

ProtPlot – A Tissue Molecular Anatomy Program Java-based Data Mining Tool: Screen Shots

**** DRAFT - undergoing revision ****

Peter F. Lemkin, Ph.D.⁽¹⁾, Djamel Medjahed Ph.D.⁽²⁾

¹NCI-Frederick; ²SAIC-Frederick, MD 21702

Home page: <http://tmap.sourceforge.net/>

Revised: 08-26-2004 Version 0.40.1-Beta



Abstract

- ProtPlot is an open-source Java-based data mining bioinformatic tool for analyzing CGAP- database derived estimated mRNA tissue EST expression in terms of a set of virtual 2D-gels
- The estimated mRNA expression is mapped to estimated “proteins”
- It is well known, mRNA expression generally does *not* correlate well with protein expression as seen in 2D-PAGE gels (Ideker et.al., *Science* 292:929-934, 2001)
- ProtPlot lets you look at the data in new ways and may help in thinking about new hypotheses for protein post-modifications or mRNA post-transcription processing.

Possible Questions

- **ProtPlot may help look at aggregates of CGAP data in new ways:**
 - Which “estimated proteins” are in a particular (pI,Mw) range?
 - Which sets of “proteins” are up or down regulated in cancer(s) and normal(s) or precancer(s)?
 - Which sets of “proteins” are entirely missing in one condition vs. the other?
 - Which sets of “proteins” cluster together across different types of cancers or normals?

ProtPlot

- It was developed initially as Virtual-2D [Proteomics J, in press], and upcoming paper on TMAP [Proteomics, in press]
- ProtPlot was derived from an open-source microarray data mining tool MAExplorer (<http://maexplorer.sourceforge.net/>) by P. Lemkin
- ProtPlot is a Java application and runs on your computer. You download and install the application and the data.

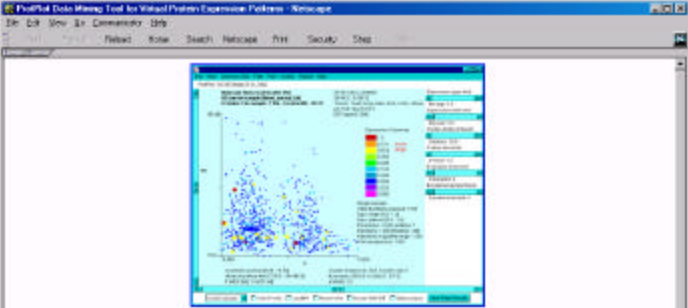
Pseudo 2D-Gel Map Expression Data

- Sample mRNA estimated expression data was obtained for a variety of human tissue and histology types (normal, pre-cancer, cancer) using the relative hit rates on cDNA clone libraries. Data from multiple libraries/tissue were merged
- Pseudo-protein data was computed by mapping the UniGene Ids in the CGAP libraries to SwissProt AC. The (pI, Mw) was computed using the SwissProt (pI,Mw) server tool
- These data are assembled into ProfPlot data files called .prp files described on the Web site.
- ProfPlot then generates an interactive pseudo 2D-gel Map (pIe,Mw) scatterplot that may be used for data mining

VIRTUAL2D home: <http://proteom.ncifcrf.gov/>

The screenshot displays the VIRTUAL2D website interface. On the left is a yellow sidebar with a 'TABLE OF CONTENTS' menu listing sections such as 'Introduction', '2D PAGE', 'EXPERIMENTAL 2D PAGE MAP', 'IDENTIFICATION OF MOLECULAR TARGETS', 'INTERACTIVE 2D PAGE MAP', 'SPOT-PROTEIN ASSIGNMENT', and 'EXPRESSION PROFILING'. The main content area features a dark background with a starry pattern and a central diagram. The diagram is titled 'VIRTUAL 2D' and shows a flow from 'GENOME' to 'PROTEOME', which then branches into 'EXPERIMENTAL 2D PAGE MAP' and 'INTERACTIVE 2D PAGE MAP'. From 'INTERACTIVE 2D PAGE MAP', the flow continues to 'SPOT-PROTEIN ASSIGNMENT', which then leads to 'EXPRESSION PROFILING' and finally 'IDENTIFICATION OF MOLECULAR TARGETS'.

TMAP HOME: <http://tmap.sourceforge.net/>



ProtPlot Data Mining Tool for Virtual Protein Expression Patterns

[Home](#) | [Introduction](#) | [Using ProtPlot](#) | [News](#) | [Estimating expression](#) | [Downloads](#) | [FAQ](#) | [Feedback](#) | [PDF documents](#) | [History of ProtPlot](#) | [Latest version](#)


Introduction

ProtPlot is a Java-based data-mining software tool for virtual 2D gels. It may be downloaded and run as a stand-alone application on your computer. Its exploratory data analysis environment provides tools for the data-mining of quantified virtual 2D gel (MTF.gel, expression) data of estimated protein expression from the CGAP mRNA expression database.

ProtPlot generates an interactive pseudo-protein 2D Map structured based on a database of [defined sequence EST sequences](#) over a variety of tissue types from data obtained from the NCI-NCI's CGAP EST Database of human cancer, prostatic and cancer mRNA sequences (CGAP is the NCI's Cancer Genome Anatomy Project <http://ncap.nci.nih.gov/>). EST is the Expressed Sequence Tag of a mRNA found in a particular tissue. The EST hit rate is a rough estimate of gene expression. These ESTs were mapped to [GENE_PRODUCT \(geneid.org\)](#) accession numbers and IGI, the M5v and p1 estimates were computed and used as estimates for correct coding proteins in a pseudo-2D-gel.

NOTE: in general, mRNA expression will not correspond quantitatively to measured protein expression because most proteins are extensively post-modified as well as many mRNAs being processed in various ways. However, it may be useful to see the levels of common gene expression across normal, prostatic and cancer.

History of ProtPlot



History ProtPlot

[TMAP home](#) | [Introduction](#) | [Using ProtPlot](#) | [News](#) | [Estimating expression](#) | [Downloads](#) | [FAQ](#) | [Feedback](#) | [PDF documents](#) | [History of ProtPlot](#) | [Latest version](#)

The new version of ProtPlot as part of the Tissue Molecular Anatomy Project was derived from an earlier version of ProtPlot as discussed in a paper on VIRTUAL2D: [O'Medjhed D, 2003a](#). A [poster](#) was given at the 2002 Swiss Proteomics Meeting in Geneva on TMAP. This has been submitted to the proceedings of the meeting: [O'Medjhed D, 2003b](#).

This version of ProtPlot is run as a stand-alone Java program. After ProtPlot is downloaded and installed from the Web server at <http://www.leitch.net/~tmap/TMAP/>, the user starts up the program the default startup icon or clicks on a specific .proptext file to start it on a previous data-mining session. You may have several copies running simultaneously if you have enough memory.

This work was produced by Peter Lenkin of the National Cancer Institute, an agency of the United States Government and Djamel Medjhed (SAJC-Fredrick). As a work of the United States Government there is no associated copyright.

ProtPlot will be offered as open source software under the [Mozilla Public License \(version 1.1\)](#) subject to the limitations noted in the accompanying [LEGAL file](#).

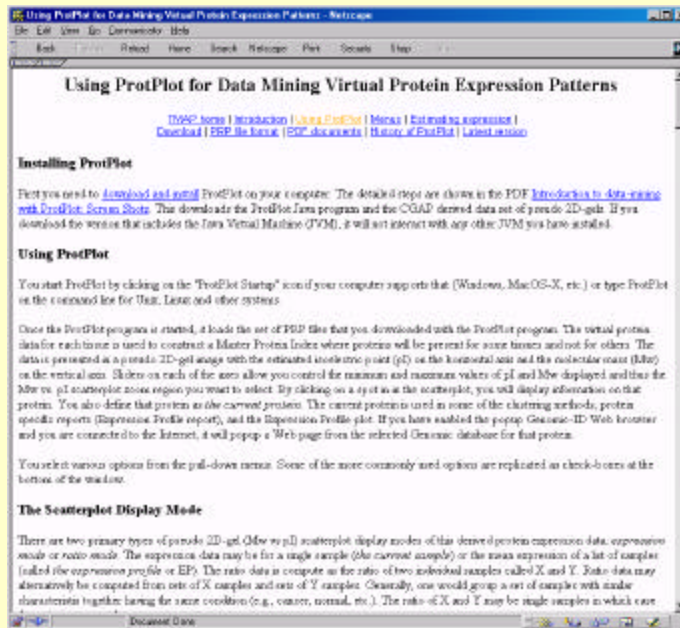
This work was derived from MAExplorer (<http://maexplorer.sourceforge.net/>) under the Mozilla 1.1 Open Source Public License by Peter Lenkin of the National Cancer Institute.

See [disclaimer](#) for more information on its usage.

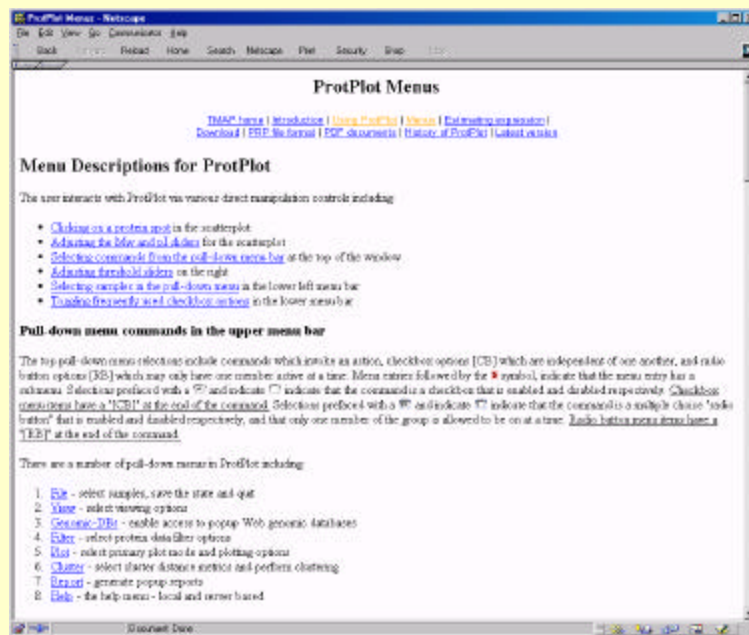
*Djamel Medjhed, LMU SAJC-Fredrick
Peter Lenkin, LBCA, NCI-Fredrick*

Revised: 02-07-2003

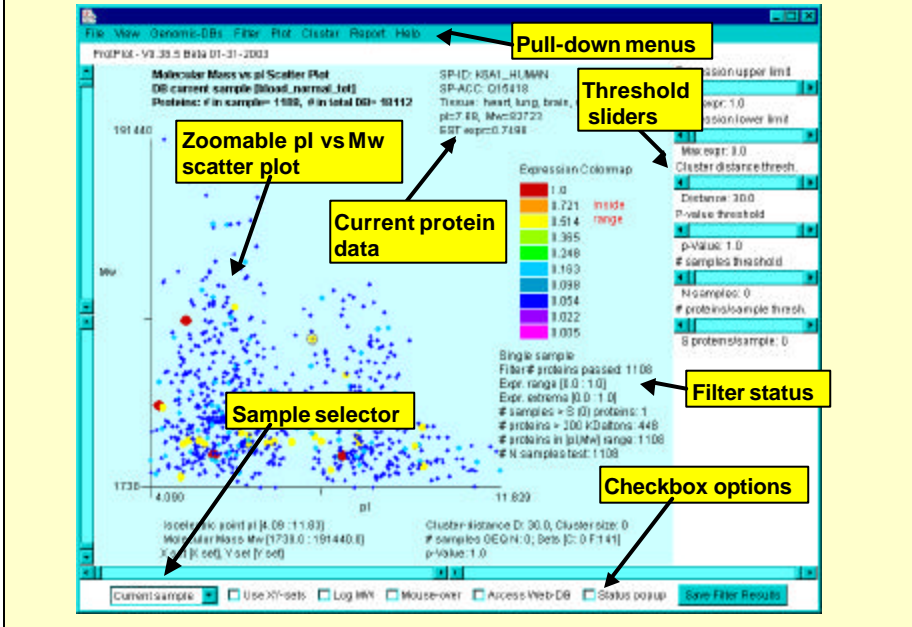
Using ProtPlot



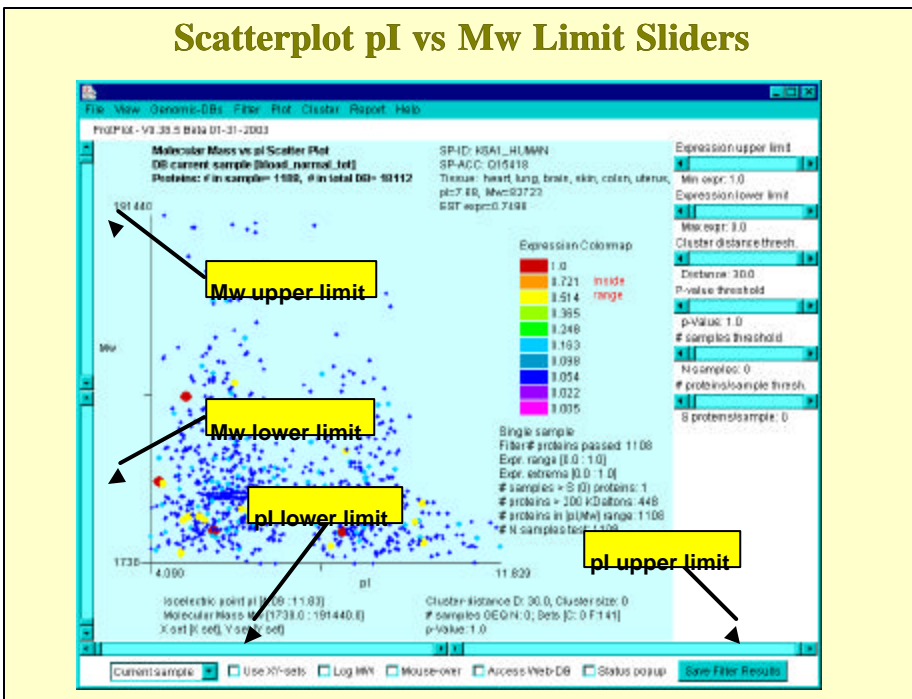
ProtPlot Menus and User Controls



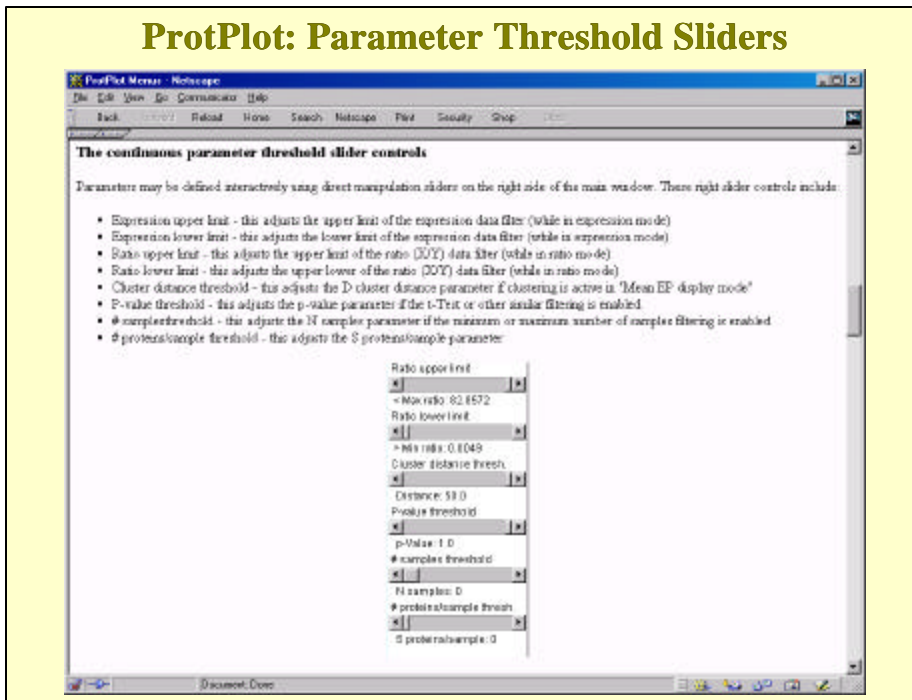
Initial Screen displaying (pI vs Mw) scatterplot



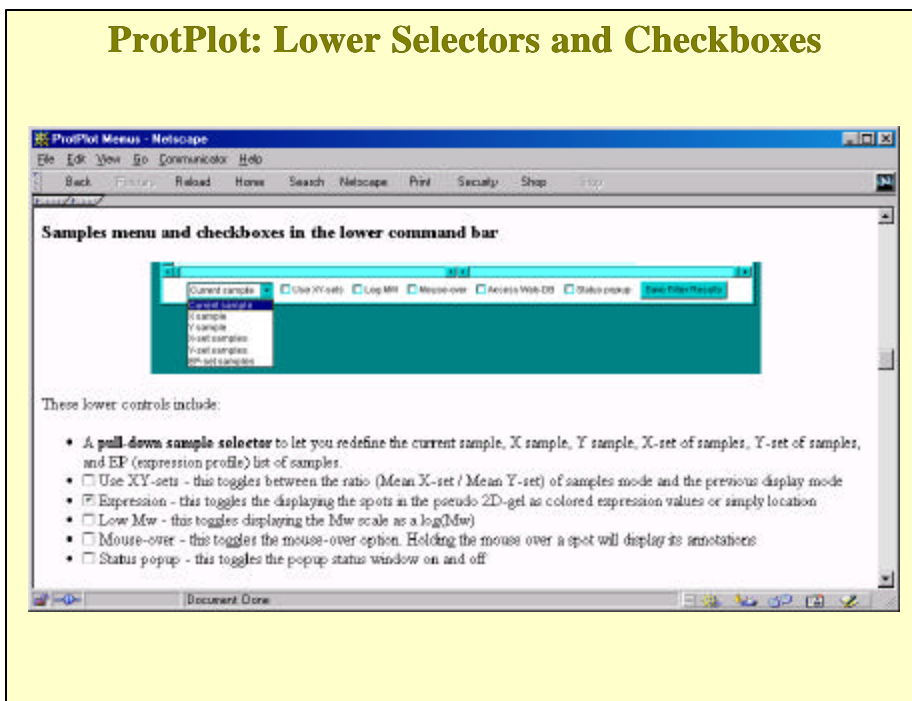
Scatterplot pI vs Mw Limit Sliders



ProtPlot: Parameter Threshold Sliders



ProtPlot: Lower Selectors and Checkboxes



Download ProtPlot

Downloads for The MicroArray Explorer Project

[TMAP home](#) | [Introduction](#) | [Using ProtPlot](#) | [Menus](#) | [Estimating expression](#) | [Download](#) | [PRP file format](#) | [PDF documents](#) | [History of ProtPlot](#) | [Latest version](#)

[Program downloads](#) | [Update programs](#) | [Update jar file](#) | [Installation hints](#)

The table below lists the various types of downloads: program installer, source code file, jar file, and information on installing the programs.

Types of download files available

You may download program installers for your particular computer for both ProtPlot. Click on the entries to download the installer or files.

Access of ProtPlot from the Web server

Program	Installer Version	Update Program Jar Version	Program installers	Jar file(s)	Source	PRP data
ProtPlot	0.39.1	0.39.1	ProtPlot	ProtPlot.jar	N.A. yet	PRP data

Click on program installers

Download ProtPlot Installer

Version 0.38.5 - ProtPlot application for TMAP (Tissue Molecular Anatomy Project)

Recommended Installation for Your Platform:
[Download of ProtPlot for Windows](#) | [Include VM in download](#)

Available Installers

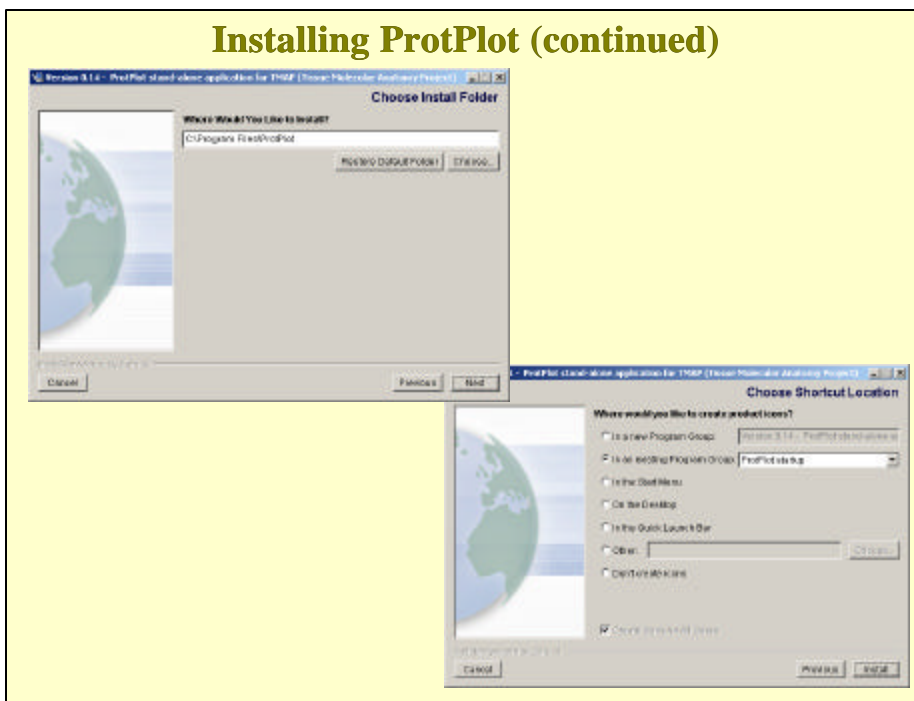
Platform	includes Java VM	without Java VM	Instructions
Windows	Download (173.19K)	Download (2.09M)	View
Misc OS	Download (7.1M)	Download (2.26M)	View
Misc OS X		Download (2.26M)	View
AKK		Download (2.26M)	View
Solaris	Download (132.26K)	Download (2.26M)	View
Linux	Download (134.69K)	Download (2.26M)	View
HP-UX		Download (2.26M)	View
UNIX with VM		Download (2.26M)	View
Any Unix Platform		Download (2.26M)	View
Other Java-enabled Platforms		Download (2.44M)	View

Click on Download button

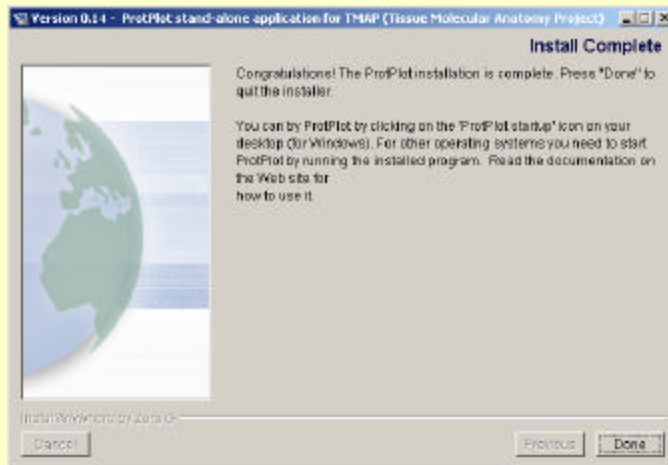
Installing ProtPlot



Installing ProtPlot (continued)

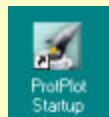


Finished Downloading ProtPlot

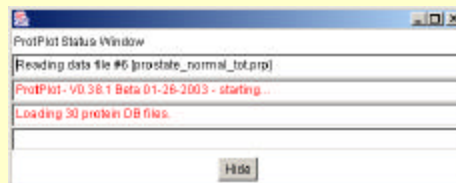


Starting ProtPlot - Click on the Startup Icon or Use the Start menu

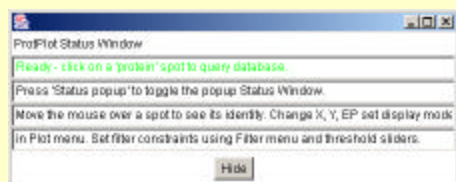
A) Click on ProtPlot Startup icon



B) Displays the loading status



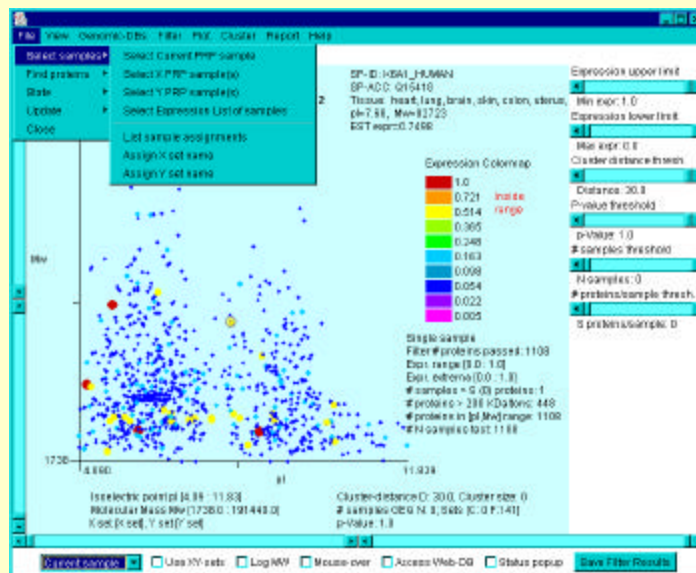
C) Press Hide button to remove



ProtPlot Menus

- **File** - select samples, save the state and quit
- **View** - select viewing options
- **Genomic-DBs** - enable access to popup Web genomic databases
- **Filter** - select protein data filter options
- **Plot** - select primary data mining and scatterplot display options
- **Cluster** - select cluster distance metrics and perform clustering
- **Report** - generate popup reports
- **Help** - popup help menu

File Menu - Selecting Single Samples, the X-set, Y-set or EP-set of Samples



File Menu - Selecting Samples using Choice Menu

The screenshot shows a software interface with a 'File' menu open. The menu is titled 'Select new X sample' and contains a list of sample types with checkboxes. A yellow box labeled 'Sample selector (picked X set)' points to the 'prostate cancer tot' option. Another yellow box labeled 'Pick specific sample or sam' points to the 'Set' button at the bottom of the menu. The background shows a scatter plot with axes labeled 'log2' and 'p'. The plot has a title 'Molecular Mass vs log Scatter Plot' and a subtitle 'DE: cancer vs normal (DE: total, normal, Prostate X in contrast - 1958, F14)'. The plot shows a cluster of points with a red dot and a yellow dot. The status bar at the bottom shows 'Cluster: 1958, Cluster size: 0'.

Selecting Subsets of Samples for Experiments

- **Current Sample** - to look at the expression for any individual sample. E.g., *prostate_cancer*
- **Sample X and Sample Y** - to look at the ratio of $exprX/exprY$ where the protein for which the ratio is defined has expression in both the X and Y individual samples. E.g., X is *prostate_cancer* and Y is *prostate_normal*
- **X set of samples and Y set of samples** - to look at the ratio of $Mean-exprX / Mean-exprY$ where the protein for which the ratio is defined has expression in both the X and Y samples for at least 1 sample in X and at least 1 in Y. E.g., X set is all *cancer* and Y is all *normal*
- **Expression Profile set of samples** - to look at the expression profile (EP plot or EP report) for any protein. The scatter plot shows mean EP expression. E.g., EP is all samples, or EP is all *cancer*, etc.

Plot Display Mode Rules

All proteins in the Master Protein Index (mPid) are displayed except for the following:

- In single sample or EP expression mode, do not show missing proteins
- In X/Y sample mode, do not show proteins that are missing in X but present in Y or vice versa. However, if the View option to display this missing data is enabled, then show the missing data as gray spots.
- In X-set/Y-set samples mode, do not show proteins unless they meet the sizing criteria N for both X and Y if enable or if using the missing sets > N filter.
- Normally, plot proteins in a (**Mw vs. pI**) scatterplot
- If in one of the X/Y ratio modes, may plot (**X vs. Y**) expression scatterplot instead of (Mw vs. pI)

Selecting the Current Sample (those with [**>S**] have more than S proteins/sample)

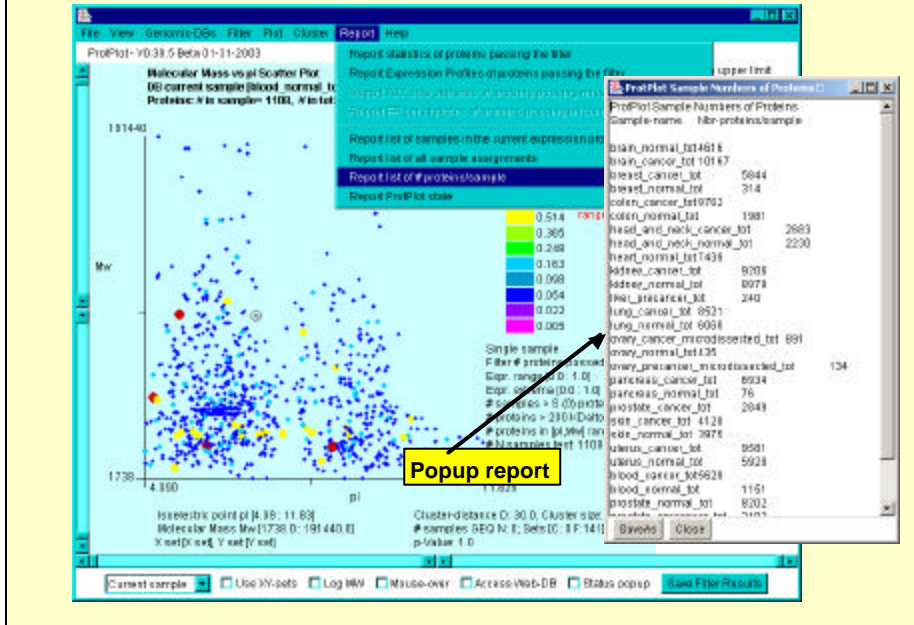
The screenshot shows a dialog box titled "Select new current sample" with a list of checkboxes for various sample types. The list includes:

- brain normal tot [**>S**]
- brain cancer tot [**>S**]
- breast cancer tot [**>S**]
- breast normal tot
- colon cancer tot [**>S**]
- colon normal tot [**>S**]
- head and neck cancer tot [**>S**]
- head and neck normal tot [**>S**]
- heart normal tot [**>S**]
- kidney cancer tot [**>S**]
- kidney normal tot [**>S**]
- liver precancer tot
- lung cancer tot [**>S**]
- lung normal tot [**>S**]
- ovary cancer microdissected tot
- ovary normal tot
- ovary precancer microdissected tot
- pancreas cancer tot [**>S**]
- pancreas normal tot
- prostate cancer tot [**>S**]
- prostate precancer tot [**>S**]
- skin cancer tot [**>S**]
- skin normal tot [**>S**]
- uterus cancer tot [**>S**]
- uterus normal tot [**>S**]
- blood normal tot [**>S**]
- prostate normal tot [**>S**]
- cervix normal tot
- cervix cancer tot

Annotations:

- A yellow box with the text "Slider to set S the # proteins/sample for the sample to be used" points to a slider control labeled "# proteins/sample thresh." with a value of 1000.
- A yellow box with the text "Pick specific sample" points to the "blood normal tot [**>S**]" checkbox.

Report Menu - Listing # Proteins in All Samples



Selecting the X Sample

The screenshot shows the 'Select new X sample' dialog box. The dialog box contains a list of sample names with checkboxes. The 'head and neck cancer tot (>5)' checkbox is checked. The list includes samples like 'uterus normal tot (>5)', 'skin normal tot (>5)', 'prostate precancer tot (>3)', 'prostate cancer tot (>5)', 'pancreas cancer tot (>8)', 'ovary normal tot', 'lung normal tot (>5)', 'liver precancer tot', 'kidney cancer tot (>8)', 'head and neck normal tot (>3)', 'color normal tot (>5)', 'cervix normal tot', 'breast normal tot', 'brain normal tot (>5)', 'blood normal tot (>5)', 'uterus cancer tot (>5)', 'skin cancer tot (>3)', 'prostate normal tot (>3)', 'pancreas normal tot', 'ovary precancer microdissected tot', 'ovary cancer microdissected tot', 'lung cancer tot (>5)', 'kidney normal tot (>8)', 'heart normal tot (>5)', 'head and neck cancer tot (>5)', 'colon cancer tot (>5)', 'cervix cancer tot', 'breast cancer tot (>5)', 'brain cancer tot (>5)', and 'blood cancer tot (>5)'. The dialog box also includes buttons for 'Define X sample set else single X sample', 'Reset', 'Clear all', 'Get normal', 'Get precancer', 'Get cancer', and 'Close'.

Selecting the X-set of Samples

Select new X sample

<input type="checkbox"/> uterus normal tot [-S]	<input checked="" type="checkbox"/> uterus cancer tot [-S]
<input type="checkbox"/> skin normal tot [-S]	<input checked="" type="checkbox"/> skin cancer tot [-S]
<input checked="" type="checkbox"/> prostate precancer tot [-S]	<input type="checkbox"/> prostate normal tot [-S]
<input checked="" type="checkbox"/> prostate cancer tot [-S]	<input type="checkbox"/> pancreas normal tot
<input checked="" type="checkbox"/> pancreas cancer tot [-S]	<input checked="" type="checkbox"/> ovary precancer microdissected tot
<input type="checkbox"/> ovary normal tot	<input checked="" type="checkbox"/> ovary cancer microdissected tot
<input type="checkbox"/> lung normal tot [-S]	<input checked="" type="checkbox"/> lung cancer tot [-S]
<input checked="" type="checkbox"/> liver precancer tot	<input type="checkbox"/> kidney normal tot [-S]
<input checked="" type="checkbox"/> kidney cancer tot [-S]	<input type="checkbox"/> heart normal tot [-S]
<input type="checkbox"/> head and neck normal tot [-S]	<input checked="" type="checkbox"/> head and neck cancer tot [-S]
<input type="checkbox"/> colon normal tot [-S]	<input checked="" type="checkbox"/> colon cancer tot [-S]
<input type="checkbox"/> cervix normal tot	<input checked="" type="checkbox"/> cervix cancer tot
<input type="checkbox"/> breast normal tot	<input checked="" type="checkbox"/> breast cancer tot [-S]
<input type="checkbox"/> brain normal tot [-S]	<input checked="" type="checkbox"/> brain cancer tot [-S]
<input type="checkbox"/> blood normal tot [-S]	<input checked="" type="checkbox"/> blood cancer tot [-S]

Define X sample set else single X sample

Pick multiple samples

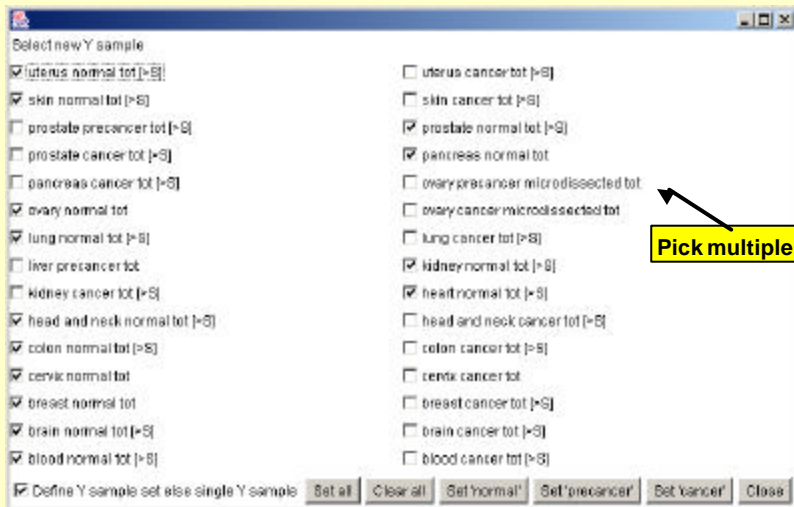
Selecting the Y Sample

Select new Y sample

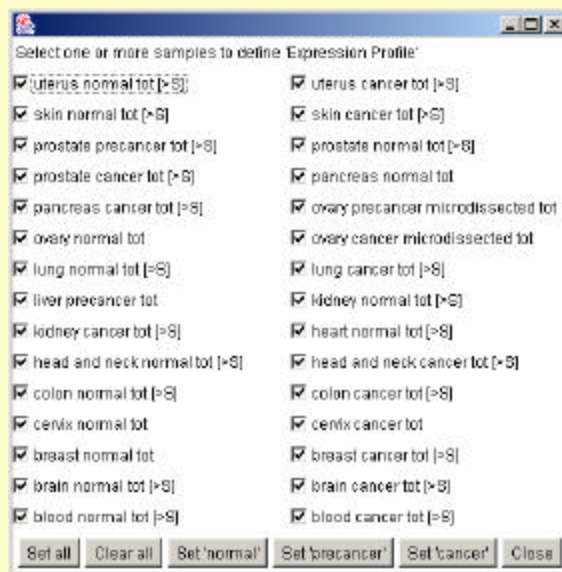
<input type="checkbox"/> uterus normal tot [-S]	<input type="checkbox"/> uterus cancer tot [-S]
<input type="checkbox"/> skin normal tot [-S]	<input type="checkbox"/> skin cancer tot [-S]
<input type="checkbox"/> prostate precancer tot [-S]	<input checked="" type="checkbox"/> prostate normal tot [-S]
<input type="checkbox"/> prostate cancer tot [-S]	<input type="checkbox"/> pancreas normal tot
<input type="checkbox"/> pancreas cancer tot [-S]	<input type="checkbox"/> ovary precancer microdissected tot
<input type="checkbox"/> ovary normal tot	<input type="checkbox"/> ovary cancer microdissected tot
<input type="checkbox"/> lung normal tot [-S]	<input type="checkbox"/> lung cancer tot [-S]
<input type="checkbox"/> liver precancer tot	<input type="checkbox"/> kidney normal tot [-S]
<input type="checkbox"/> kidney cancer tot [-S]	<input type="checkbox"/> heart normal tot [-S]
<input type="checkbox"/> head and neck normal tot [-S]	<input type="checkbox"/> head and neck cancer tot [-S]
<input type="checkbox"/> colon normal tot [-S]	<input type="checkbox"/> colon cancer tot [-S]
<input type="checkbox"/> cervix normal tot	<input type="checkbox"/> cervix cancer tot
<input type="checkbox"/> breast normal tot	<input type="checkbox"/> breast cancer tot [-S]
<input type="checkbox"/> brain normal tot [-S]	<input type="checkbox"/> brain cancer tot [-S]
<input type="checkbox"/> blood normal tot [-S]	<input type="checkbox"/> blood cancer tot [-S]

Define Y sample set else single Y sample

Selecting the Y-set of Samples



Selecting the Expression Profile (EP) Set of Samples



Listing Sample Assignments

Condition-Set	Sample-name	MIs-proteins
Current sample	blood_normal_bot	1151
X-sample	prostate_cancer_bot	2848
Y-sample	pancreas_normal_bot	76
X set: set	brain_cancer_bot_10187	
X set: set	breast_cancer_bot	5844
X set: set	colon_cancer_bot9762	
X set: set	head_and_neck_cancer_bot	2863
X set: set	kidney_cancer_bot	9209
X set: set	lung_cancer_bot_9521	
X set: set	ovary_cancer_mitochondria_bot	891
X set: set	pancreas_cancer_bot	6934
X set: set	prostate_cancer_bot	2848
X set: set	skin_cancer_bot_4128	
X set: set	stomach_cancer_bot	9581
X set: set	blood_normal_bot5628	
X set: set	colon_cancer_bot1567	
Y set: set	brain_normal_bot4618	
Y set: set	breast_normal_bot	214
Y set: set	colon_normal_bot	1981
Y set: set	head_and_neck_normal_bot	2230
Y set: set	heart_normal_bot7438	
Y set: set	kidney_normal_bot	8879
Y set: set	lung_normal_bot_6068	
Y set: set	ovary_normal_bot435	
Y set: set	pancreas_normal_bot	76
Y set: set	skin_normal_bot_3975	
Y set: set	stomach_normal_bot	5929
Y set: set	blood_normal_bot	1151
Y set: set	prostate_normal_bot	8202
Y set: set	colon_normal_bot	878
Exp+Profile list	brain_normal_bot4618	
Exp+Profile list	brain_cancer_bot_10187	
Exp+Profile list	breast_cancer_bot	5844
Exp+Profile list	breast_normal_bot	214

Defining the X and Y Condition Set Names

A.1
(default X set)

Dialog
Enter Set X set condition name

X set

Ok Cancel

A.2
set to 'cancer'

Dialog
Enter Set X set condition name

Cancer

Ok Cancel

B.1
(default Y set)

Dialog
Enter Set Y set condition name

Y set

Ok Cancel

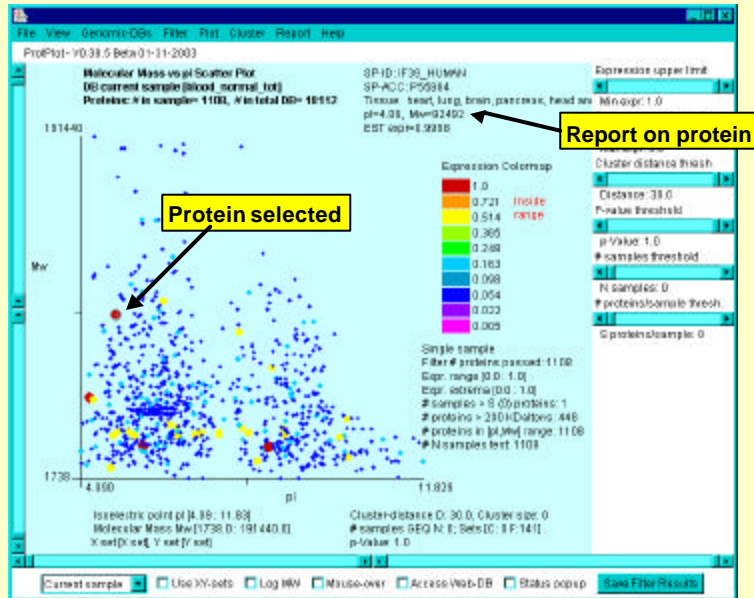
B.2
set to 'normal'

Dialog
Enter Set Y set condition name

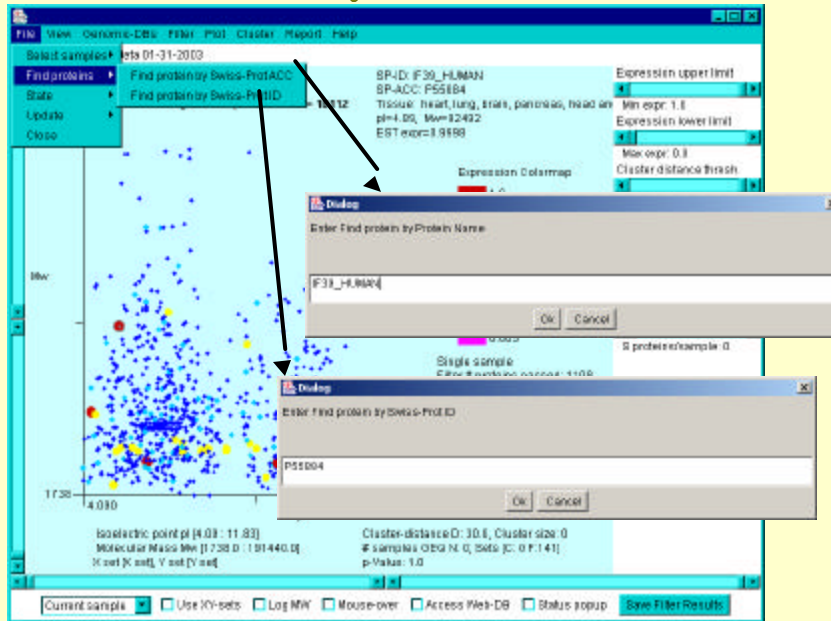
Normal

Ok Cancel

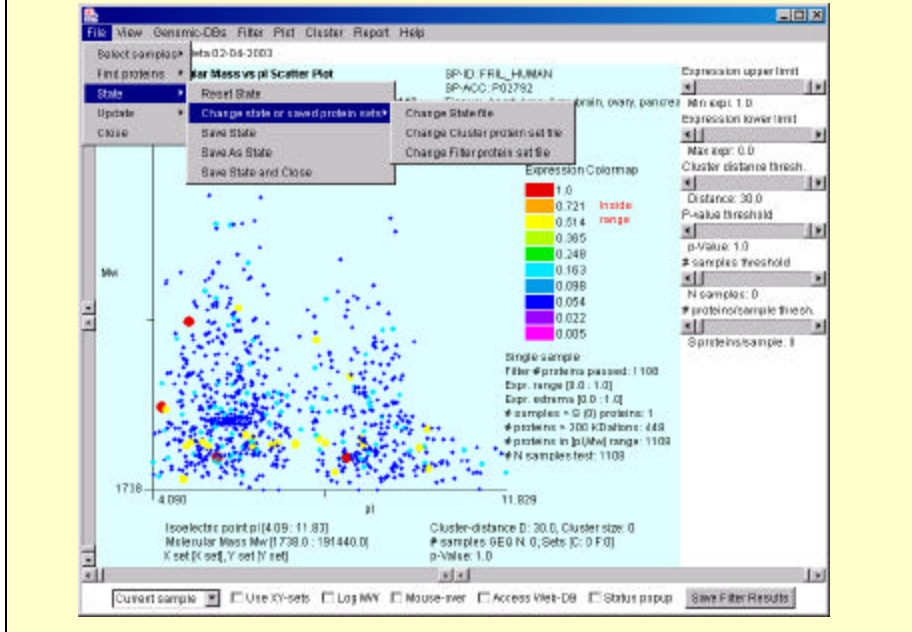
Click on a Spot to Select the Protein



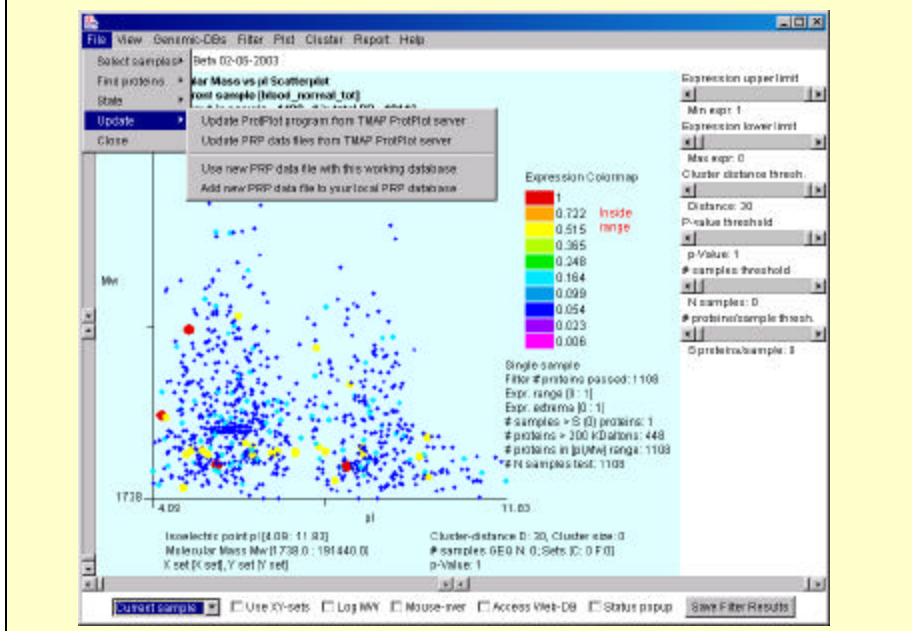
Select a Protein by SwissProt ID or ACC



File Menu - Save & Restore the Data Mining State



File Menu - Updating the Program and PRP data



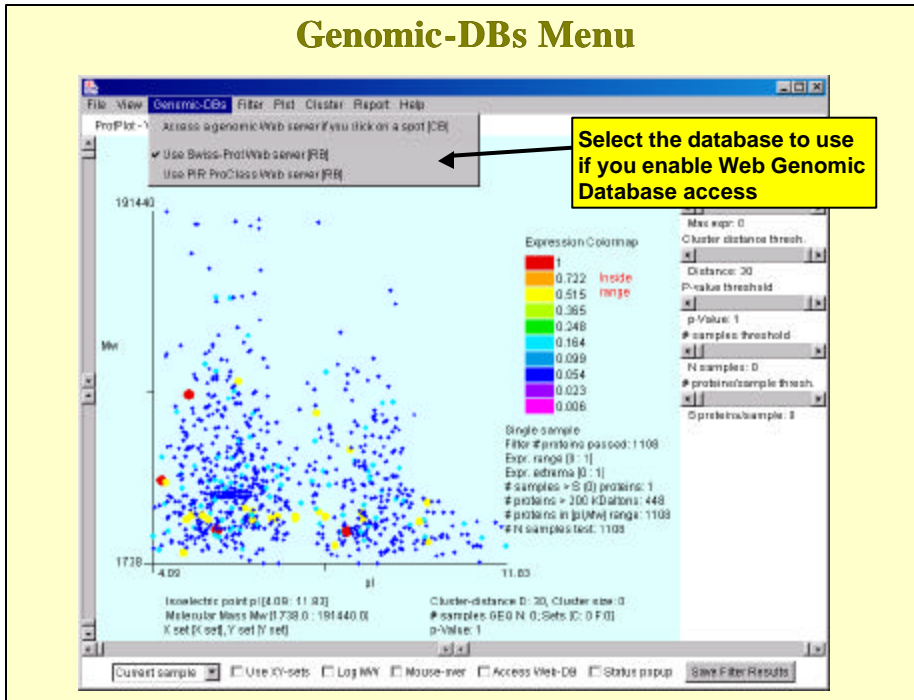
Updating ProtPlot Program from the Proteom Server

The screenshot shows the ProtPlot software interface. The 'Update' menu is open, displaying options: 'Update ProtPlot program from TMAP ProtPlot server', 'Update PRP data files from TMAP ProtPlot server', 'Use new PRP data file with this working database', and 'Add new PRP data file to your local PRP database'. A yellow callout box points to the 'Update ProtPlot program...' option with the text: 'Asks you to verify that you want to update the program'. Below the menu, a dialog box is open with the text: 'Update ProtPlot (v 0.38.1) from http://www.bioinformatics.ncf.gov/TMAP/ Web site - are you sure?' and 'Yes' and 'No' buttons. The background shows a scatter plot of Molecular Mass (Mw) vs. pI with a color-coded expression heatmap. The plot includes axes for Mw (17.20 to 4.00) and pI (4.00 to 11.00). A legend for 'Expression Colormap' shows a color scale from 1.0 (red) to 0.005 (magenta). The status bar at the bottom indicates '#N samples test: 1108'.

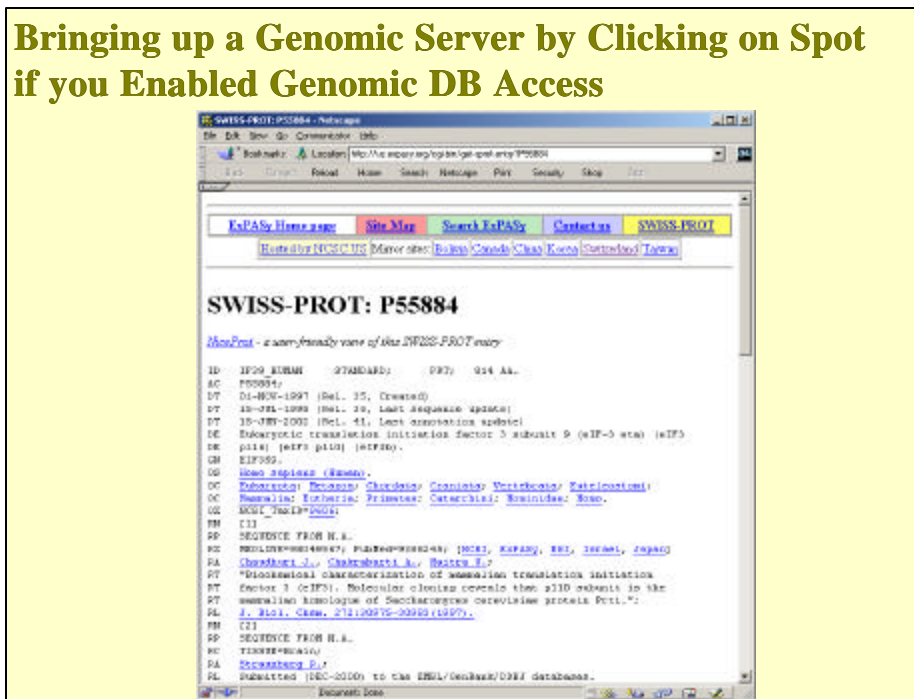
View Menu - Display Options

The screenshot shows the ProtPlot software interface with the 'View' menu open. The menu items are: 'Show Status Window', 'Hide Status Window', 'Use pseudocolor for expression else grayscale [CB]', 'Use log of Mw [CB]', 'Enable mouse-over to see data [CB]', 'Show expression data also just position [CB]', 'Use radius proportional to expression value constant [CB]', 'Show clustered proteins by black boxes [CB]', 'Show all X and Y samples if doing XY filtering [CB]', and 'Auto update reports if the state changes [CB]'. A yellow callout box points to the menu with the text: 'Modifies how data is displayed. Some of the options are also in the checkboxes below'. The background shows the same scatter plot as the previous image, but with a different color scale for the expression heatmap, ranging from 1.0 (red) to 0.005 (magenta). The status bar at the bottom indicates '#N samples test: 1108'.

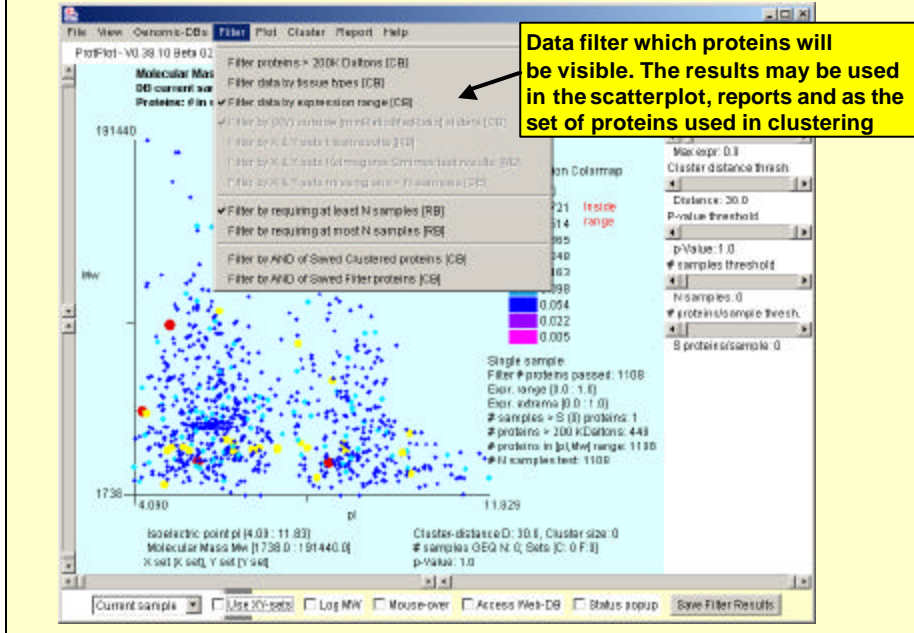
Genomic-DBs Menu



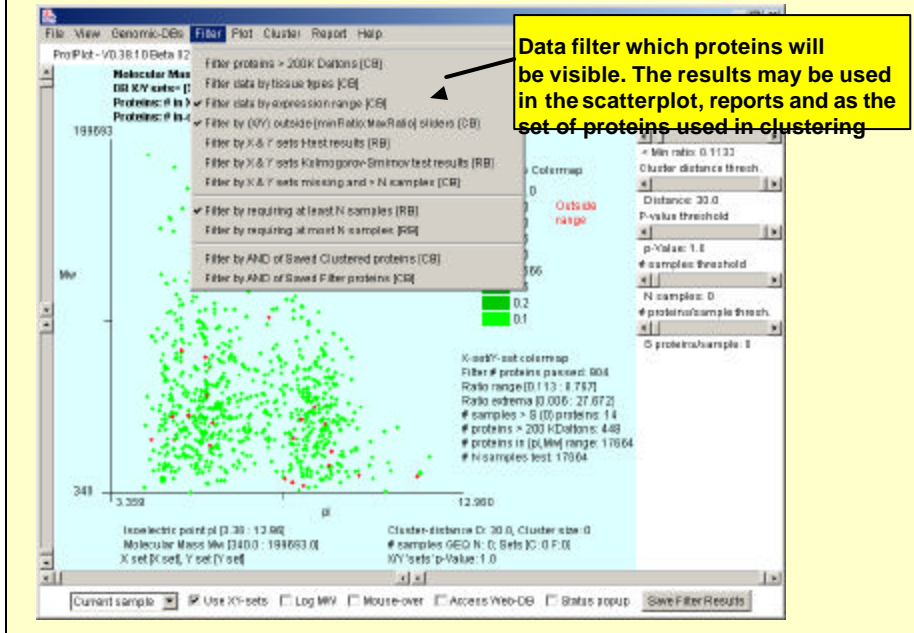
Bringing up a Genomic Server by Clicking on Spot if you Enabled Genomic DB Access



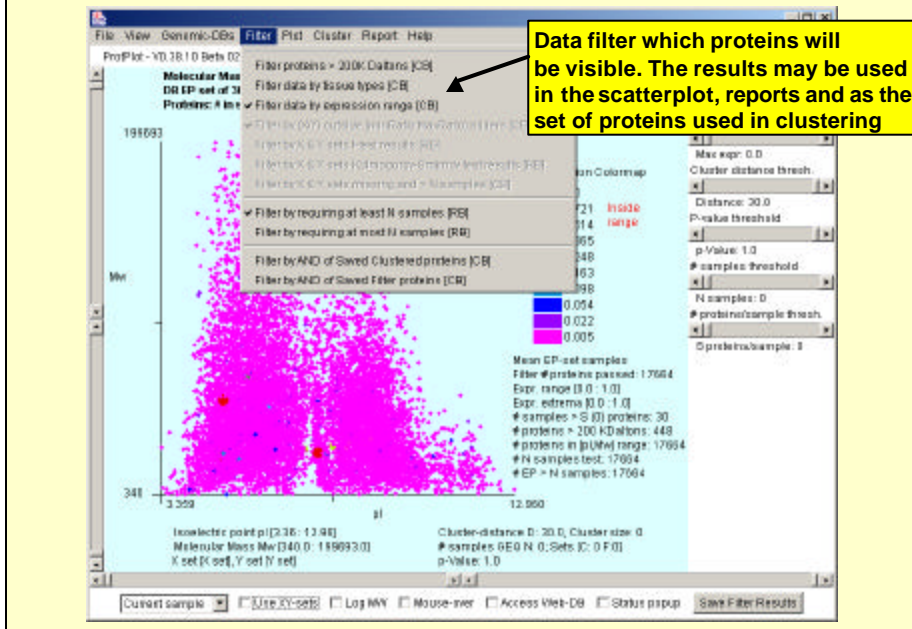
Filter Menu - Data Filter Options for Single Sample



Filter Menu - Data Filter Options for X/Y Ratio



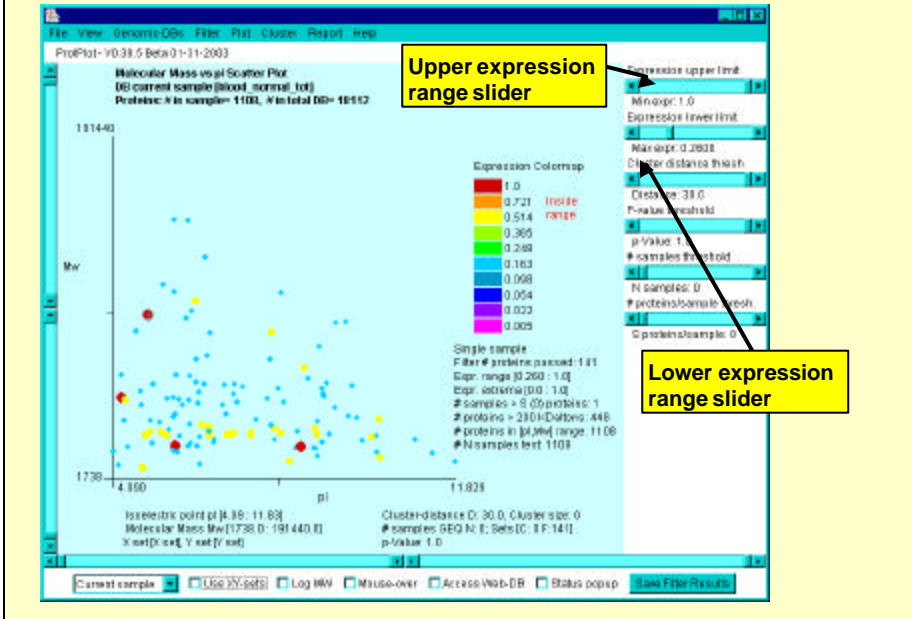
Filter Menu - Data Filter Options for EP-set



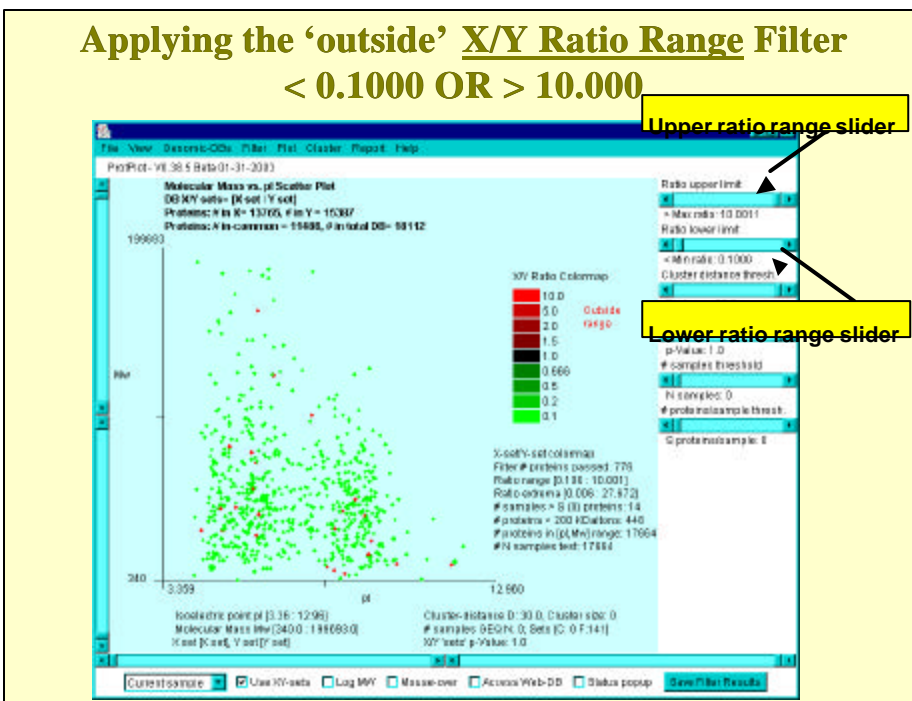
Filter Types - Available

- By Proteins > 200 Kdaltons , Mw and pI within ranges
- By tissue types
- By expression value range
- By expression X/Y ratio range (either inside or outside range)
- By t-Test of X-set and Y-Set samples < p-value threshold
- By min # samples in X & Y or EP sets > N samples threshold
- By missing proteins in X or Y set with other set > N samples threshold
- By number of samples for the protein > N samples threshold or < N samples threshold

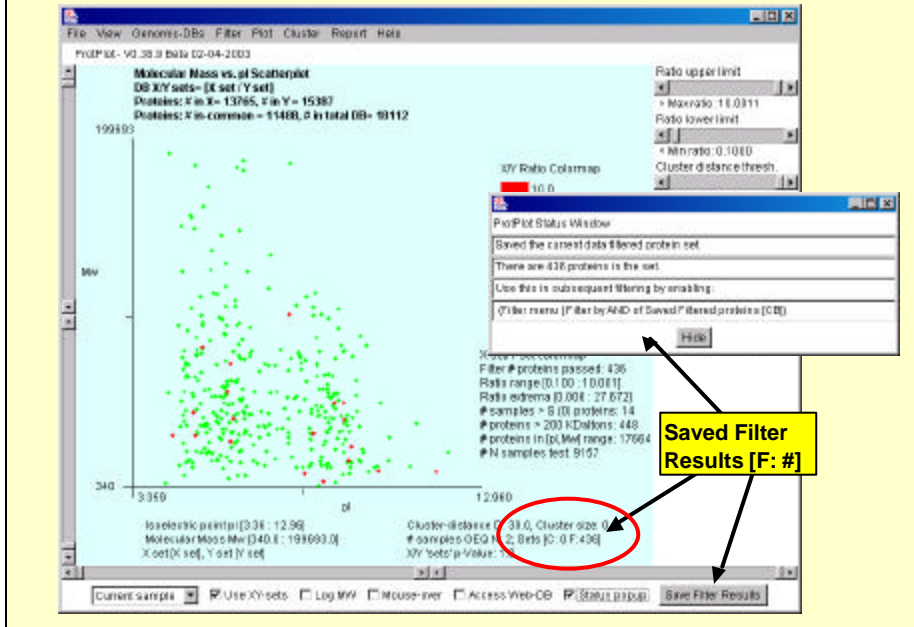
Applying Expression Range Filter [0.455 : 1.0]



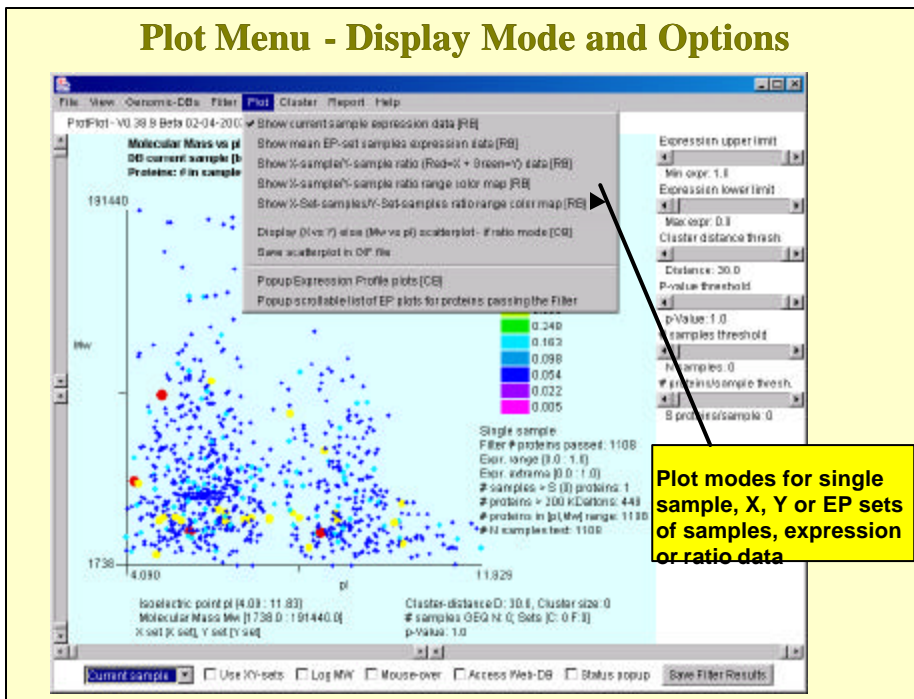
Applying the 'outside' X/Y Ratio Range Filter < 0.1000 OR > 10.000



Saving Filter Set of Proteins - For Future Filtering



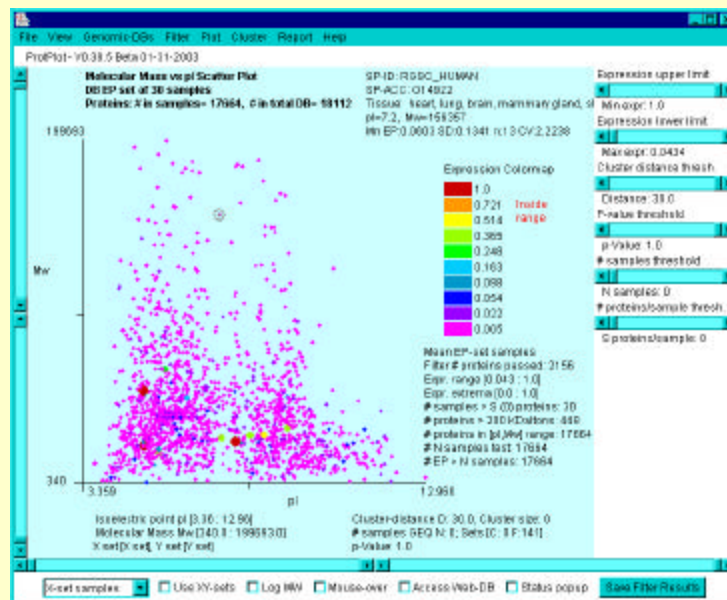
Plot Menu - Display Mode and Options



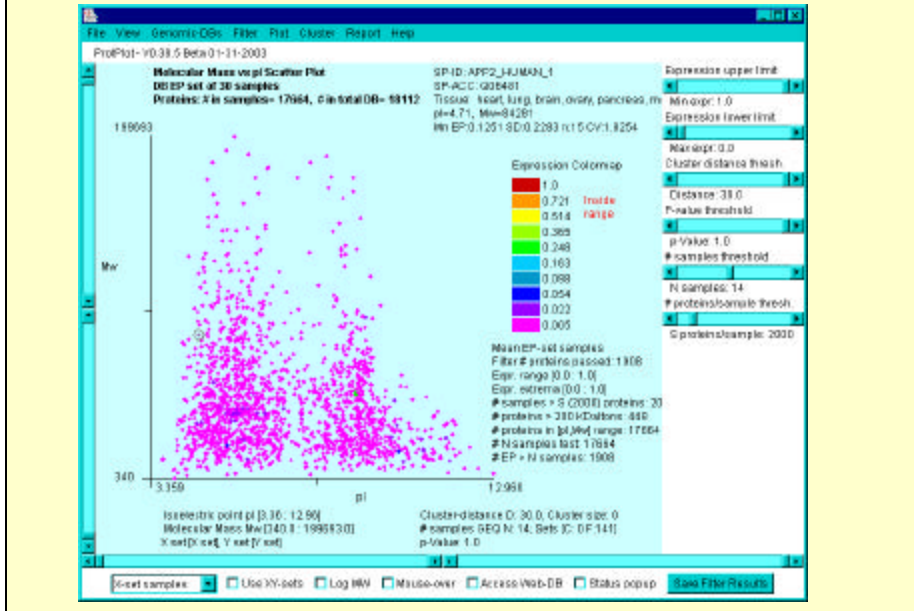
Plotting Display Modes

- **Show Current Sample** - to look at the **expression** for a single sample
- **Show Mean Expression-Profile set of samples** - to look at the **mean expression** for a subset of samples
- **Show X-Sample /Y-Sample Y** - to look at the **ratio** of two individual samples
- **Show X-set samples / Y-set samples** - to look at the **ratio** of $Mean\text{-}exprX / Mean\text{-}exprY$ for two **sets** of samples (X and Y sets)
- If in one of the X/Y ratio modes, may plot (**X vs Y**) expression scatterplot instead of default (**Mw vs. pI**) scatterplot

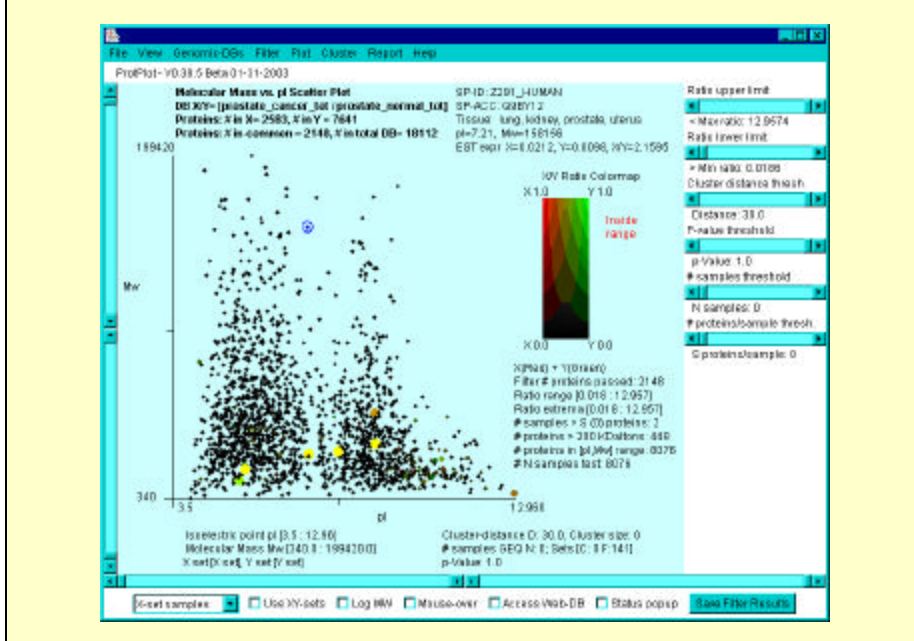
Plot Display Mode - Current Sample



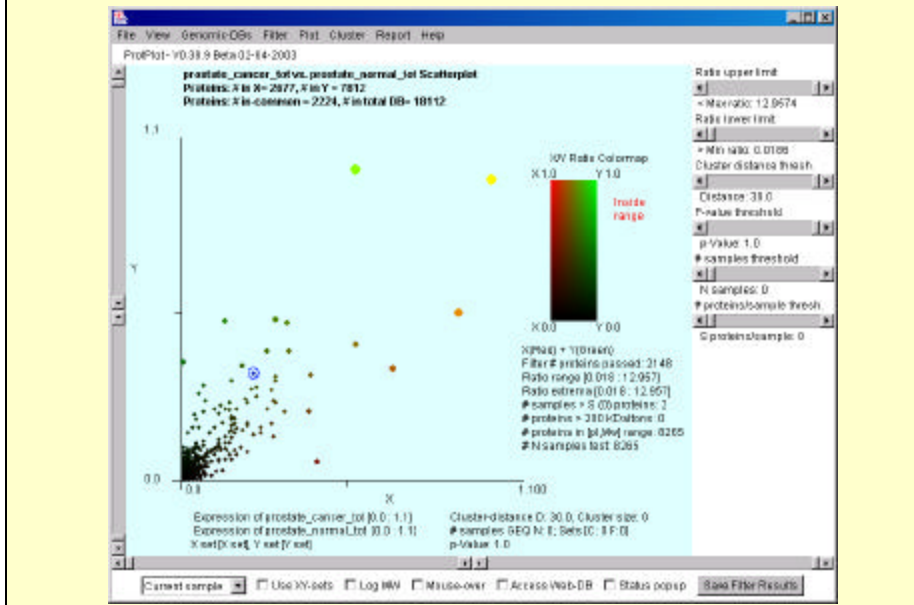
Plot Display Mode - Mean of EP Set of Samples (N >= 14, S >= 2000)



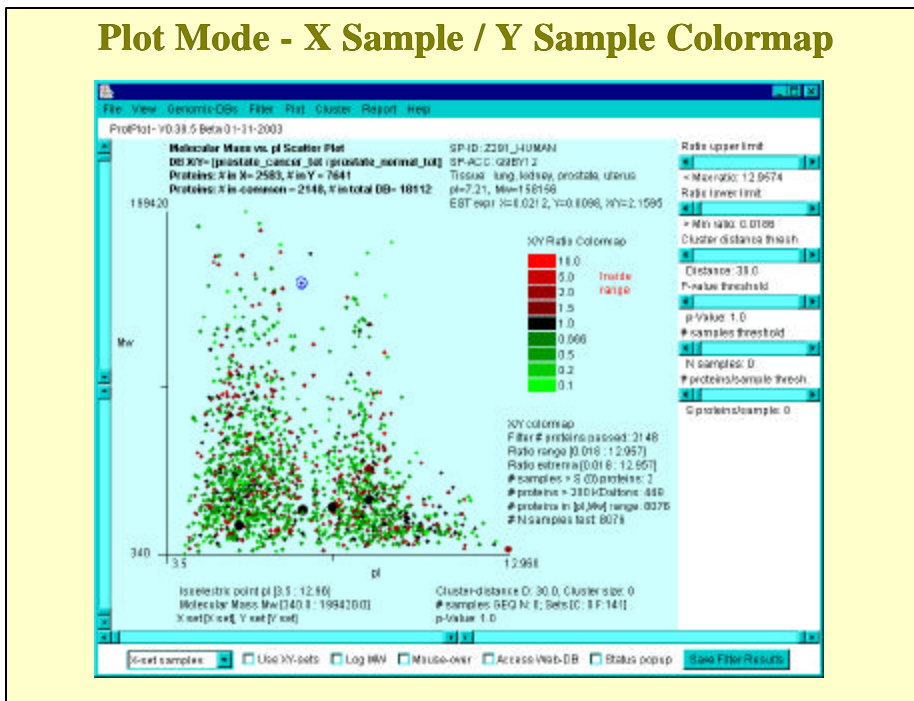
Plot Display Mode - X Sample (Red) + Y Sample (Green)



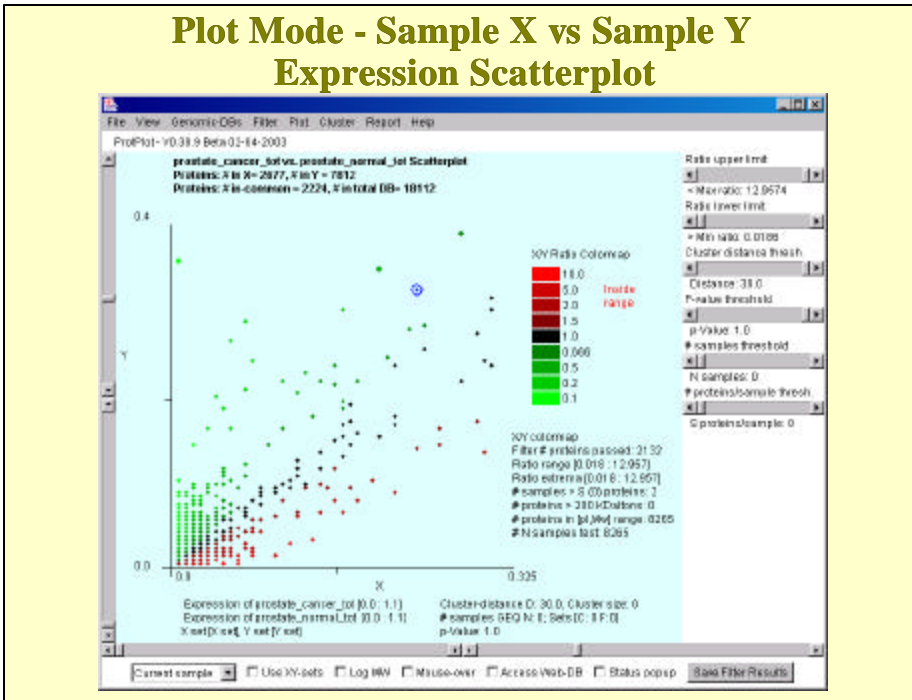
Plot Mode - Sample Xvs Sample Y Expression Scatterplot



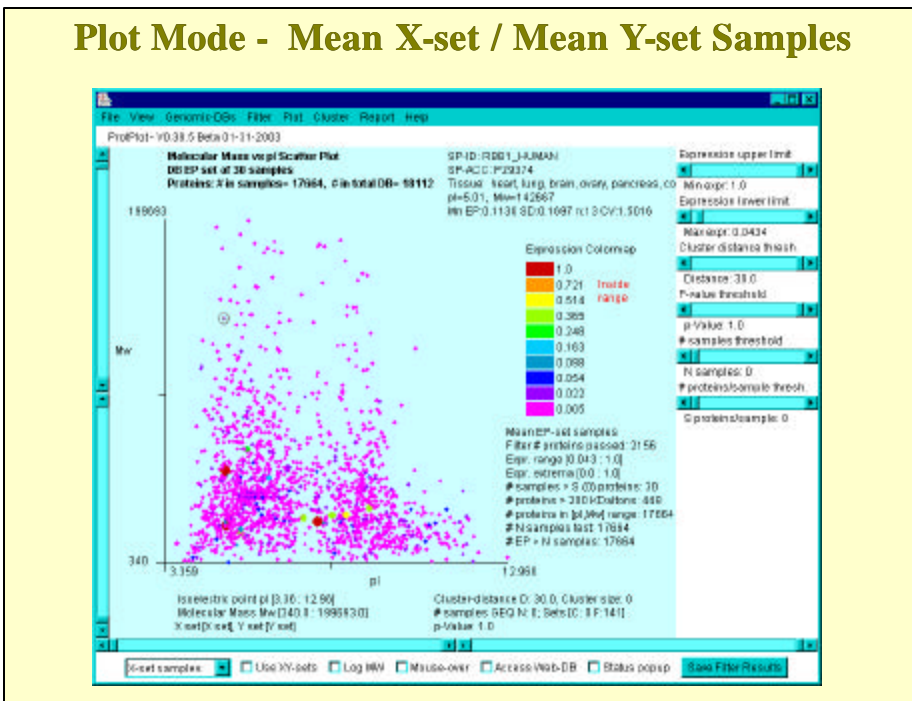
Plot Mode - X Sample / Y Sample Colormap



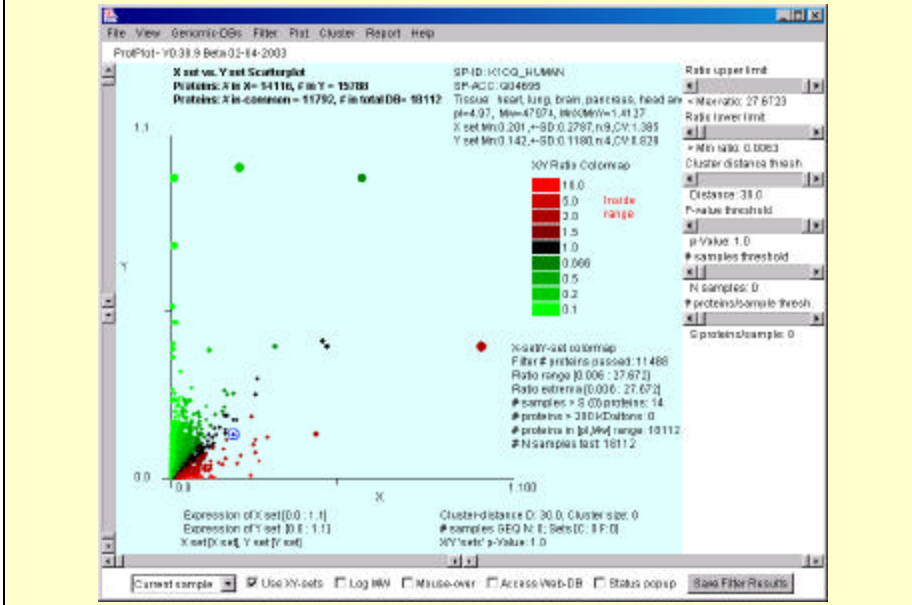
Plot Mode - Sample X vs Sample Y Expression Scatterplot



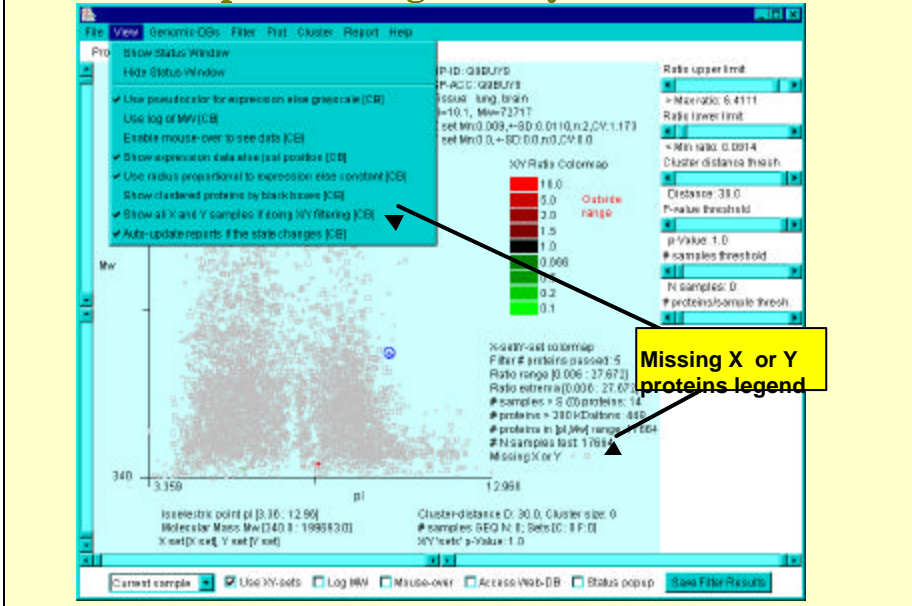
Plot Mode - Mean X-set / Mean Y-set Samples



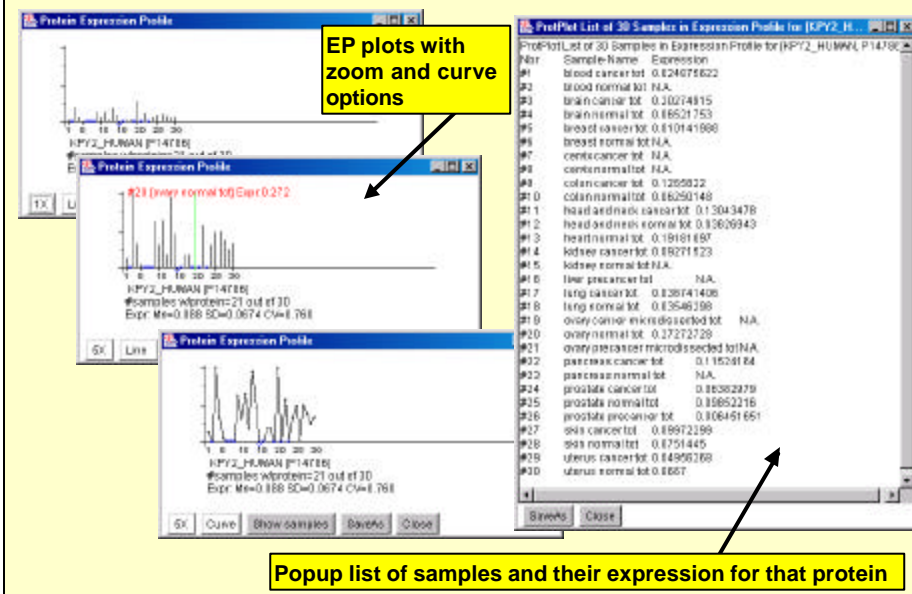
Plot Mode - Mean X-set vs Mean Y-set Expression Scatterplot



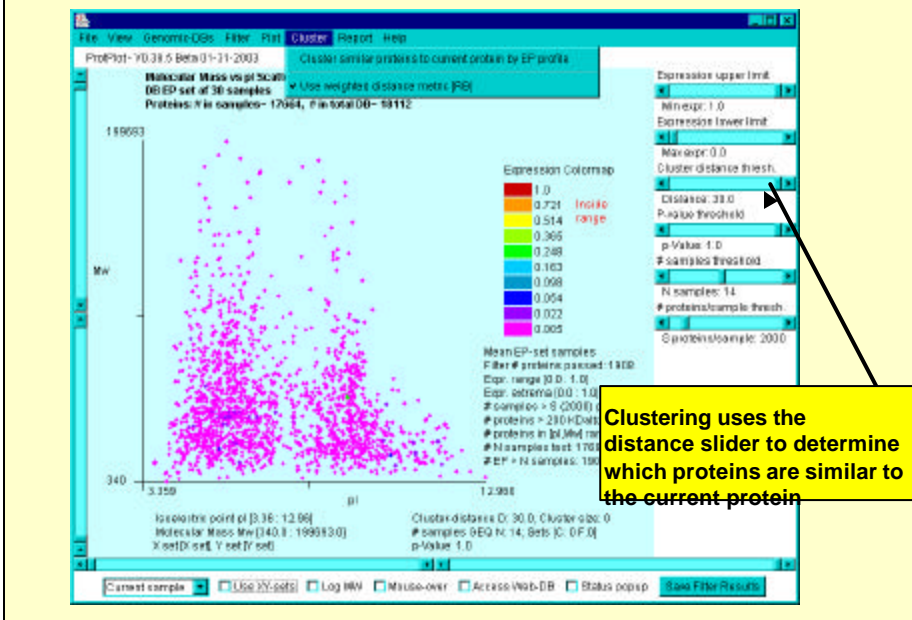
Plot Mode - Showing Proteins With Either X or Y Samples Missing as Gray '+' or Boxes



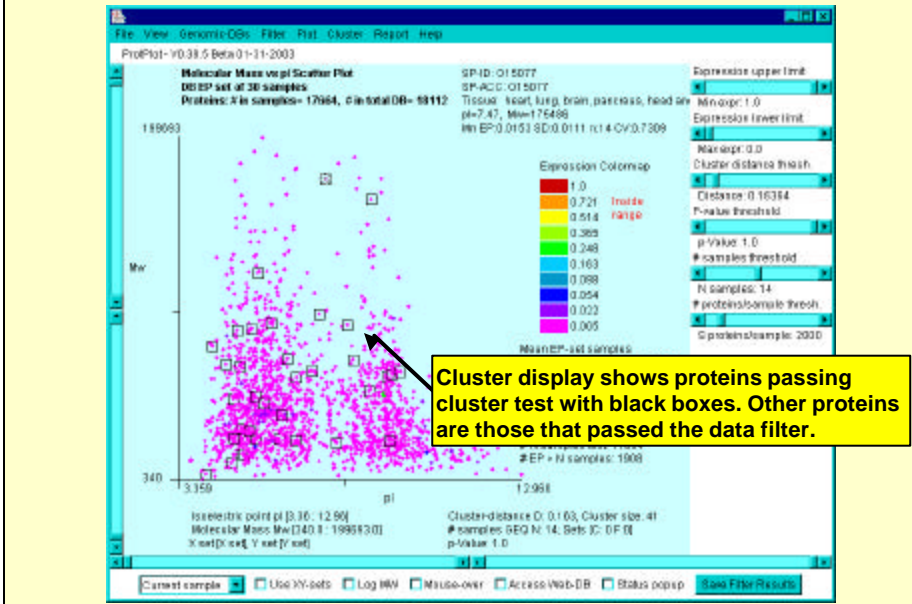
Plot Mode - Popup Expression Profile Plot for 1 Protein - Click on a Different Spot to Change the Plot



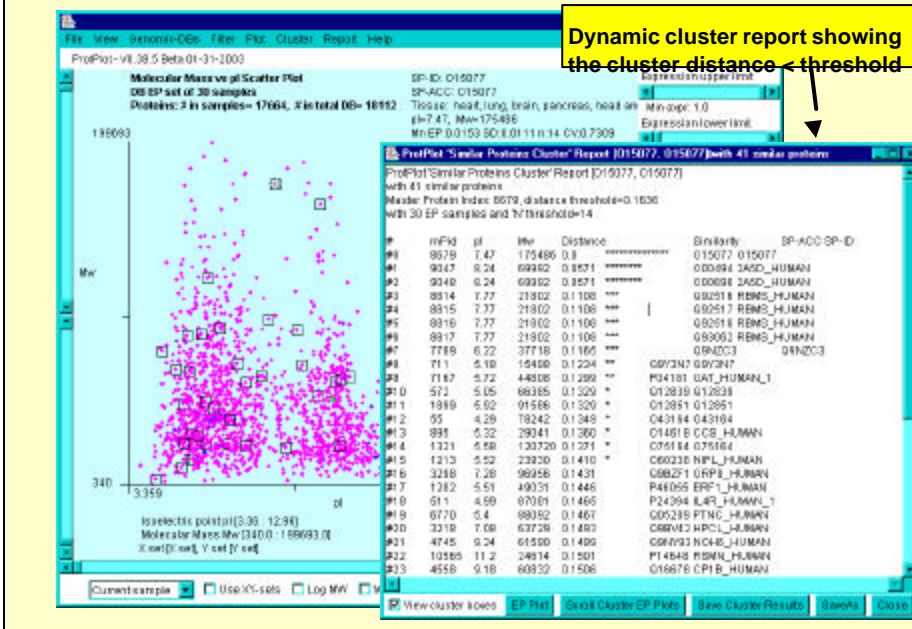
Cluster Menu - Find Proteins with Similar Expression



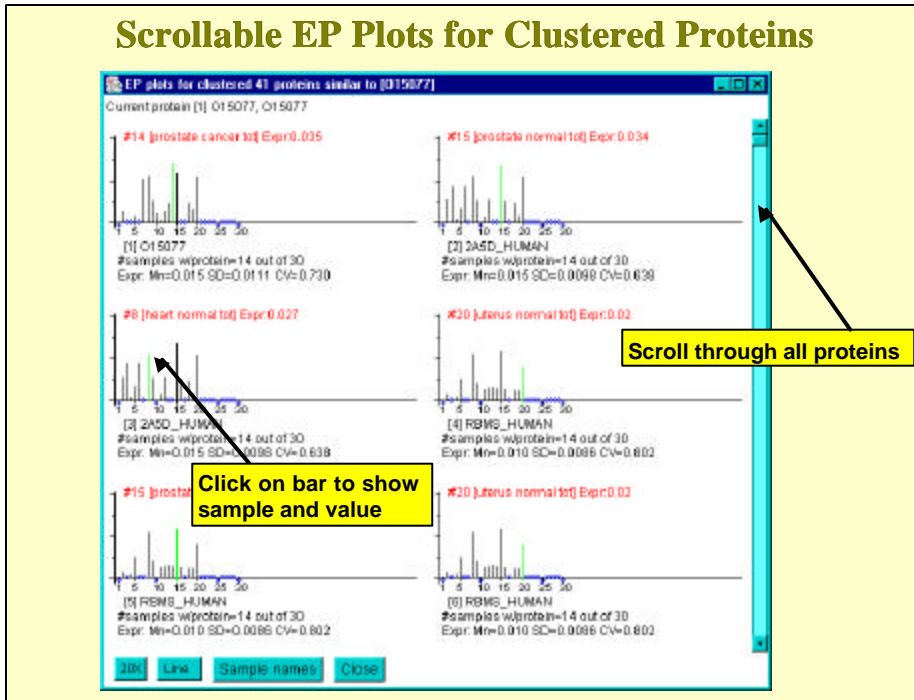
Clustering on Selected Protein - Scatterplot with Cluster Member Proteins Shown with Black Boxes



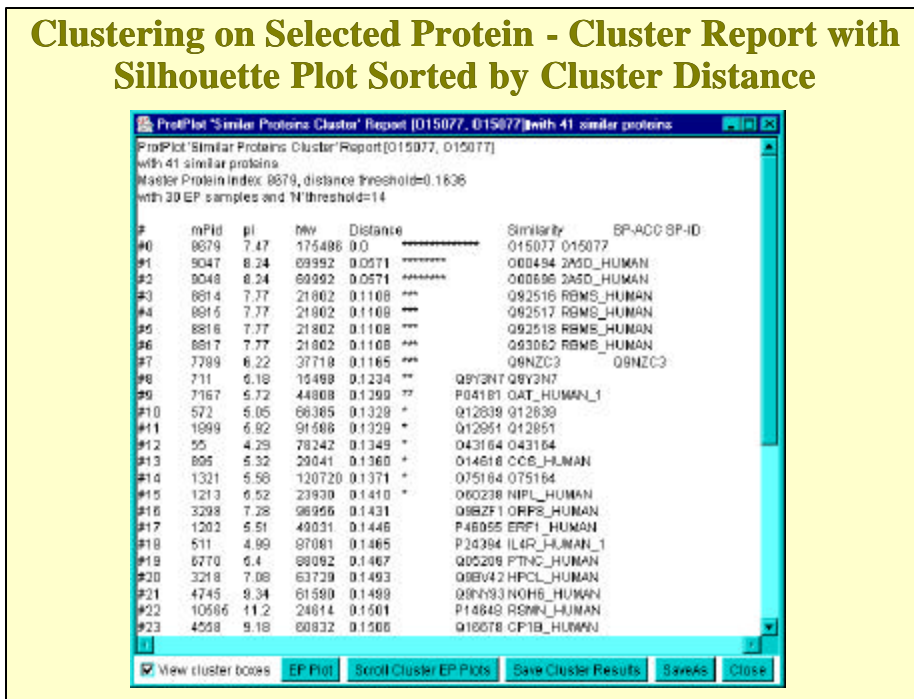
Clustering on Selected Protein (All Samples) $D < 0.69$



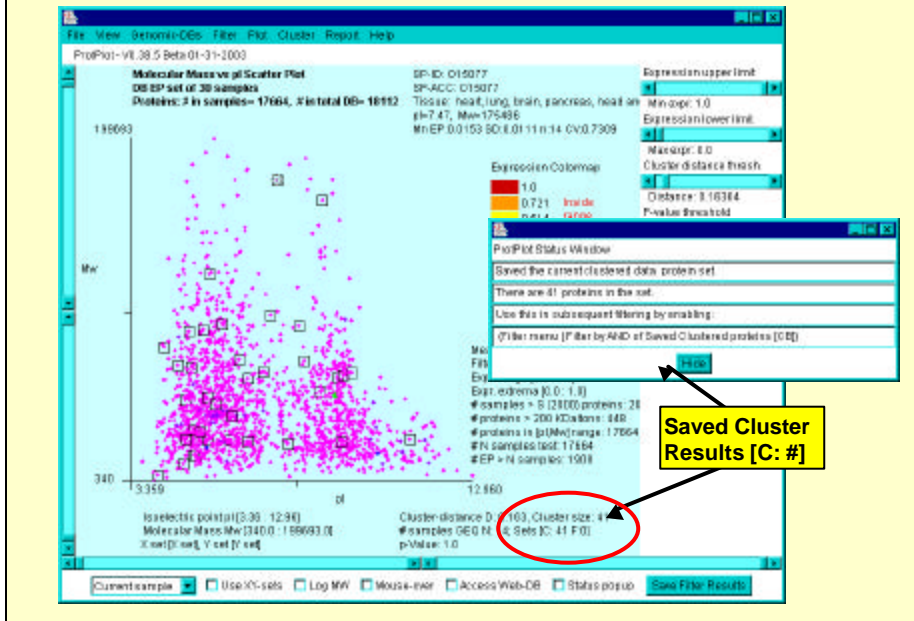
Scrollable EP Plots for Clustered Proteins



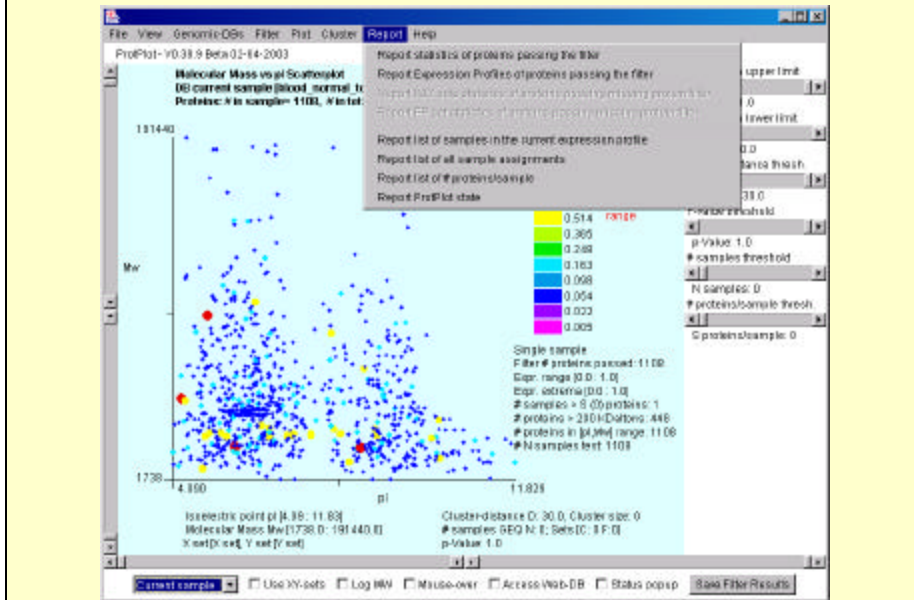
Clustering on Selected Protein - Cluster Report with Silhouette Plot Sorted by Cluster Distance



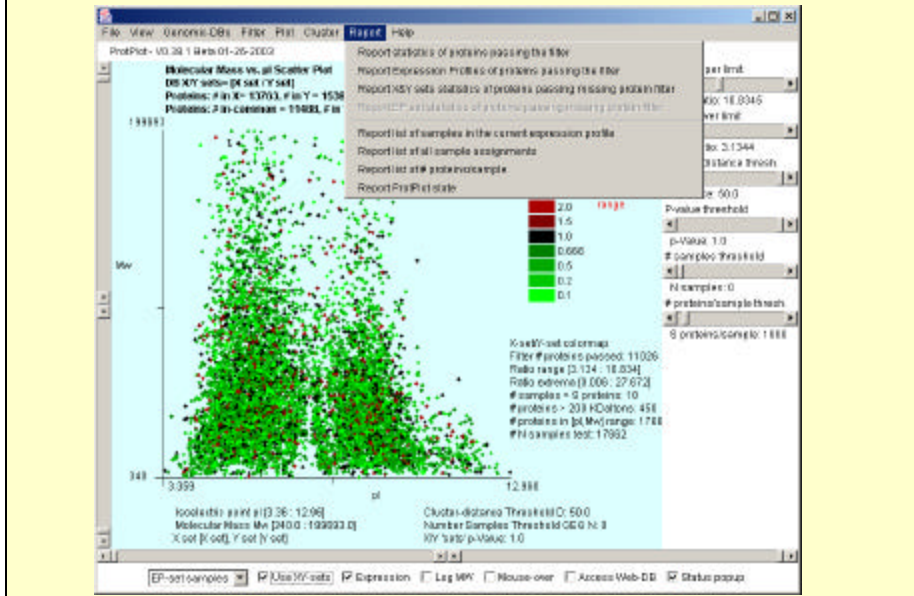
Saving Cluster Set of Proteins - For Future Filtering



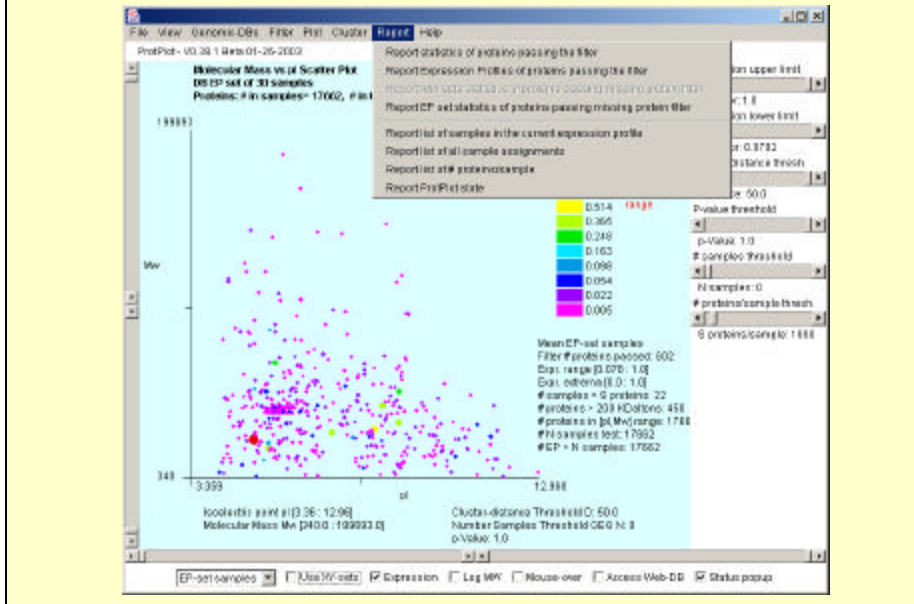
Report Menu - Options are Display Mode Dependent Current Sample Mode



Report Menu - Options are Display Mode Dependent Ratio Mode options



Report Menu - Options are Display Mode Dependent Mean EP Expression Mode



Popup Report for the Filter X/Y sets Minimum S>=3465 Proteins/Sample

ProPlot Filtered Protein Report [protplot] 3537 proteins

ProPlot Filtered Protein Report [protplot] 3537 proteins

(1) X-set samples: "brain cancer tot" "breast cancer tot" "colon cancer tot" "kidney cancer tot" "lung cancer tot" "pancreas ca tot"

(2) Y-set samples: "brain normal tot" "heart normal tot" "kidney normal tot" "lung normal tot" "prostate normal tot" "skin noncancer tot"

#Fid	pIa	FW	Ratio	Min-X	Max-X	Mean-X	Mean-Y	StdDev-X	StdDev-Y	CV-X	CV-Y	Mrx-X	Mrx-Y	SP-ACC	SP-ID	Tax
#470	5.27	39285	0.0351	0.0351	0.0	0.0450	0.0	1.2797	0	3	0.9993	SH3_HUMAN				heart, lung, brain, pro
#938	5.69	74145	0.0652	0.0652	0.0	0.0	0.0	0.8	0	1	0.9872	GATP_HUMAN				brain
#925	6.39	80245	0.0652	0.0652	0.0	0.0	0.0	0.8	0	1	0.9206	CDKL4_HUMAN				brain
#1221	3.36	27035	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9186	CDR1_HUMAN				brain
#1559	5.45	32515	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9823	QSOX2_HUMAN				lung, brain, kidney, uterus
#1572	5.48	62558	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9563	ETAS_HUMAN				brain
#1573	5.48	62558	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9564	ETAS_HUMAN				brain
#1574	5.48	62558	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9563	ETAS_HUMAN				brain
#1761	5.97	112077	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9794	QSOX2_HUMAN				brain
#1833	6.04	34531	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9542	ZNF4_HUMAN				brain
#1904	6.18	145660	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9793	QSOX2_HUMAN				brain
#2012	6.43	72648	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9807	QSOX2_HUMAN				brain
#3104	6.67	93888	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9794	QSOX2_HUMAN				brain
#2131	6.74	12497	0.0242	0.0242	0.0	0.0272	0.0	1.1263	0	2	0.9517	BMTB_HUMAN				lung, brain, prostate
#2172	6.92	29658	0.0373	0.0373	0.0	0.0086	0.0	0.2314	0	2	0.9366	CAH7_HUMAN				brain, colon
#2209	7.31	25662	0.0314	0.0314	0.0	0.0187	0.0	0.5961	0	3	0.9439	DTMS_HUMAN				heart, brain, prostate
#2250	7.73	27582	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9826	DTMS_HUMAN				lung, brain
#2263	8.20	30200	0.0434	0.0434	0.0	0.0	0.0	0.8	0	1	0.9793	DTMS_HUMAN				brain

Buttons: [Save] [Close]

Popup Report for the t-test (p=0.05) Filter X/Y sets Min 7 Samples for X and Y, S>=2000 proteins/sample

ProPlot Filtered Protein Report [protplot] 104 proteins

ProPlot Filtered Protein Report [protplot] 104 proteins

(1) X-set samples: "breast cancer tot" "breast cancer tot" "breast cancer tot" "breast cancer tot" "breast cancer tot" "breast cancer tot" "breast cancer tot"

(2) Y-set samples: "breast cancer tot" "breast cancer tot" "breast cancer tot" "breast cancer tot" "breast cancer tot" "breast cancer tot" "breast cancer tot"

#Fid	pIa	FW	p-value	Ratio	Min-X	Max-X	Mean-X	Mean-Y	StdDev-X	StdDev-Y	CV-X	CV-Y	Mrx-X	Mrx-Y	SP-ACC	SP-ID	Tax
#712	11.25	17779	1.7512	1.1401	1.1401	1.0268	1.0623	0.7160	0.4440	11	7	0.2678	F30663	PL24_HUMAN		heart, lung, brain, ovary, l	
#9428	5.68	24923	4.2390	1.9394	1.9394	1.0139	1.0460	0.6091	0.3408	11	0	0.2528	000289	CL11_HUMAN		heart, lung, brain, ovary, l	
#9946	6.21	18581	5.4805	1.1117	1.1117	1.0277	1.1127	0.7625	0.4029	12	7	0.1584	325528	CSF1_HUMAN		heart, lung, brain, ovary, l	
#2271	6.89	28747	7.8851	1.0314	1.0314	1.0791	1.0668	0.8122	0.4591	11	7	0.1737	796819	FA2L2_HUMAN		heart, lung, brain, ovary, l	
#29	6.83	24239	8.9321	1.1108	1.1108	1.0142	1.0029	0.7505	0.7000	0	0	0.1678	740253	HEAT_HUMAN		heart, lung, brain, ovary, l	
#517	5.89	11654	8.9332	1.0361	1.0361	1.0179	1.0494	0.5404	0.5742	0	7	0.2289	350814	SDS1_HUMAN		heart, lung, brain, ovary, l	
#2158	6.74	34885	8.9323	1.1313	1.1313	1.0220	1.0901	0.6154	0.7470	11	7	0.2562	700681	LEI1_HUMAN		heart, lung, brain, pancreas	
#2841	10.49	26884	8.9322	1.8620	1.8620	1.0117	1.0420	0.6128	0.8798	10	7	0.2522	080811	080811		heart, lung, brain, pancreas, testis	
#2793	6.23	21893	8.9341	1.0351	1.0351	1.0096	1.0029	0.6530	0.7086	0	7	0.1584	707283	POU2_HUMAN		heart, lung, brain, pancreas	
#204	10.21	16714	8.9347	1.1225	1.1225	1.0425	1.0651	1.0679	0.5317	11	7	0.3283	717088	SD14_HUMAN		heart, lung, brain, ovary, l	
#1561	5.5	36766	8.9351	1.9999	1.9999	1.0134	1.0714	0.6704	0.7153	9	7	0.1688	351571	SD9_HUMAN		heart, lung, brain, ovary, l	
#4042	6.87	62123	8.9360	1.1709	1.1709	1.0497	1.1260	0.7006	0.6722	12	0	0.2468	700228	SD8_HUMAN		heart, lung, brain, ovary, l	
#2024	6.23	4842	8.9361	1.8678	1.8678	1.0229	1.0406	1.1068	0.8988	10	7	0.2374	607193	SD7_HUMAN		heart, lung, brain, ovary, l	
#4089	10.16	18687	8.9371	1.1605	1.1605	1.0208	1.1374	0.7050	0.6506	11	7	0.2617	700226	SD7_HUMAN		heart, lung, brain, ovary, l	
#232	6.45	27352	8.9372	1.8509	1.8509	1.0152	1.0511	0.7652	0.8477	9	7	0.2481	700874	SD1_HUMAN		heart, lung, brain, ovary, l	
#2774	6.27	13214	8.9376	1.1716	1.1716	1.0355	1.1114	0.7107	0.6571	11	0	0.2374	300044	SD7_HUMAN		heart, lung, brain, ovary, l	

Buttons: [Save] [Close]

PopUp Report Expression Profile Values for Filtered Proteins (min N>= 14 samples, S >=2000)

ProPlot Expression Profile List of 1308 Filtered Proteins (with 30 samples passing 3 Filter)

SP-ACC	SP-ID	"brain normal tot"	"brain cancer tot"	"breast cancer tot"	"colon cancer tot"	"head and neck cano"
P04720	EF11_HUMAN	1.0	0.8900	0.0750	1.0	0.4347
P07680	D76800	0.2826	0.0412	0.0060	0.0506	0.0
P13473	LMP2_HUMAN_1	0.2391	0.0584	0.0040	0.0295	0.0
P02794	FRIB_HUMAN	0.2391	0.2611	0.0425	0.4219	0.3913
P02511	CRAB_HUMAN	0.2391	0.0340	0.0040	0.0084	0.0
P07900	H9A_HUMAN	0.2173	0.0687	0.0223	0.1223	0.1086
P02571	ACTE_HUMAN_1	0.2173	0.5532	0.0344	0.3037	0.2608
P27816	MAP4_HUMAN	0.2173	0.0721	0.0040	0.0633	0.0
Q13082	MAP4_HUMAN	0.2173	0.0721	0.0040	0.0633	0.0
Q96A76	MAP4_HUMAN	0.2173	0.0721	0.0040	0.0633	0.0
P04083	ARPC_HUMAN	0.2173	0.0206	0.0020	0.0084	0.0434
P40925	MEW1_HUMAN	0.2173	0.0274	0.0141	0.0064	0.0
P15173	PCNA_HUMAN	0.1956	0.0274	0.0040	0.0169	0.0217
P10041	ADP2_HUMAN	0.1956	0.0378	0.0040	0.039	0.0217
P35209	NFL1_HUMAN	0.1739	0.1168	0.0	0.0506	0.0434
Q9UB95	GIL2_HUMAN	0.1739	0.0481	0.0020	0.0169	0.0
Q9UG13	GIL2_HUMAN	0.1739	0.0481	0.0020	0.0169	0.0
P29360	143E_HUMAN	0.1521	0.0859	0.0101	0.0886	0.0
P42655	143E_HUMAN	0.1521	0.0859	0.0101	0.0886	0.0
Q96FZ2	Q96FZ2	0.1521	0.0103	0.0020	0.0464	0.0
Q96FZ6	Q96FZ6	0.1521	0.0103	0.0020	0.0464	0.0

PopUp Report of Samples in Expression Profile Set for the Currently Selected Protein

ProPlot List of 30 Samples in Expression Profile for [IF39_HUMAN, P55884] (N=45)

Idx	Sample-Name	Expression
#1	brain normal tot	N.A.
#2	brain cancer tot	0.04467354
#3	breast cancer tot	N.A.
#4	breast normal tot	N.A.
#5	colon cancer tot	0.0253
#6	colon normal tot	0.03125074
#7	head and neck cancer tot	0.02178913
#8	head and neck normal tot	0.005181347
#9	heart normal tot	0.054803133
#10	kidney cancer tot	0.009933775
#11	kidney normal tot	N.A.
#12	liver precancer tot	N.A.
#13	lung cancer tot	0.006389809
#14	lung normal tot	0.007092590
#15	ovary cancer microdissected tot	N.A.
#16	ovary normal tot	N.A.
#17	ovary precancer microdissected tot	N.A.
#18	pancreas cancer tot	0.011152416
#19	pancreas normal tot	N.A.
#20	prostate cancer tot	0.021276595
#21	prostate normal tot	0.014778324
#22	skin cancer tot	0.011080332
#23	skin normal tot	0.02890173
#24	uterus cancer tot	0.011661807
#25	uterus normal tot	0.0333
#26	blood cancer tot	0.004975124
#27	blood normal tot	0.999865
#28	prostate precancer tot	0.010354052
#29	cervix normal tot	N.A.
#30	cervix cancer tot	N.A.

Popup Report # of proteins/sample for All Samples

Protein Sample Numbers of Proteins

Sample-name	Nbr-proteins/sample
brain_normal_tot	4616
brain_cancer_tot	10167
breast_cancer_tot	5844
breast_normal_tot	314
colon_cancer_tot	9762
colon_normal_tot	1981
head_and_neck_cancer_tot	2683
head_and_neck_normal_tot	2230
heart_normal_tot	7439
kidney_cancer_tot	9206
kidney_normal_tot	8978
liver_precancer_tot	240
lung_cancer_tot	8521
lung_normal_tot	6068
ovary_cancer_microdissected_tot	891
ovary_normal_tot	435
ovary_precancer_microdissected_tot	134
pancreas_cancer_tot	6934
pancreas_normal_tot	76
prostate_cancer_tot	2849
prostate_normal_tot	8202
skin_cancer_tot	4128
skin_normal_tot	3975
uterus_cancer_tot	9581
uterus_normal_tot	5928
blood_cancer_tot	1629
blood_normal_tot	1151
prostate_precancer_tot	2493
cervix_normal_tot	879
cervix_cancer_tot	887

Buttons: SaveAs, Close

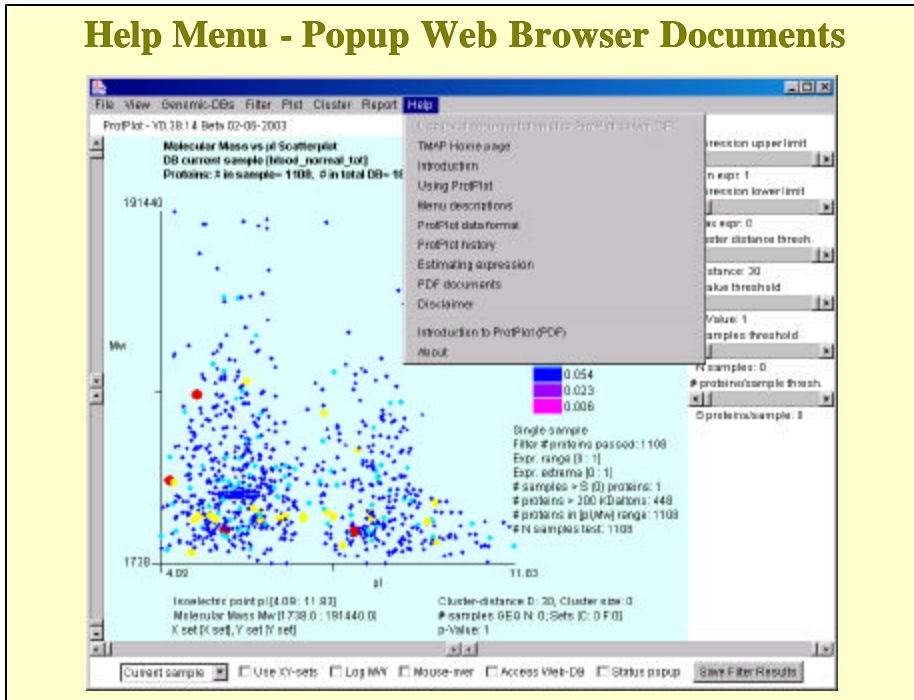
Popup Report All X, Y, EP Sets Sample Assignments

Protein Sample Assignment Lists

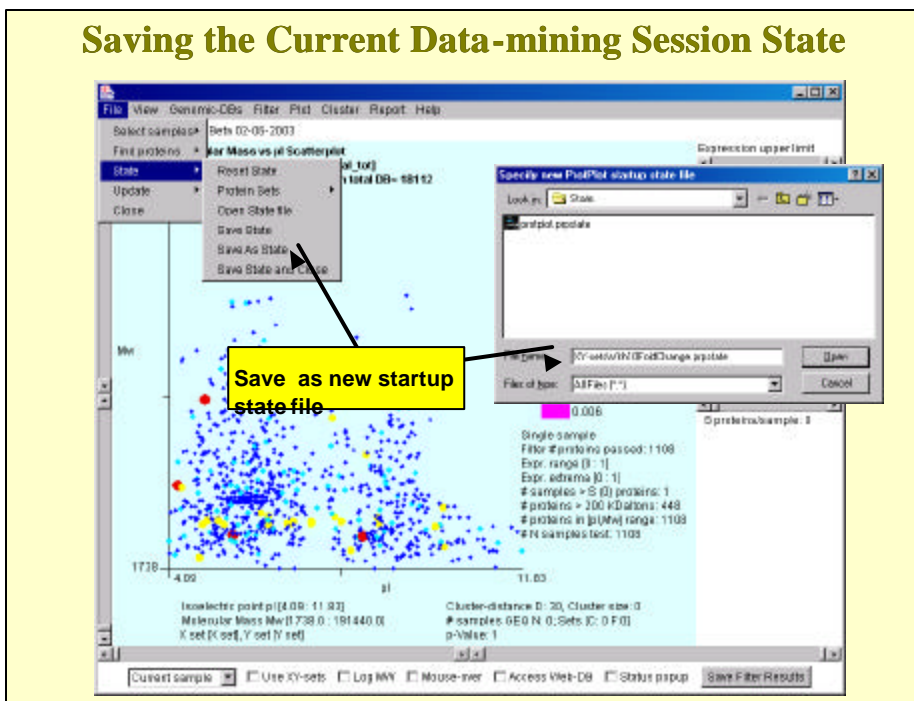
Condition-list	Sample-name	Nbr-proteins
Current-sample	blood_normal_tot	1151
X-sample	prostate_cancer_tot	2849
Y-sample	prostate_normal_tot	8202
X set[X set]	brain_cancer_tot	10167
X set[X set]	breast_cancer_tot	5844
X set[X set]	colon_cancer_tot	9762
X set[X set]	head_and_neck_cancer_tot	2683
X set[X set]	kidney_cancer_tot	9206
X set[X set]	liver_precancer_tot	240
X set[X set]	lung_cancer_tot	8521
X set[X set]	ovary_cancer_microdissected_tot	891
X set[X set]	ovary_precancer_microdissected_tot	134
X set[X set]	pancreas_cancer_tot	6934
X set[X set]	prostate_cancer_tot	2849
X set[X set]	skin_cancer_tot	4128
X set[X set]	uterus_cancer_tot	9581
X set[X set]	blood_cancer_tot	1629
X set[X set]	prostate_precancer_tot	2493
X set[X set]	cervix_cancer_tot	887
Y set[Y set]	brain_normal_tot	4616
Y set[Y set]	breast_normal_tot	314
Y set[Y set]	colon_normal_tot	1981
Y set[Y set]	head_and_neck_normal_tot	2230
Y set[Y set]	heart_normal_tot	7439
Y set[Y set]	kidney_normal_tot	8978
Y set[Y set]	lung_normal_tot	6068
Y set[Y set]	ovary_normal_tot	435
Y set[Y set]	pancreas_normal_tot	76
Y set[Y set]	prostate_normal_tot	8202
Y set[Y set]	skin_normal_tot	3975
Y set[Y set]	uterus_normal_tot	5928
Y set[Y set]	blood_normal_tot	1151
Y set[Y set]	cervix_normal_tot	879
Eppe-ProFile list	brain_normal_tot	4616
Eppe-ProFile list	brain_cancer_tot	10167
Eppe-ProFile list	breast_cancer_tot	5844

Buttons: SaveAs, Close

Help Menu - Popup Web Browser Documents



Saving the Current Data-mining Session State



Changing the State to a Previous Data-mining Session

The screenshot shows the software interface with the 'State' menu open. The 'Open State file' option is highlighted. A dialog box titled 'Specify new PlotPlot startup state file' is open, showing a file path and a 'File of type' dropdown set to 'State File (*.*)'. A yellow callout box with the text 'Opening the data-mining state to a previous session' has arrows pointing to the 'Open State file' menu item and the 'Open' button in the dialog box.

Changing the Filter Set of Proteins to a Previously Saved Filter Set

The screenshot shows the software interface with the 'Filter Sets' menu open. The 'Open Filter protein set from .set file' option is highlighted. A dialog box titled 'Specify new PlotPlot Filter protein set file' is open, showing a file path and a 'File of type' dropdown set to 'State File (*.*)'. A yellow callout box with the text 'Changing the Filter set of proteins to previously saved set' has arrows pointing to the 'Open Filter protein set from .set file' menu item and the 'Open' button in the dialog box.

References

- Medjahed D, Luke BT, Tontesh TS, Smythers GW, Munroe DJ, Lemkin PF, TMAP poster, Swiss Proteomics Meeting, Geneva, Dec, 2002.
- Medjahed D, Smythers GW, Powell DA, Stephens RM, Lemkin PF, Munroe DJ, VIRTUAL2D: A Web-accessible predictive database for proteomics analysis, *Proteomics*, 2003, (in press, Feb).
- Medjahed D, Luke BT, Tontesh TS, Smythers GW, Munroe DJ, Lemkin PF, "TMAP" (Tissue Molecular Anatomy Project), an expression database for comparative cancer proteomics. *Proteomics*, 2003, (in press, June).

References

