalization Step updated	Xudong D. 12-Apr-2011
Interactome	Transcription Factor Prediction Step created	Mariano A. 02-Mar-2011
Network survey	Correlation calculation Step updated	Andrew B. 25-Mar-2011

Conversations and Events for this Project

[View All](#)

[Add a note](#)

Expression Normalization step in Network Generation updated [Xudong D.](#) 12-Apr-2011

Thanks Brig, can you have a look at my new results?
Xudong D. 12-Apr-2011

I've looked at this data and found a huge batch effect. You may get cleaner results if you use my normalization method, let me know if you need help.
Brig M. 09-Nov-2010

Looks good to me.
Marino A. 22-Mar-2011

New analysis Network Survey created [Andrew Beck](#) 21-Mar-2011

Can someone please have a look at the way we're

Project History

[View All](#)

Action	User	Date
Step updated	Xudong D.	12-Apr-2011
Document updated	Eric S.	11-Apr-2011
Document updated	Marino A.	11-Apr-2011
Document updated	Xudong D.	09-Apr-2011
Step created	Andrew B.	21-Mar-2011
Step created	Andrew B.	20-Mar-2011
Analysis created	Andrew B.	20-Mar-2011
Document created	Xudong D.	19-Mar-2011
Member added	Antonina M.	18-Mar-2011
Step updated	Mariano A.	02-Mar-2011
Step created	Mariano A.	01-Mar-2011
Step created	Mariano A.	01-Mar-2011
Step created	Mariano A.	01-Mar-2011

Network Generation Analysis

Privacy Setting: [Public](#) » [Learn more...](#)

Overview

[Follow this analysis](#)

In this analysis, we generate a list of key drivers for the network model governing prostate cancer. The methodology is based on the Rosetta Bayesian network approach

» [Read more...](#)

Created by: Xudong Dai on 01-Dec-2011
 Last modified by: Xudong Dai on 12-Apr-2011
 Version: 2.0
 Public datasets used: [MSKCC Prostate Cancer - exon expression](#)
[MSKCC Prostate Cancer - QC'd phenotypes](#)
[MSKCC Prostate Cancer - CNV](#)
 Public tools used: [Coexpression Module](#)
[Bayesian Network Generation](#)
 Public networks generated: [Warburg Cancer Network](#)

Notes for this analysis

[View All](#)
[Add a note](#)

Expression Normalization rerun [Xudong Dai](#)
 12-Apr-2011

Gene List Selection run [Xudong Dai](#) 16-Feb-2011

Network Module run [Xudong Dai](#) 15-Feb-2011

OK, ready to go.
 Xudong Dai 20-Jan-2011

We need to build the priors for sample matching before .
 Jun Zhu 20-Jan-2011

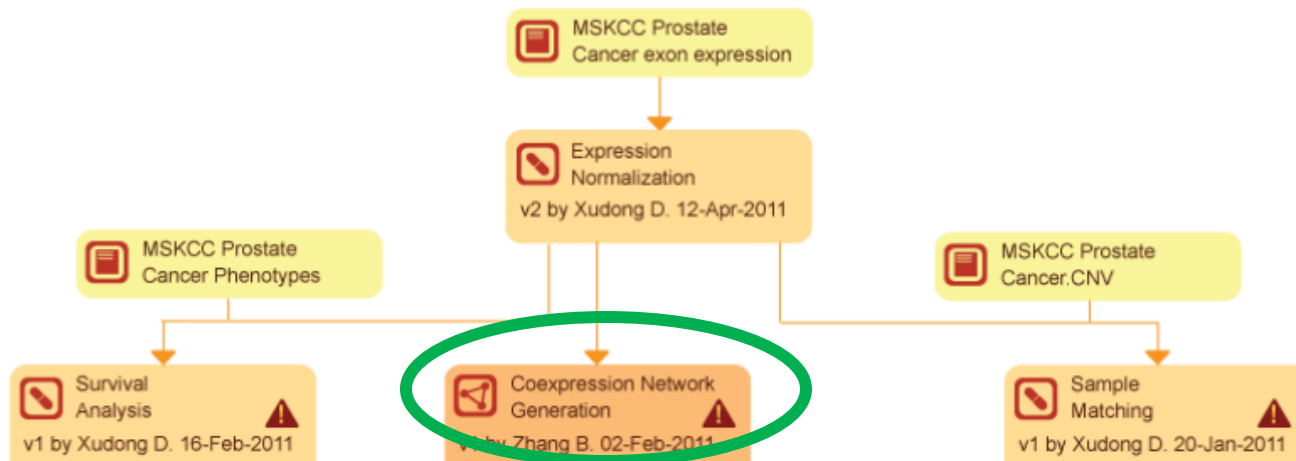
Analysis

Select a step to view its details and resources

[Load in Taverna](#)
[Add a step](#)

Step types

- Dataset
- Normalization
- Correlation
- Network
- Gene List



Datasets

Tools

Networks

People

Projects


[Advanced Search](#)

[Home](#) > [All Projects](#) > [Federation Warburg Project](#) > [Network Generation Analysis](#) > Coexpression Network Generation Step

Coexpression Network Generation Step

Privacy Setting: [Public](#) » [Learn more...](#)

Overview

 Rerun this step

Warning: This input has changed since the latest version of this analysis step.

In this step, we generate a coexpression network based on gene expression levels as measured by affymetrix arrays.

» [Read more...](#)

Created by: Bin Zhang on 15-Feb-2011

Last modified by: Bin Zhang on 15-Feb-2011


Inputs: [Normalized Expression](#)


Outputs: [Warburg Cancer Network](#)

Tool: [CoexNet.SNW](#)

Documentation: [CoexNet.SNW](#)



Notes for this step

 View All

 Add a note

Analysis

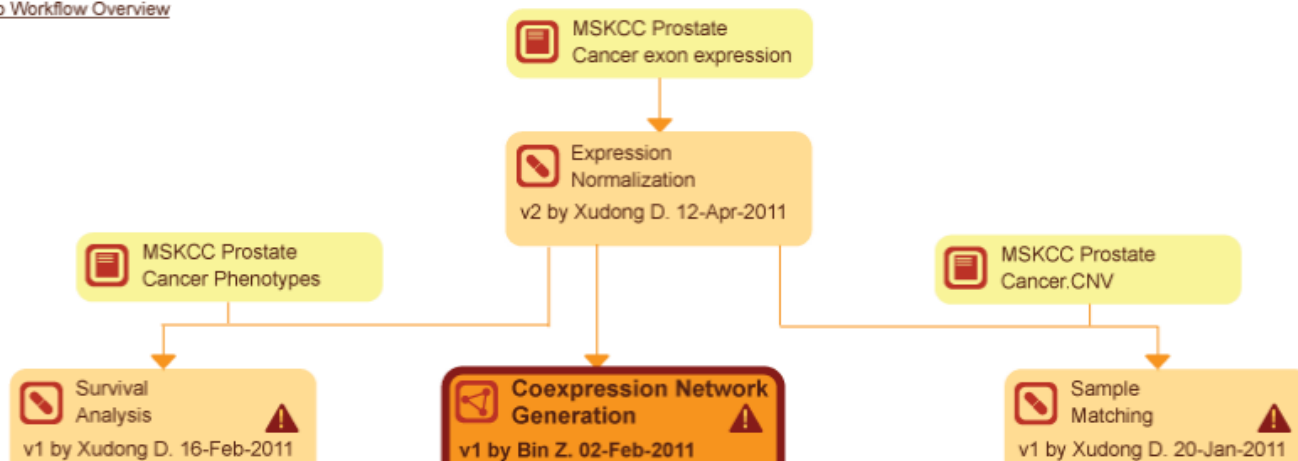
Select a step to view its details and resources

 Load in Taverna
 Add a step

Step types

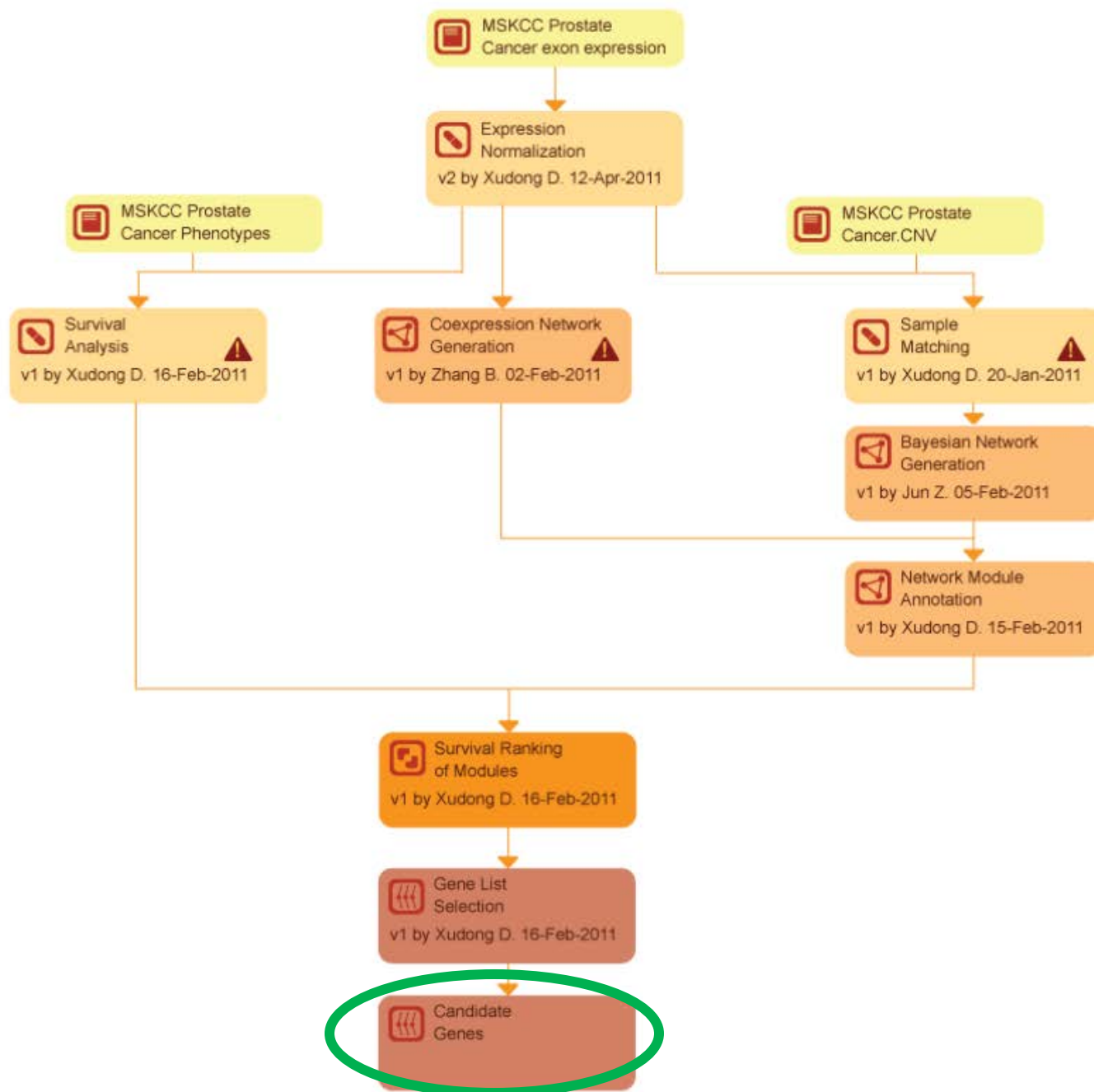
< [Return to Workflow Overview](#)

-  Dataset
-  Normalization
-  Correlation
-  Network
-  Gene List



Step types

-  Dataset
-  Normalization
-  Correlation
-  Network
-  Gene List



Candidate Genes

Privacy Setting: [Public](#) » [Learn more...](#)

Overview

[Follow this analysis](#)

The following genes have been identified as candidate genes in this analysis.

Gene Name
MSMB
ATF3
IL1
TLR4
BRCA2

Created by: Xudong Dai on 16-Feb-2011

Last modified by: Xudong Dai on 16-Feb-2011

Inputs: [Ranked Modules](#)

Outputs: [Warburg Gene List](#)

Notes for this analysis

[View All](#)

[Add a note](#)

Expression Normalization rerun [Xudong Dai](#)
12-Apr-2011

Gene List Selection run [Xudong Dai](#) 16-Feb-2011

Network Module run [Xudong Dai](#) 15-Feb-2011

OK, ready to go.
Xudong Dai 20-Jan-2011

We need to build the priors for sample matching before .
Jun Zhu 20-Jan-2011

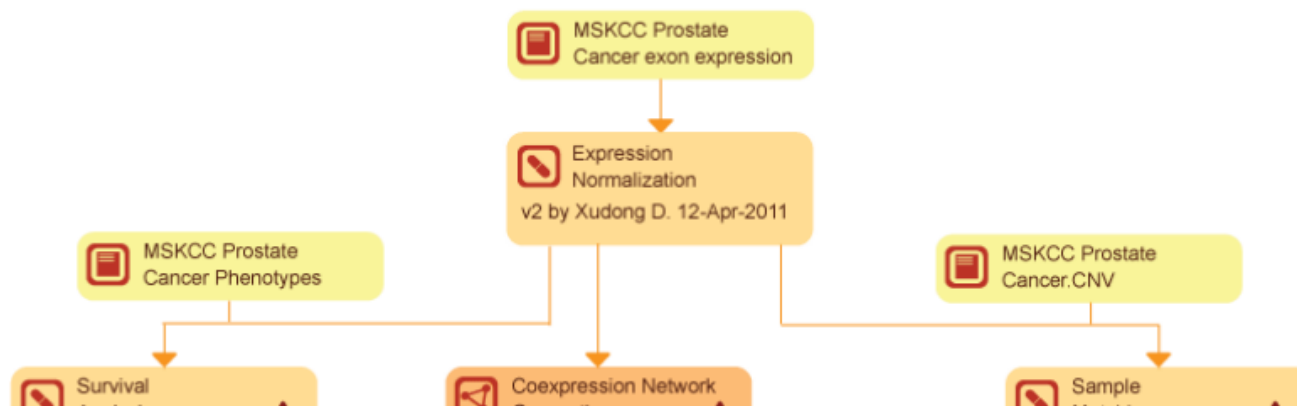
Analysis

Select a step to view its details and resources

[Load in Taverna](#)
[Add a step](#)

Step types

-  Dataset
-  Normalization
-  Correlation
-  Network
-  Gene List



RStudio

File Edit View Workspace Plots Tools Help

signELUA.R x metaGeo.R x DemoScript.R x

Source on Save Run Source

```
1 library(synapseclient)
2 synapseLogin()
3
4 ### do this ahead of time
5 project <- createEntity(Project(list(name="Mikes Playground")))
6 dataset <- createEntity(Dataset(list(name="Results Dataset", parentId = propertyValue(proj
7 ###
8
9 myDataset <- propertyValue(dataset, "id")
10
11 ## download the sawyers phenotypes and list the contents
12 sawyersPhenotypes <- loadEntity(4635)
13 names(sawyersPhenotypes)
14 ls(sawyersPhenotypes$objects)
15
16 ## pull out the data and modify it
17 ptCopy <- sawyersPhenotypes$objects$phenotypes
18 ptCopy$gleasonScore <- "fakeScore"
19
20
```

2:1 (Top Level) R Script

Workspace History

Load Save Import Dataset Clear All

Values

entityId	15627
layer	PhenotypeLayer [1]
parent	Dataset [1]

Files Plots Packages Help

Install Packages Check for Updates

affy	Methods for Affymetrix Oligonucleotide Arrays	
affydata	Affymetrix Data for Demonstration Purpose	
affyio	Tools for parsing Affymetrix data files	
affyPLM	Methods for fitting probe-level models	
affyQCReport	QC Report Generation for affyBatch objects	
annaffy	Annotation tools for Affymetrix biological metadata	
annotate	Annotation for microarrays	
AnnotationDbi	Annotation Database Interface	
Biobase	Biobase: Base functions for Bioconductor	
biomaRt	Interface to BioMart databases (e.g. Ensembl, COSMIC, Wormbase and Gramene)	
Biostrings	String objects representing biological sequences, and matching algorithms	
bitops	Functions for Bitwise operations	
boot	Bootstrap Functions (originally by Angelo Canty for S)	
caret	Classification and Regression Training	
class	Functions for Classification	
cluster	Cluster Analysis Extended Rousseeuw et al.	
codetools	Code Analysis Tools for R	
compiler	The R Compiler Package	
<input checked="" type="checkbox"/> datasets	The R Datasets Package	

Console ~/R/

Copyright (C) 2011 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: x86_64-pc-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[workspace loaded from ~/R/.RData]

```
> library(synapseclient)
>
```


RStudio

File Edit View Workspace Plots Tools Help

signELUA.R x metaGeo.R x DemoScript.R x

Source on Save Run Source

7% Welcome! Please login.

 **Synapse**
CONTRIBUTE to the CURE

Email Address

Password

Login Cancel

```
1 library(synapseclient)
2 synapseLogin()
3
4 ### connect to synapse
5 project = synapse.resolveProjectId("synapse")
6 dataset = synapse.resolveDatasetId("synapse", parentId = propertyvalue(project, "datasetId"))
7
8
9 myData = synapse.resolveDatasetId("synapse", parentId = propertyvalue(project, "datasetId"))
10
11 ## download data
12 sawyers = synapse.resolveDatasetId("synapse", parentId = propertyvalue(project, "datasetId"))
13 names = synapse.resolveDatasetId("synapse", parentId = propertyvalue(project, "datasetId"))
14 ls(sawyers)
15
16 ## pull out the data and modify it
17 ptCopy <- sawyersPhenotypes$objects$phenotypes
18 ptCopy$gleasonScore <- "fakeScore"
19
20
```

4:1 (Top Level) R Script

Console ~/R/

Platform: x86_64-pc-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[workspace loaded from ~/R/.RData]

```
> library(synapseclient)
> synapseLogin()
Loading Tcl/Tk interface ... done
```

Workspace History

Load Save Import Dataset Clear All

Values

entityId	15627
layer	PhenotypeLayer [1]
parent	Dataset [1]

Files Plots Packages Help

Install Packages Check for Updates

affy	Methods for Affymetrix Oligonucleotide Arrays
affydata	Affymetrix Data for Demonstration Purpose
affyio	Tools for parsing Affymetrix data files
affyPLM	Methods for fitting probe-level models
affyQCReport	QC Report Generation for affyBatch objects
annaffy	Annotation tools for Affymetrix biological metadata
annotate	Annotation for microarrays
AnnotationDbi	Annotation Database Interface
Biobase	Biobase: Base functions for Bioconductor
biomaRt	Interface to BioMart databases (e.g. Ensembl, COSMIC, Wormbase and Gramene)
Biostrings	String objects representing biological sequences, and matching algorithms
bitops	Functions for Bitwise operations
boot	Bootstrap Functions (originally by Angelo Canty for S)
caret	Classification and Regression Training
class	Functions for Classification
cluster	Cluster Analysis Extended Rousseeuw et al.
codetools	Code Analysis Tools for R
compiler	The R Compiler Package
<input checked="" type="checkbox"/> datasets	The R Datasets Package