MicroArray Explorer - a Tool for Data Mining of Microarrays:

Examples

http://www.lecb.ncifcrf.gov/MAExplorer

PF Lemkin[1], GC Thornwall[2], K Walton[3], L Hennighausen[3]
(1) LECB, NCI/FCRDC (2) SAIC/FCRDC (3) LGP, NIDDK
mail: lemkin@ncifcrf.gov

Revised: 10-27-2001

MAExplorer Home Page
http://www.lecb.ncifcrf.gov/MAExplorer

MAExplorer

- Introduction
- Demonstrations

Documentation
- Manual (on right)
- Manual (new window)
- Manual (7Mb Zip)
- Manual (entire)
- Newsletters

- Short Tutorial
- Advanced tutorial
- Menu summary
- Quick Start
- Glossary
- Index
- Help Desk

- Overview (PDF)
- Examples (PDF)
- Intro Data Mining (PDF)
- N.A.R. paper (PDF)
- Use with mAdb (PDF)

Downloading MAExplorer
- User's array data
- Cvt2Mae data converter
- Stand-alone version
- NEW Revision notes
- Installer information

- download

Disclaimer

MGAP-MAExplorer home

Contact [Lemkin | lemkin@ncifcrf.gov | (LECB, NCI/FCRDC)]

Revised: $Date: 2001/08/10 18:05:25 $

MAExplorer - MicroArray Explorer

MicroArray Explorer for Data Mining Gene Expression Patterns

The Microarray Explorer (MAExplorer) is a Java-based data-mining facility for cDNA microarray databases. It may be freely downloaded and run as a stand-alone application on your computer, or run as an applet in your Web browser. The exploratory data analysis environment provides tools for the data-mining of quantitative cDNA expression profiles across multiple microarrays.

With this program it is possible to: 1) analyze the expression of individual genes; 2) analyze the expression of gene families and clusters; 3) compare expression patterns and outliers; 4) directly access other genomic databases for clones of interest. In the applet version, data is downloaded as required from the server to the user's Web browser where real-time analyses are performed. The stand-alone version uses previously quantified array data copied to the local computer where it may save data from data mining sessions.

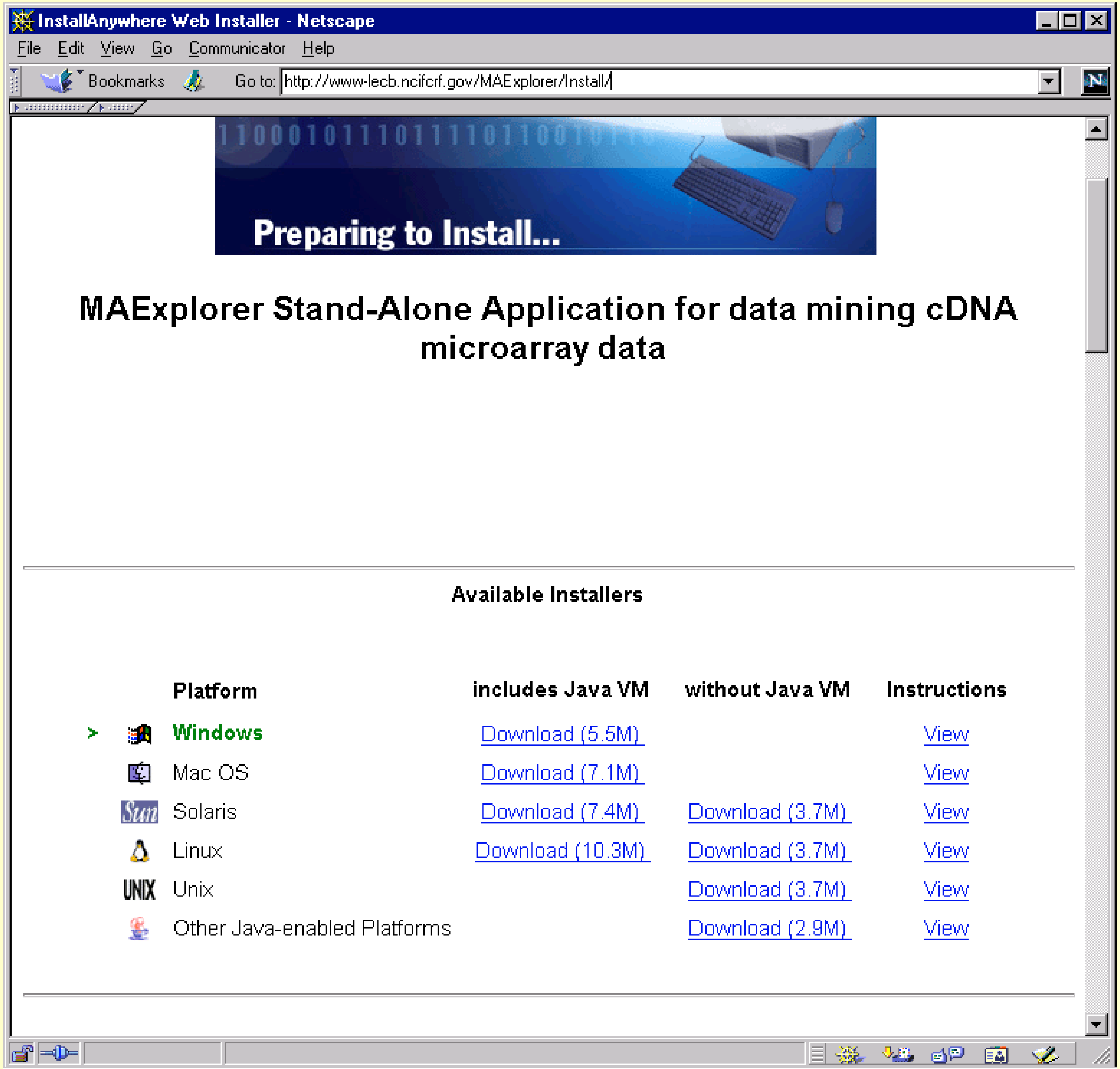
Microarray data may be viewed and directly manipulated in array pseudoimages, scatter plots, histograms, expression profile plots, cluster analyses (similar clones, K-means, hierarchical clusters, etc.), and reports. A key feature is the clone data filters for constraining a working set of clones to those passing a variety of user-specifed tests. Reports may be generated with hypertext Web access to genomic databases such as UniGene, GenBank, dbEST, I.M.A.G.E., NCI/CIT mAdb Clone DB and other Internet databases for sets of clones found to be of interest.

A major focus of this tool is interactive data mining with access to other supporting Web genomic databases. The emphasis on direct manipulation of clones and sets of clones in graphics and tables provides a high level of interaction with the data making it easier for investigators to test ideas when looking for patterns.

It was developed by the NCI Laboratory of Experimental and Computational Biology (LECB) in collaboration with the NIDDK Laboratory of Genetics and Physiology (LGP). MAExplorer was created to help analyze microarray data for the LGP's Mammary Genome Anatomy Program (MGAP) designed to identify and understand genetic pathways operative during normal mammary gland development and tumorigenesis. Note that 38 hybridizations from the MGAP database are included as a demonstration database when you download the stand-alone version of MAExplorer.

MAExplorer may be used as an applet to access the Mammary Genome Anatomy Project (MGAP) microarray Web data through http://mammary.nih.gov/mgap or directly at http://www.lecb.ncifcrf.gov/mae.

Download Stand-alone Web page



MAExplorer Stand-Alone Application for data mining cDNA microarray data

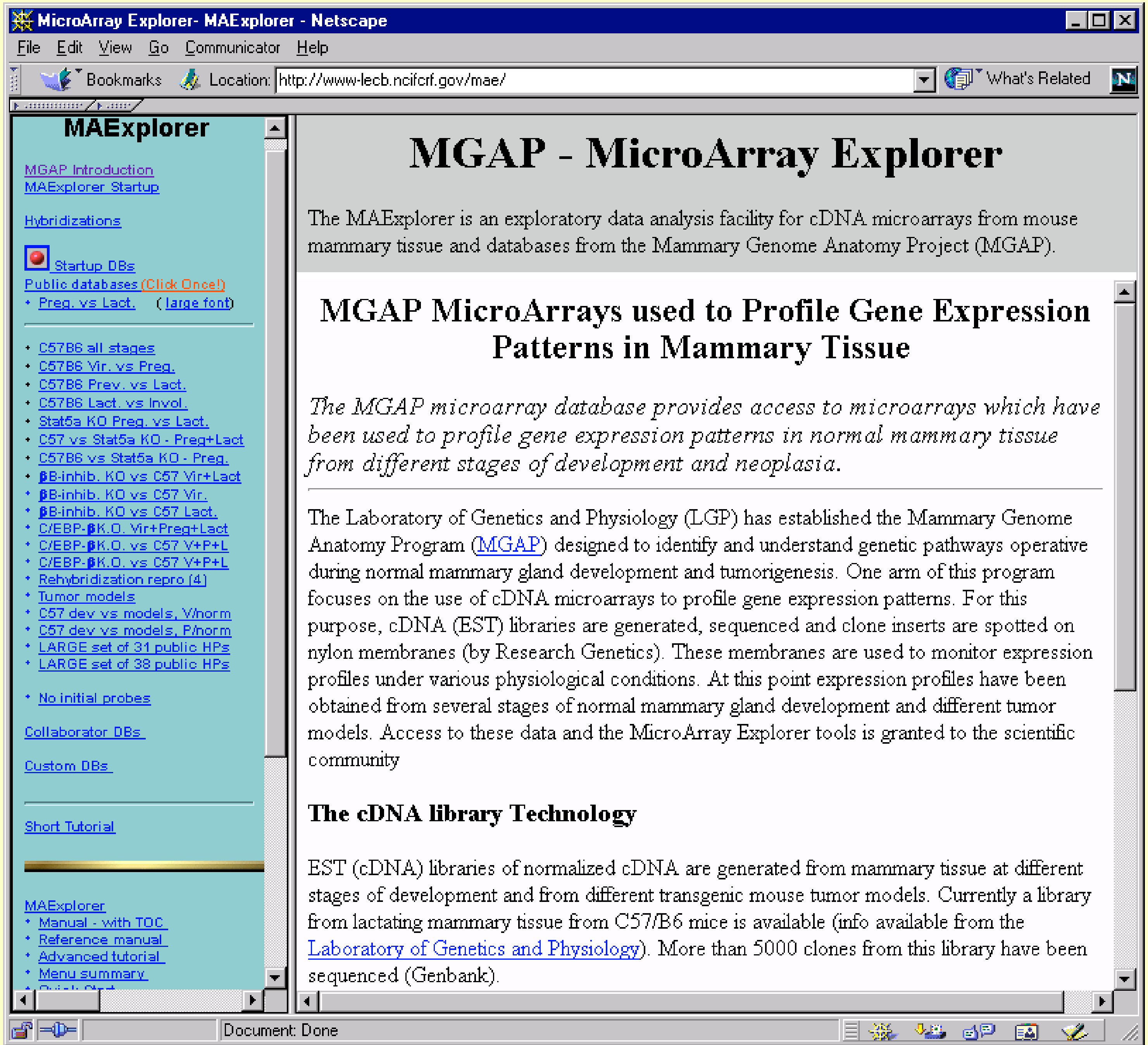
---

Available Installers

	Platform	includes Java VM	without Java VM	Instructions
>	**Windows**	Download (5.5M)		View
	Mac OS	Download (7.1M)		View
	Solaris	Download (7.4M)	Download (3.7M)	View
	Linux	Download (10.3M)	Download (3.7M)	View
	Unix		Download (3.7M)	View
	Other Java-enabled Platforms		Download (2.9M)	View

MGAP page for MAExplorer database

- Mammary Genome Anatomy Project microarray database using MAExplorer



MAExplorer

MGAP Introduction
MAExplorer Startup

Hybridizations

Startup DBs
Public databases (Click Once!)
- Preg. vs Lact. (large font)

- C57B6 all stages
- C57B6 Vir. vs Preg.
- C57B6 Prev. vs Lact.
- C57B6 Lact. vs Invol.
- Stat5a KO Preg. vs Lact.
- C57 vs Stat5a KO - Preg+Lact
- C57B6 vs Stat5a KO - Preg.
- βB-inhib. KO vs C57 Vir+Lact
- βB-inhib. KO vs C57 Vir.
- βB-inhib. KO vs C57 Lact.
- C/EBP-βK.O. Vir+Preg+Lact
- C/EBP-βK.O. vs C57 V+P+L
- C/EBP-βK.O. vs C57 V+P+L
- Rehybridization repro (4)
- Tumor models
- C57 dev vs models, V/norm
- C57 dev vs models, P/norm
- LARGE set of 31 public HPs
- LARGE set of 38 public HPs

- No initial probes

Collaborator DBs

Custom DBs

Short Tutorial

MAExplorer
- Manual - with TOC
- Reference manual
- Advanced tutorial
- Menu summary

MGAP - MicroArray Explorer

The MAExplorer is an exploratory data analysis facility for cDNA microarrays from mouse mammary tissue and databases from the Mammary Genome Anatomy Project (MGAP).

MGAP MicroArrays used to Profile Gene Expression Patterns in Mammary Tissue

The MGAP microarray database provides access to microarrays which have been used to profile gene expression patterns in normal mammary tissue from different stages of development and neoplasia.

The Laboratory of Genetics and Physiology (LGP) has established the Mammary Genome Anatomy Program (MGAP) designed to identify and understand genetic pathways operative during normal mammary gland development and tumorigenesis. One arm of this program focuses on the use of cDNA microarrays to profile gene expression patterns. For this purpose, cDNA (EST) libraries are generated, sequenced and clone inserts are spotted on nylon membranes (by Research Genetics). These membranes are used to monitor expression profiles under various physiological conditions. At this point expression profiles have been obtained from several stages of normal mammary gland development and different tumor models. Access to these data and the MicroArray Explorer tools is granted to the scientific community

The cDNA library Technology

EST (cDNA) libraries of normalized cDNA are generated from mammary tissue at different stages of development and from different transgenic mouse tumor models. Currently a library from lactating mammary tissue from C57/B6 mice is available (info available from the Laboratory of Genetics and Physiology). More than 5000 clones from this library have been sequenced (Genbank).

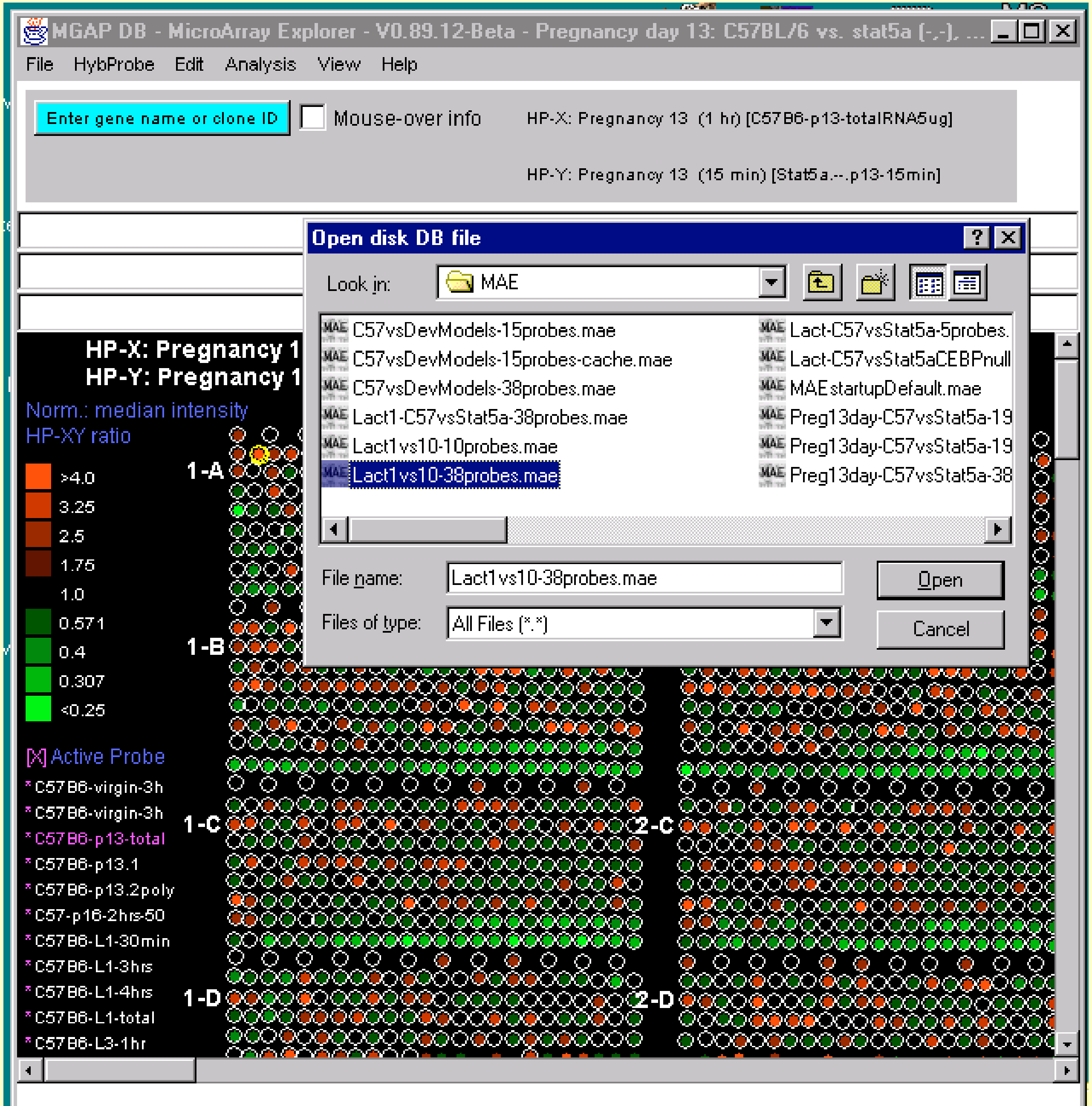
Examples of MAExplorer

- The following examples demonstrate some of its capabilities
- Note: many more examples and discussion of the various analysis plots and reports may be found in the on-line reference manual at

  http://www.lecb.ncifcrf.gov/MAExplorer/hmaeHelp.html

1a. Opening a database from local disk

- In stand-alone mode, you can browse a project database containing a set of startup databases.



MGAP DB - MicroArray Explorer - V0.89.12-Beta - Pregnancy day 13: C57BL/6 vs. stat5a (-,-), ...

HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]

HP-X: Pregnancy 1
HP-Y: Pregnancy 1

Norm.: median intensity
HP-XY ratio

- >4.0
- 3.25
- 2.5
- 1.75
- 1.0
- 0.571
- 0.4
- 0.307
- <0.25

[X] Active Probe

- C57B6-virgin-3h
- C57B6-virgin-3h
- C57B6-p13-total
- C57B6-p13.1
- C57B6-p13.2poly
- C57-p16-2hrs-50
- C57B6-L1-30min
- C57B6-L1-3hrs
- C57B6-L1-4hrs
- C57B6-L1-total
- C57B6-L3-1hr

Open disk DB file

Look in: MAE

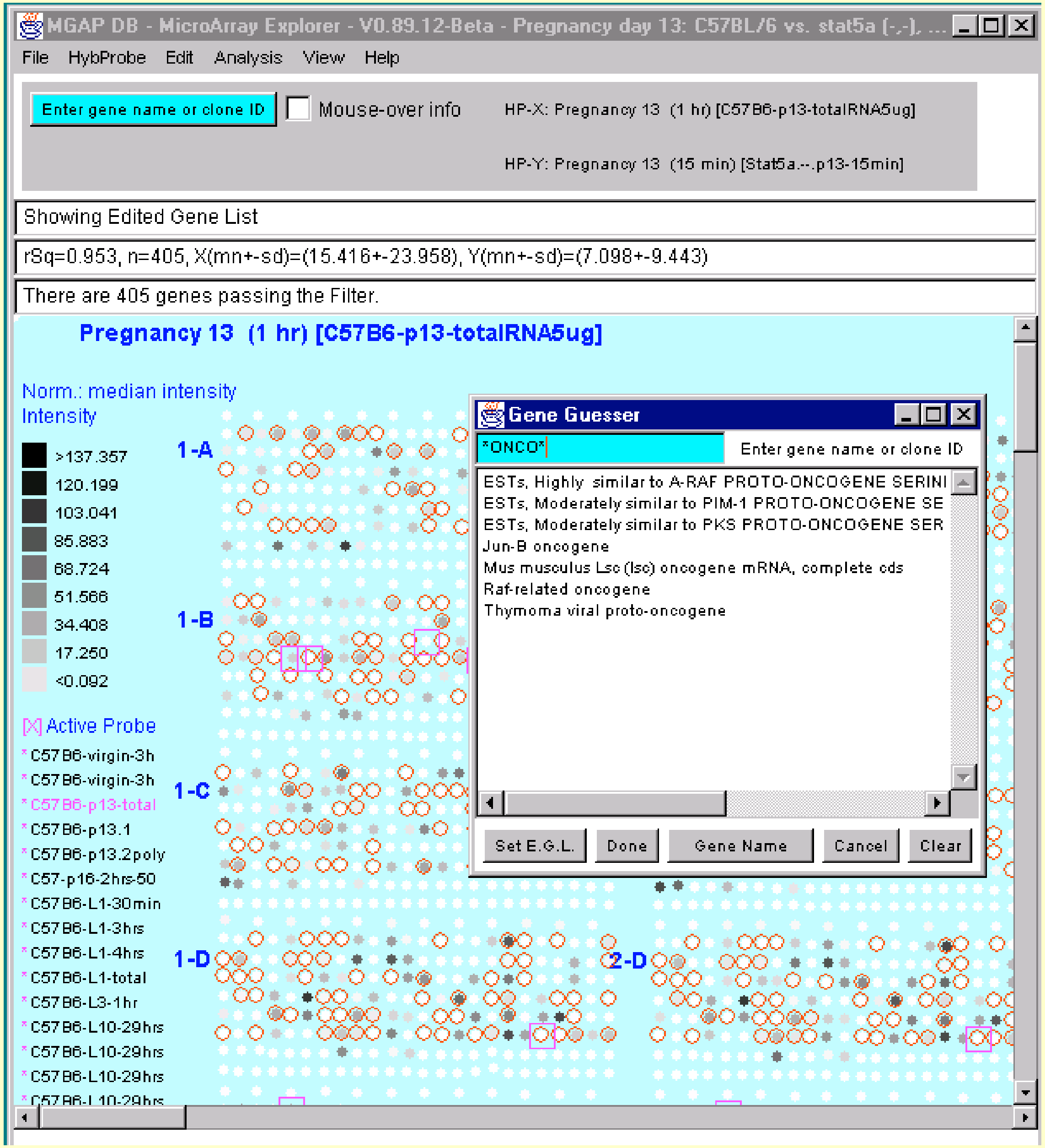
- C57vsDevModels-15probes.mae
- C57vsDevModels-15probes-cache.mae
- C57vsDevModels-38probes.mae
- Lact1-C57vsStat5a-38probes.mae
- Lact1vs10-10probes.mae
- Lact1vs10-38probes.mae
- Lact-C57vsStat5a-5probes.
- Lact-C57vsStat5aCEBPnull
- MAEstartupDefault.mae
- Preg13day-C57vsStat5a-19
- Preg13day-C57vsStat5a-19
- Preg13day-C57vsStat5a-38

File name: Lact1vs10-38probes.mae

Files of type: All Files (*.*)

1b. Specify Gene or Gene Subset by Name

- Specify gene or gene subset by gene name guesser using wildcard sub-strings eg. “*ONCO*” indicated by magenta boxes - saved in ‘Edited gene List’. [MGAP DB]



HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]

Showing Edited Gene List

rSq=0.953, n=405, X(mn+-sd)=(15.416+-23.958), Y(mn+-sd)=(7.098+-9.443)

There are 405 genes passing the Filter.

Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

ONCO

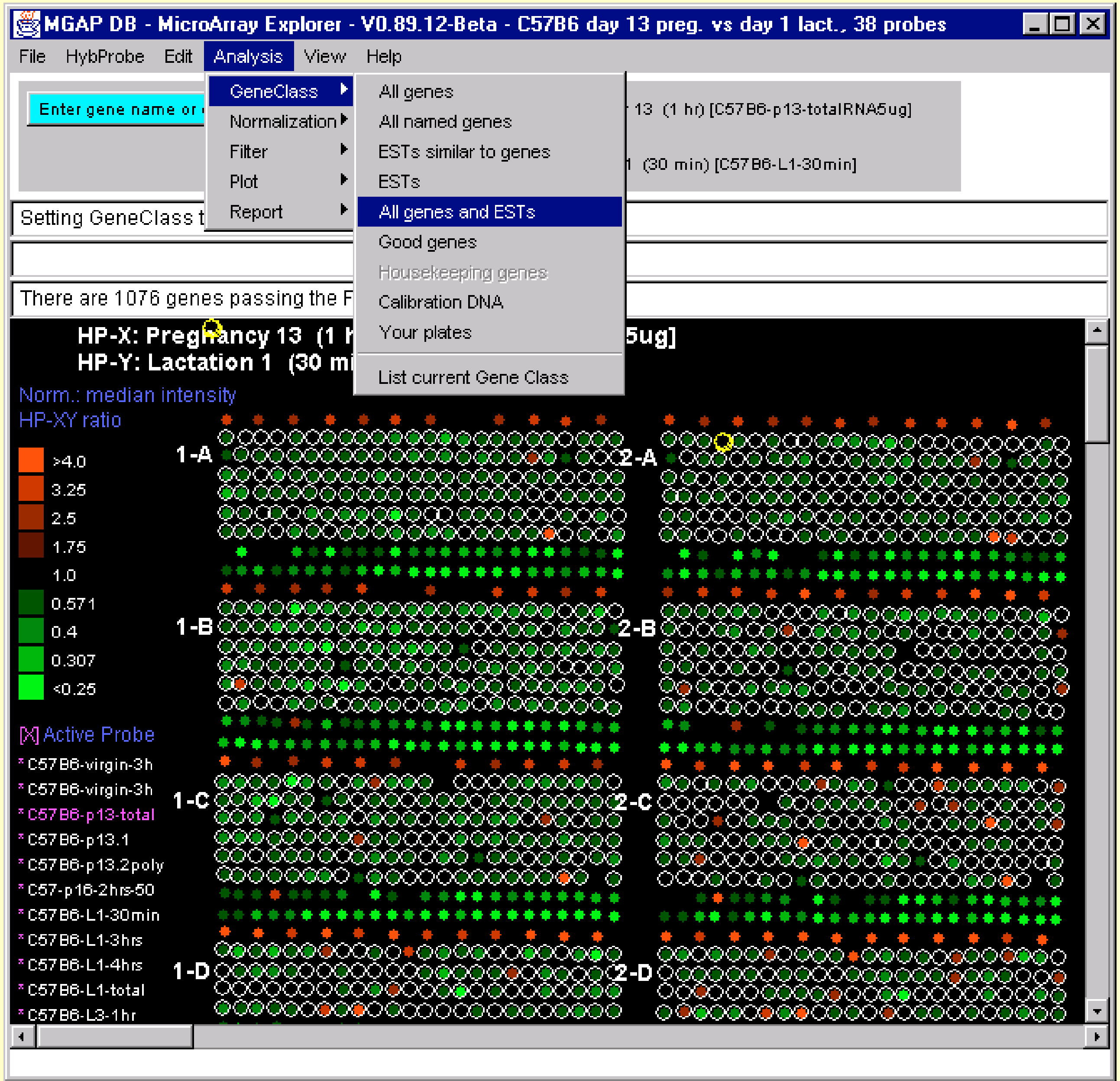
ESTs, Highly similar to A-RAF PROTO-ONCOGENE SERINI
ESTs, Moderately similar to PIM-1 PROTO-ONCOGENE SE
ESTs, Moderately similar to PKS PROTO-ONCOGENE SER
Jun-B oncogene
Mus musculus Lsc (lsc) oncogene mRNA, complete cds
Raf-related oncogene
Thymoma viral proto-oncogene

MAExplorer User Interface

- The MAExplorer menus are similar to most Windows PC applications where pull-down menu selections are used to invoke operations. These menus are fully documented in the on-line reference manual. The current hybridization probe array is displayed as a pseudo image of spot intensity. Clicking on a spot assigns it as the current gene with data being reported in the top most message area. The names of the current HP-X and HP-Y probes are listed above that area. Clicking on the “Enter gene name or Clone ID” button pops up a dialog box to assign the current gene (or set of genes) by name or wildcard. In general, clicking on spots, points in plots or cells in spreadsheet reports will assign the current gene and access Web genomic databases. The MGAP microarrays contain 1,700 duplicated ^{33}P-labeled clones indicated as fields 1 and 2 in the array pseudo image. Duplicated grids of cDNA spots are labeled as 1-A, 2-A, 1-B, 2-B, etc.

2a. Named Genes and ESTs

- Specify sets of genes for *all named genes* and *all ESTs* indicated in the microarray by white circles. [MGAP data]



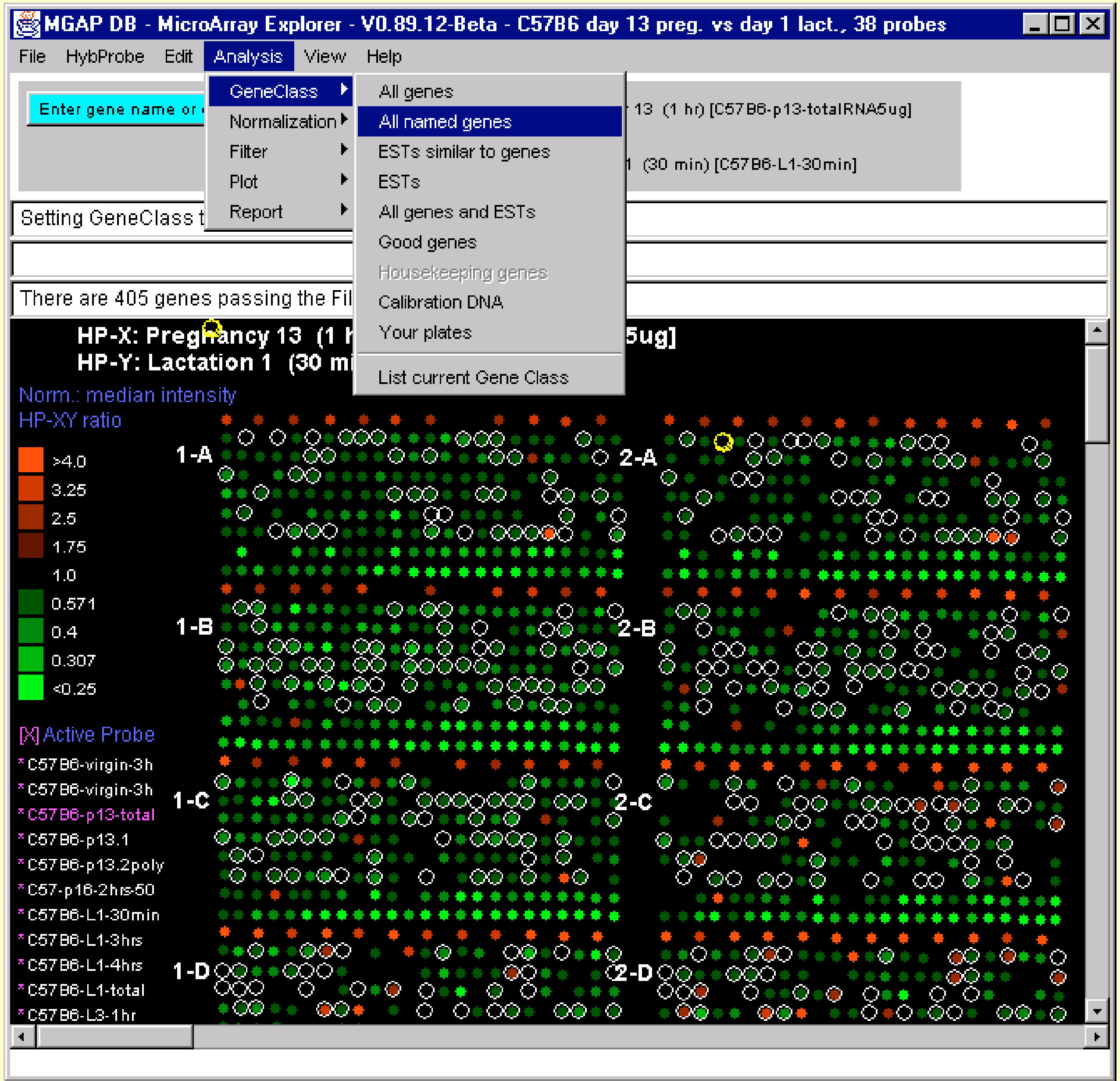
MGAP DB - MicroArray Explorer - V0.89.12-Beta - C57B6 day 13 preg. vs day 1 lact., 38 probes

Setting GeneClass t

There are 1076 genes passing the F

2b. Named Genes

- Specify sets of genes for *all named genes* indicated in ratio X/Y array plot by white circles

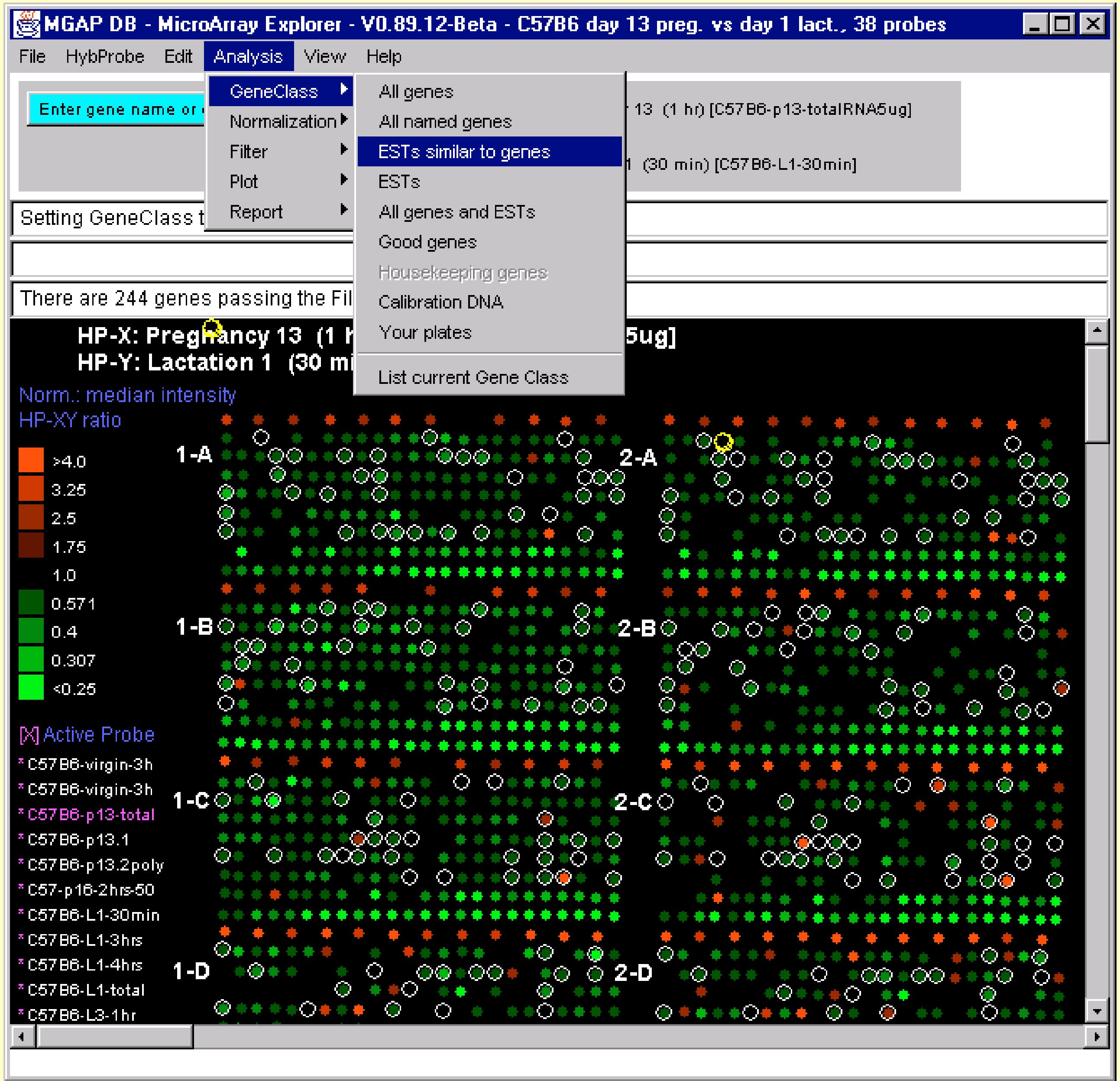


Setting GeneClass t

There are 405 genes passing the Fil

2c. ESTs similar to named genes

- Specify sets of genes for *all ESTs similar to named genes* indicated in the microarray by white circles



MGAP DB - MicroArray Explorer - V0.89.12-Beta - C57B6 day 13 preg. vs day 1 lact., 38 probes

Enter gene name or

13 (1 hr) [C57B6-p13-totalRNA5ug]

1 (30 min) [C57B6-L1-30min]

Setting GeneClass t

There are 244 genes passing the Fil

HP-X: Pregnancy 13 (1 h 5ug]

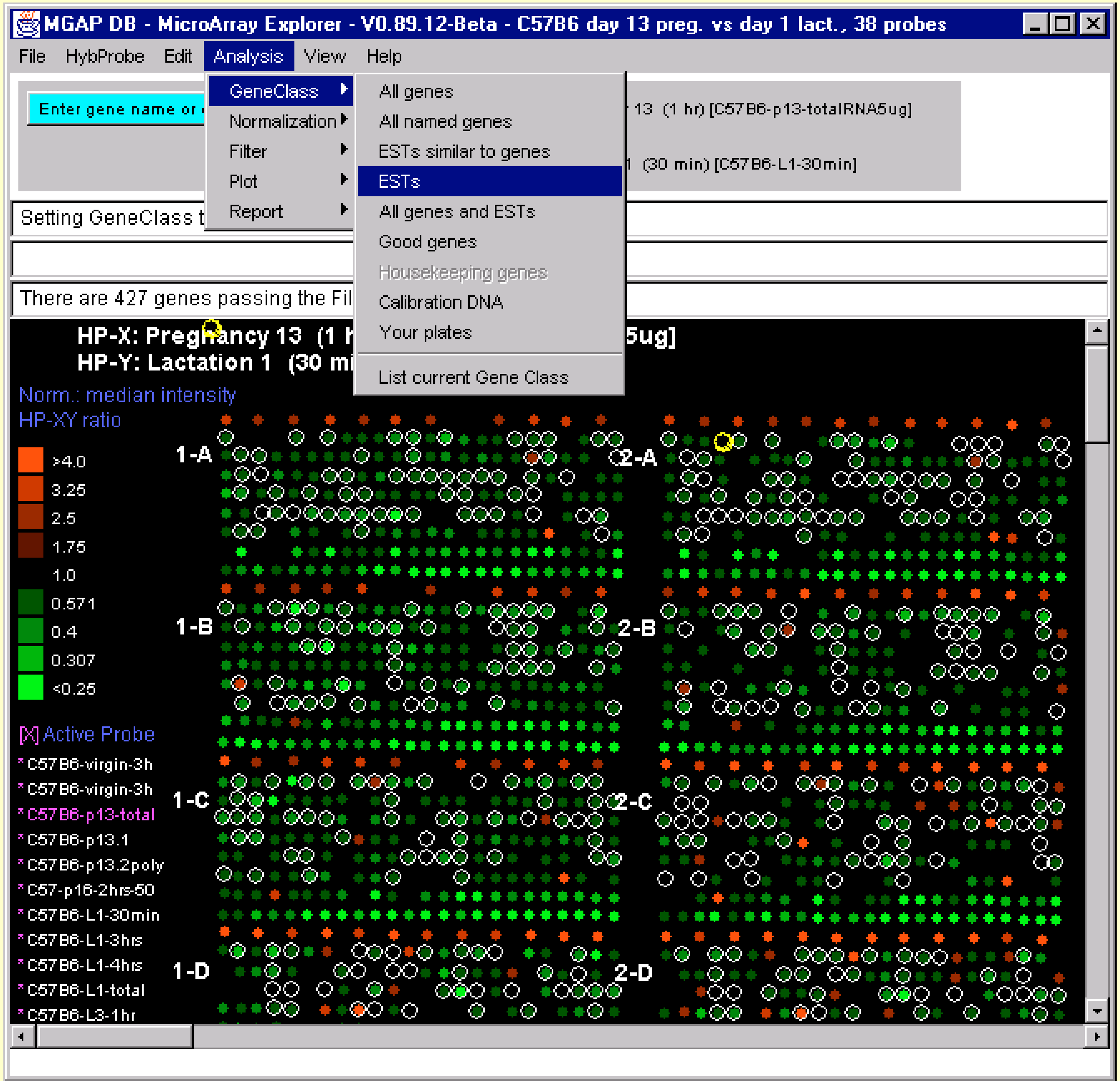
HP-Y: Lactation 1 (30 mi

Norm.: median intensity

HP-XY ratio

2d. Unknown ESTs

- Specify sets of genes for *unknown ESTs* indicated in the microarray by white circles



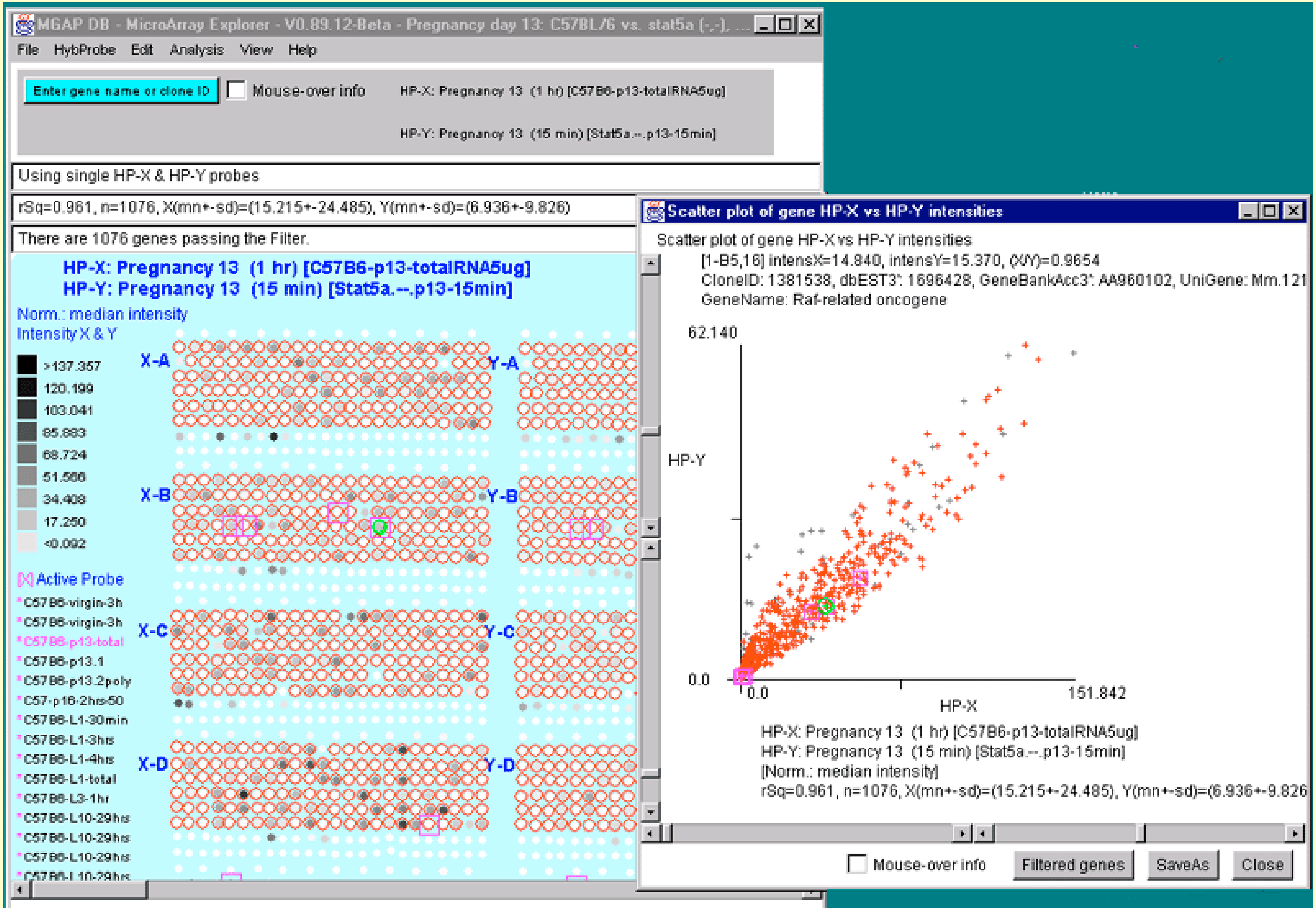
MGAP DB - MicroArray Explorer - V0.89.12-Beta - C57B6 day 13 preg. vs day 1 lact., 38 probes

Setting GeneClass t

There are 427 genes passing the Fil

3.a Scatter Plots of Two Conditions

- X-Y scatter plot of ‘sets’ of 2-probes C57B6 vs Stat5a (-) 13-day pregnancy in array [MGAP]. Current gene (green circle) & Edited Gene List (magenta squares) in plot



MGAP DB - MicroArray Explorer - V0.89.12-Beta - Pregnancy day 13: C57BL/6 vs. stat5a (-,-), ...

HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]

Using single HP-X & HP-Y probes

rSq=0.961, n=1076, X(mn+-sd)=(15.215+-24.485), Y(mn+-sd)=(6.936+-9.826)

There are 1076 genes passing the Filter.

HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]
HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]

Norm.: median intensity
Intensity X & Y

■	>137.357
■	120.199
■	103.041
■	85.883
■	68.724
■	51.566
■	34.408
■	17.250
■	<0.092

[X] Active Probe

- C57B6-virgin-3h
- C57B6-virgin-3h
- C57B6-p13-total
- C57B6-p13.1
- C57B6-p13.2poly
- C57-p16-2hrs-50
- C57B6-L1-30min
- C57B6-L1-3hrs
- C57B6-L1-4hrs
- C57B6-L1-total
- C57B6-L3-1hr
- C57B6-L10-29hrs
- C57B6-L10-29hrs
- C57B6-L10-29hrs
- C57B6-L10-29hrs

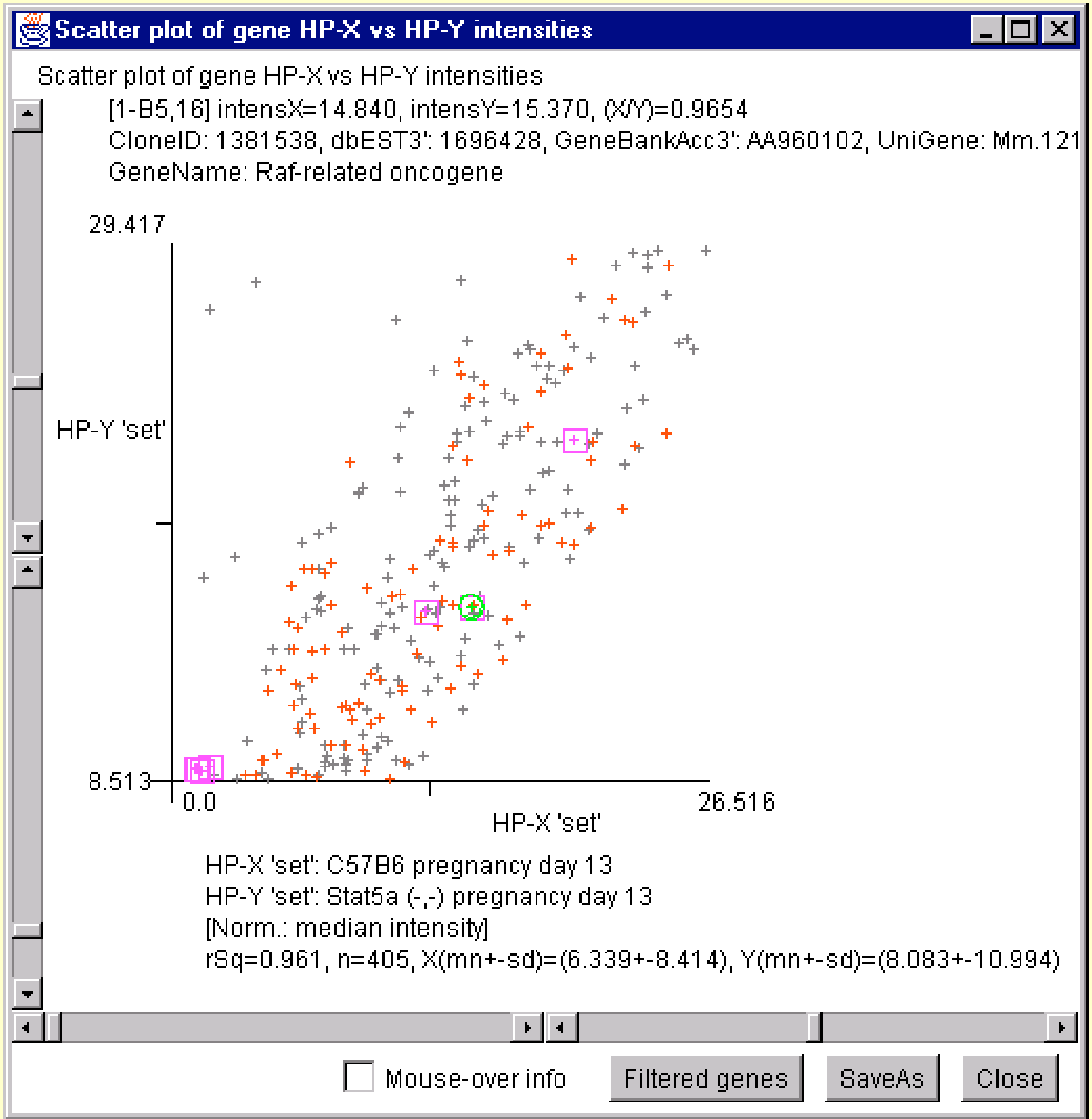
Scatter plot of gene HP-X vs HP-Y intensities

Scatter plot of gene HP-X vs HP-Y intensities
[1-B5,16] intensX=14.840, intensY=15.370, (X/Y)=0.9654
CloneID: 1381538, dbEST3': 1696428, GeneBankAcc3': AA960102, UniGene: Mm.121
GeneName: Raf-related oncogene

HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]
HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]
[Norm.: median intensity]
rSq=0.961, n=1076, X(mn+-sd)=(15.215+-24.485), Y(mn+-sd)=(6.936+-9.826

3.b Zoomed X-Y Scatter Plot (of 3.a)

- Zoomed in on Raf-related oncogene using scrollbars. Genes not passing Filter are grayed out in the plot

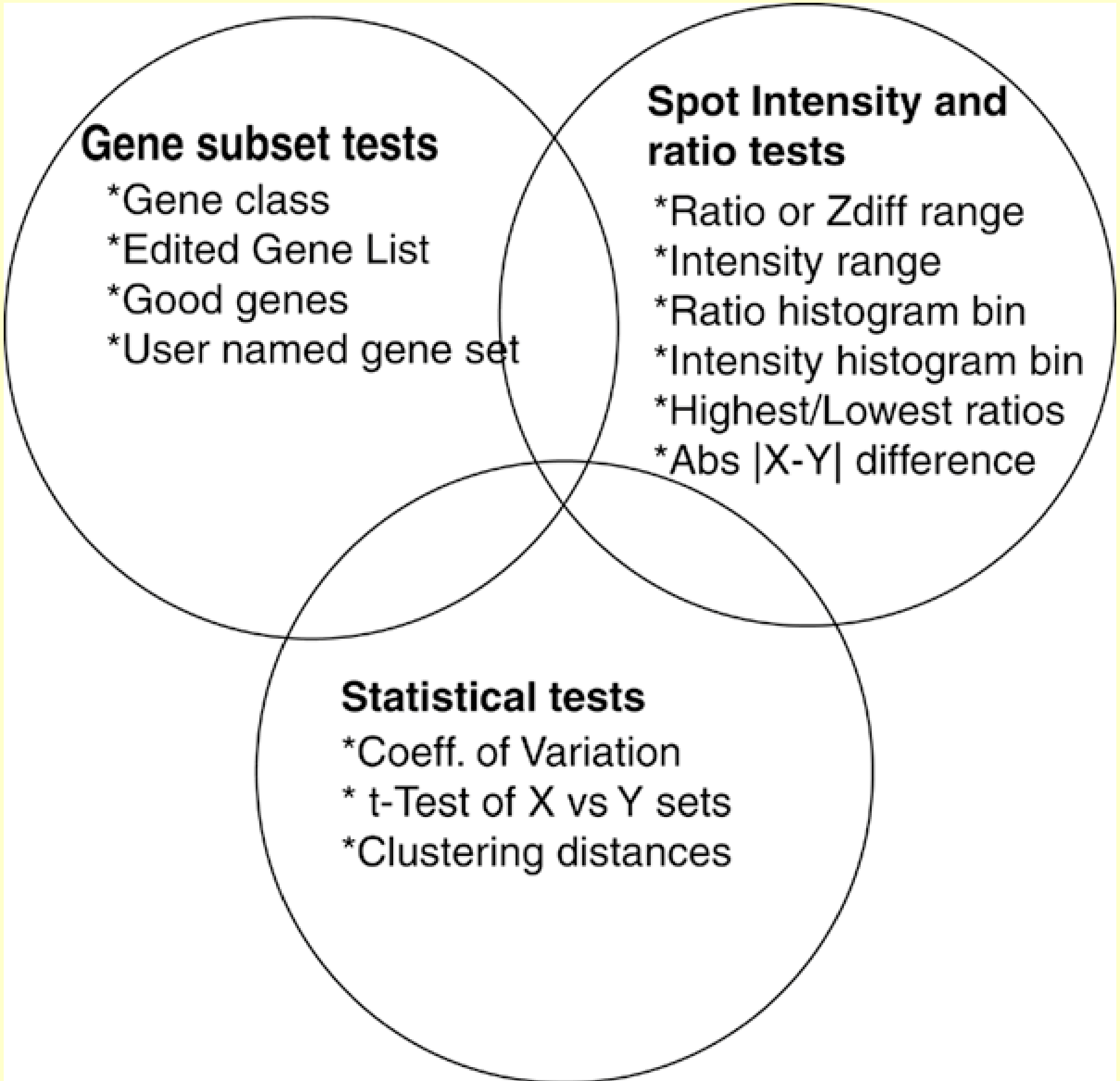


```
Scatter plot of gene HP-X vs HP-Y intensities
        [1-B5,16] intensX=14.840, intensY=15.370, (X/Y)=0.9654
        CloneID: 1381538, dbEST3': 1696428, GeneBankAcc3': AA960102, UniGene: Mm.121
        GeneName: Raf-related oncogene
```

```
HP-X 'set': C57B6 pregnancy day 13
HP-Y 'set': Stat5a (-,-) pregnancy day 13
[Norm.: median intensity]
rSq=0.961, n=405, X(mn+-sd)=(6.339+-8.414), Y(mn+-sd)=(8.083+-10.994)
```

4a. Gene Data Filter is Intersection of Tests

- Current set of genes is **intersection** of gene sets each passing selected filter tests
- Filtered gene subset is used as **pre-filter** for subsequent clustering, plots, and tables
- Changing any filter parameters causes the data filter to be re-computed



Gene subset tests

*Gene class
*Edited Gene List
*Good genes
*User named gene set

Spot Intensity and ratio tests

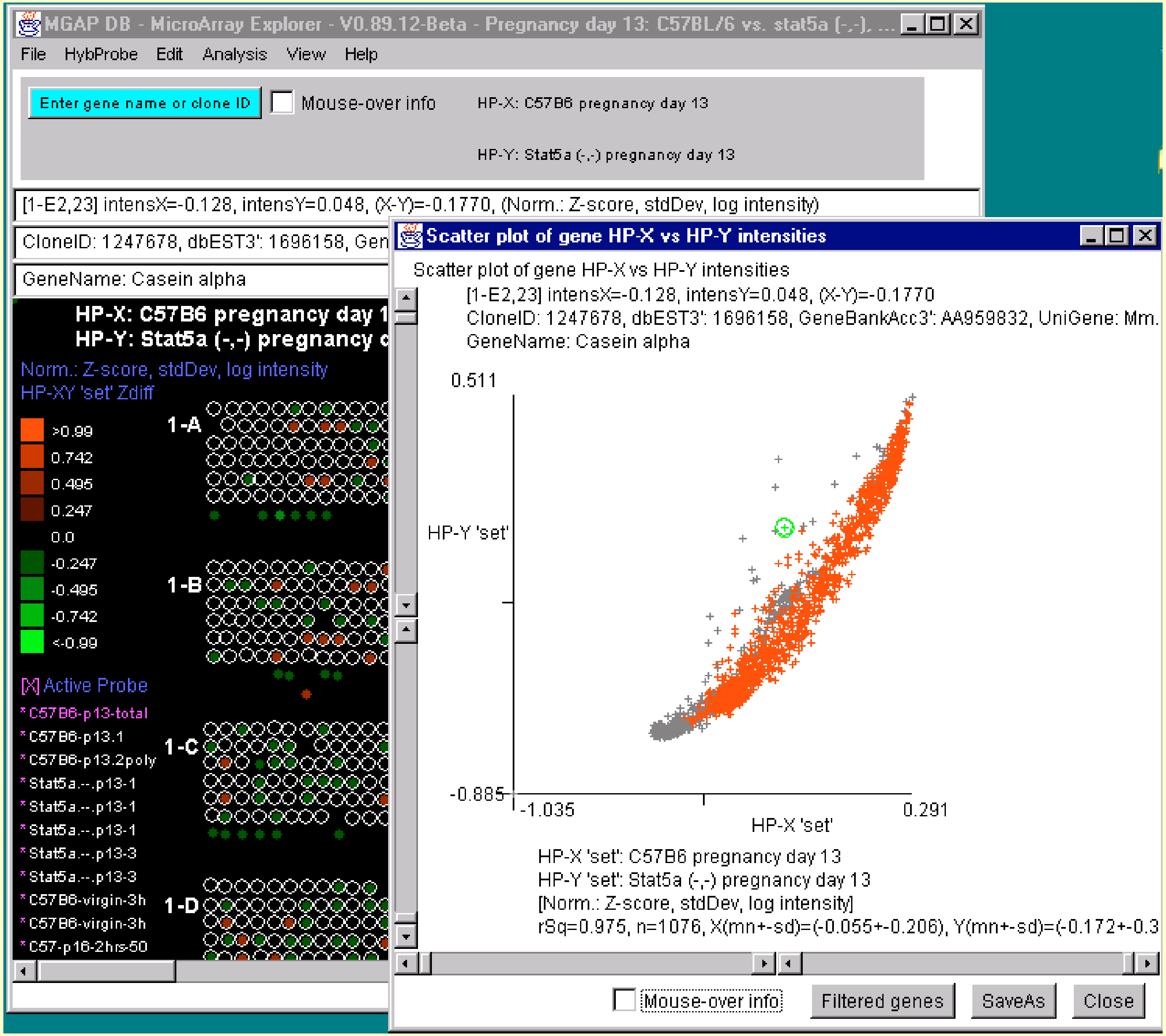
*Ratio or Zdiff range
*Intensity range
*Ratio histogram bin
*Intensity histogram bin
*Highest/Lowest ratios
*Abs |X-Y| difference

Statistical tests

*Coeff. of Variation
* t-Test of X vs Y sets
*Clustering distances

4.b Genes Filtered by Gene Class Set

- Genes class subset *named genes and ESTs* in both array & scatter plot normalized by Zscore of log intensity.



MGAP DB - MicroArray Explorer - V0.89.12-Beta - Pregnancy day 13: C57BL/6 vs. stat5a (-,-), ...

HP-X: C57B6 pregnancy day 13

HP-Y: Stat5a (-,-) pregnancy day 13

[1-E2,23] intensX=-0.128, intensY=0.048, (X-Y)=-0.1770, (Norm.: Z-score, stdDev, log intensity)

CloneID: 1247678, dbEST3': 1696158, Gen

GeneName: Casein alpha

HP-X: C57B6 pregnancy day 1

HP-Y: Stat5a (-,-) pregnancy d

Norm.: Z-score, stdDev, log intensity

HP-XY 'set' Zdiff

Scatter plot of gene HP-X vs HP-Y intensities

[1-E2,23] intensX=-0.128, intensY=0.048, (X-Y)=-0.1770

CloneID: 1247678, dbEST3': 1696158, GeneBankAcc3': AA959832, UniGene: Mm.

GeneName: Casein alpha

HP-X 'set': C57B6 pregnancy day 13

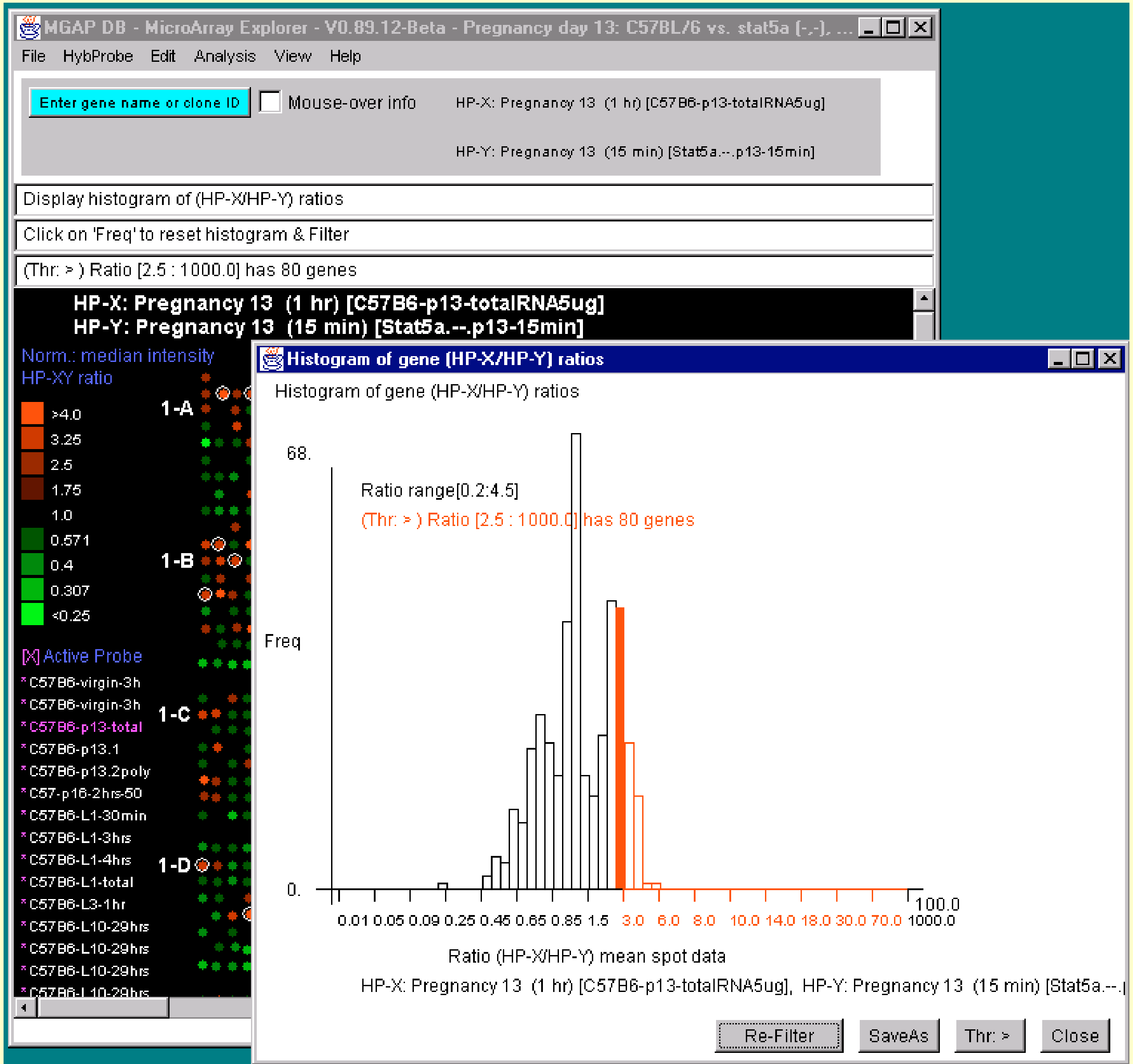
HP-Y 'set': Stat5a (-,-) pregnancy day 13

[Norm.: Z-score, stdDev, log intensity]

rSq=0.975, n=1076, X(mn+-sd)=(-0.055+-0.206), Y(mn+-sd)=(-0.172+-0.3

4.c Genes Filtered by Ratio-Histogram Bin

- Genes filtered by HP-X/HP-Y C57B5-preg / Stat5a(-,-) ratio-histogram bin-range [2.5:1000]. Histogram is for all named genes and for ESTs.



MGAP DB - MicroArray Explorer - V0.89.12-Beta - Pregnancy day 13: C57BL/6 vs. stat5a (-,-), ...

HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]

Display histogram of (HP-X/HP-Y) ratios

Click on 'Freq' to reset histogram & Filter

(Thr: >) Ratio [2.5 : 1000.0] has 80 genes

HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]
HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]

Norm.: median intensity
HP-XY ratio

🟥	>4.0
🟥	3.25
🟥	2.5
🟥	1.75
	1.0
🟩	0.571
🟩	0.4
🟩	0.307
🟩	<0.25

[X] Active Probe

- *C57B6-virgin-3h
- *C57B6-virgin-3h
- *C57B6-p13-total
- *C57B6-p13.1
- *C57B6-p13.2poly
- *C57-p16-2hrs-50
- *C57B6-L1-30min
- *C57B6-L1-3hrs
- *C57B6-L1-4hrs
- *C57B6-L1-total
- *C57B6-L3-1hr
- *C57B6-L10-29hrs
- *C57B6-L10-29hrs
- *C57B6-L10-29hrs
- *C57B6-L10-29hrs

Histogram of gene (HP-X/HP-Y) ratios

Histogram of gene (HP-X/HP-Y) ratios

Ratio range[0.2:4.5]

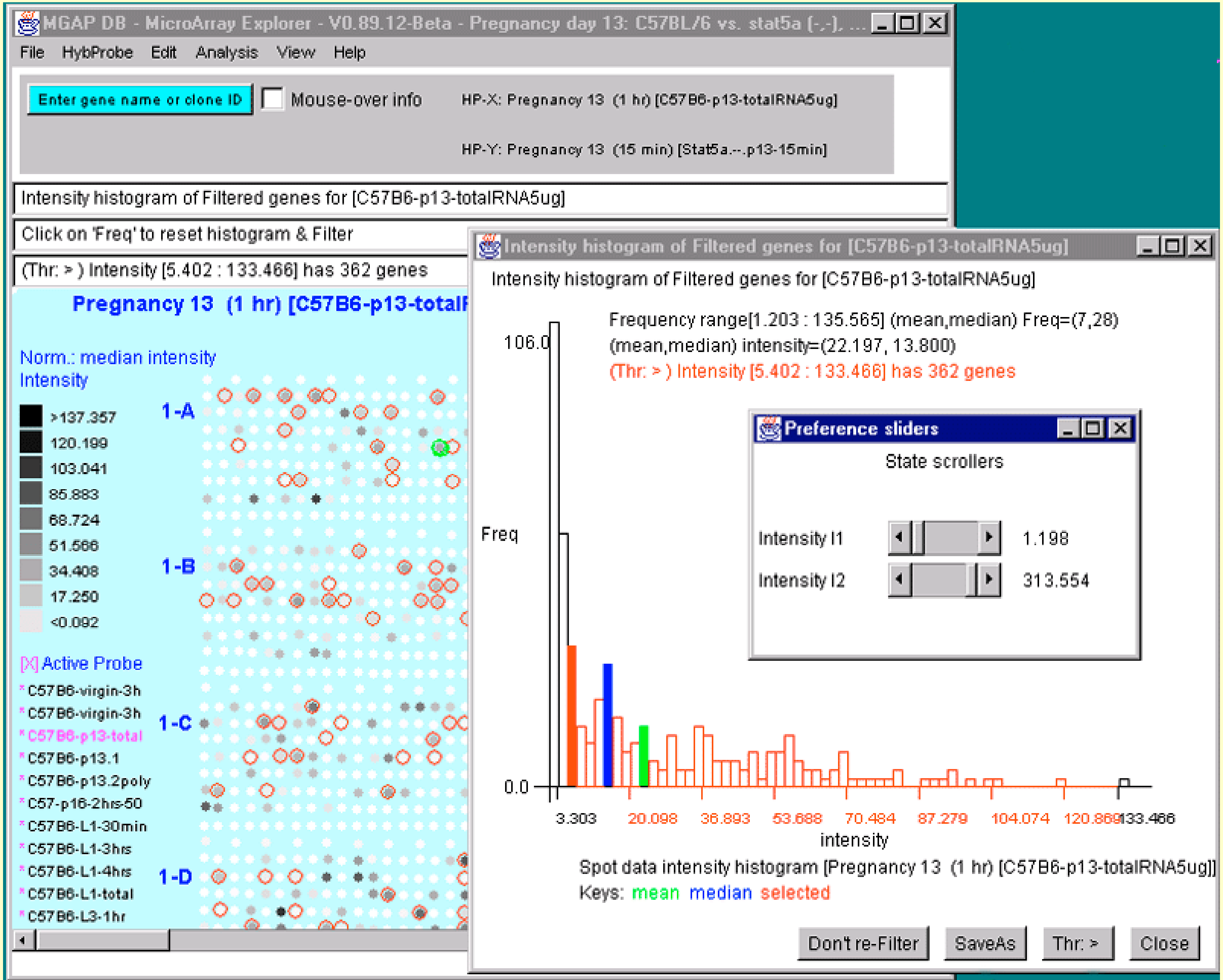
(Thr: >) Ratio [2.5 : 1000.0] has 80 genes

Ratio (HP-X/HP-Y) mean spot data

HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug], HP-Y: Pregnancy 13 (15 min) [Stat5a.--.

4.d Genes Filtered by Intensity-Histogram Bin

- Genes filtered by intensity to remove low signal strength sample genes.



HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]

Intensity histogram of Filtered genes for [C57B6-p13-totalRNA5ug]

Click on 'Freq' to reset histogram & Filter

(Thr: >) Intensity [5.402 : 133.466] has 362 genes

Pregnancy 13 (1 hr) [C57B6-p13-totalR

Norm.: median intensity

Intensity

■ >137.357
■ 120.199
■ 103.041
■ 85.883
■ 68.724
■ 51.566
■ 34.408
■ 17.250
■ <0.092

[X] Active Probe

- C57B6-virgin-3h
- C57B6-virgin-3h
- C57B6-p13-total
- C57B6-p13.1
- C57B6-p13.2poly
- C57-p16-2hrs-50
- C57B6-L1-30min
- C57B6-L1-3hrs
- C57B6-L1-4hrs
- C57B6-L1-total
- C57B6-L3-1hr

Intensity histogram of Filtered genes for [C57B6-p13-totalRNA5ug]

Frequency range[1.203 : 135.565] (mean,median) Freq=(7,28)
(mean,median) intensity=(22.197, 13.800)
(Thr: >) Intensity [5.402 : 133.466] has 362 genes

State scrollers

Intensity I1: 1.198

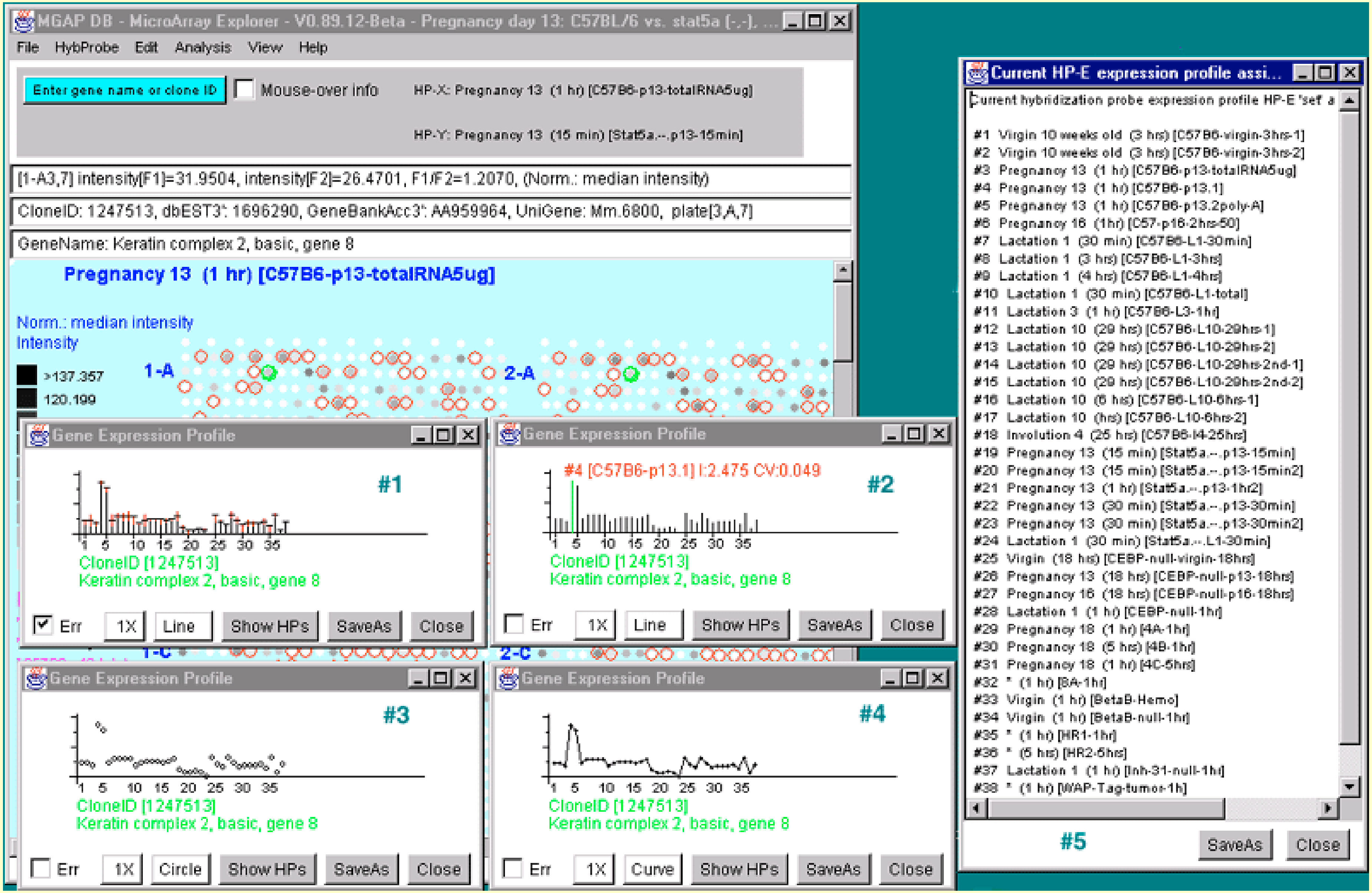
Intensity I2: 313.554

Spot data intensity histogram [Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]]

Keys: mean median selected

5.a Expression Profile Plots of N-conditions

- Expression profile plot of 38-conditions of current gene (green) . Note numbered list of probes. Intensity data for probe #4 is indicated in red - by clicking on a line in plot



MGAP DB - MicroArray Explorer - V0.89.12-Beta - Pregnancy day 13: C57BL/6 vs. stat5a [-,-]. ...

HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]

[1-A3,7] intensity[F1]=31.9504, intensity[F2]=26.4701, F1/F2=1.2070, (Norm.: median intensity)

CloneID: 1247513, dbEST3': 1696290, GeneBankAcc3': AA959964, UniGene: Mm.6800, plate[3,A,7]

GeneName: Keratin complex 2, basic, gene 8

Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

Norm.: median intensity

#4 [C57B6-p13.1] I:2.475 CV:0.049

CloneID [1247513]
Keratin complex 2, basic, gene 8

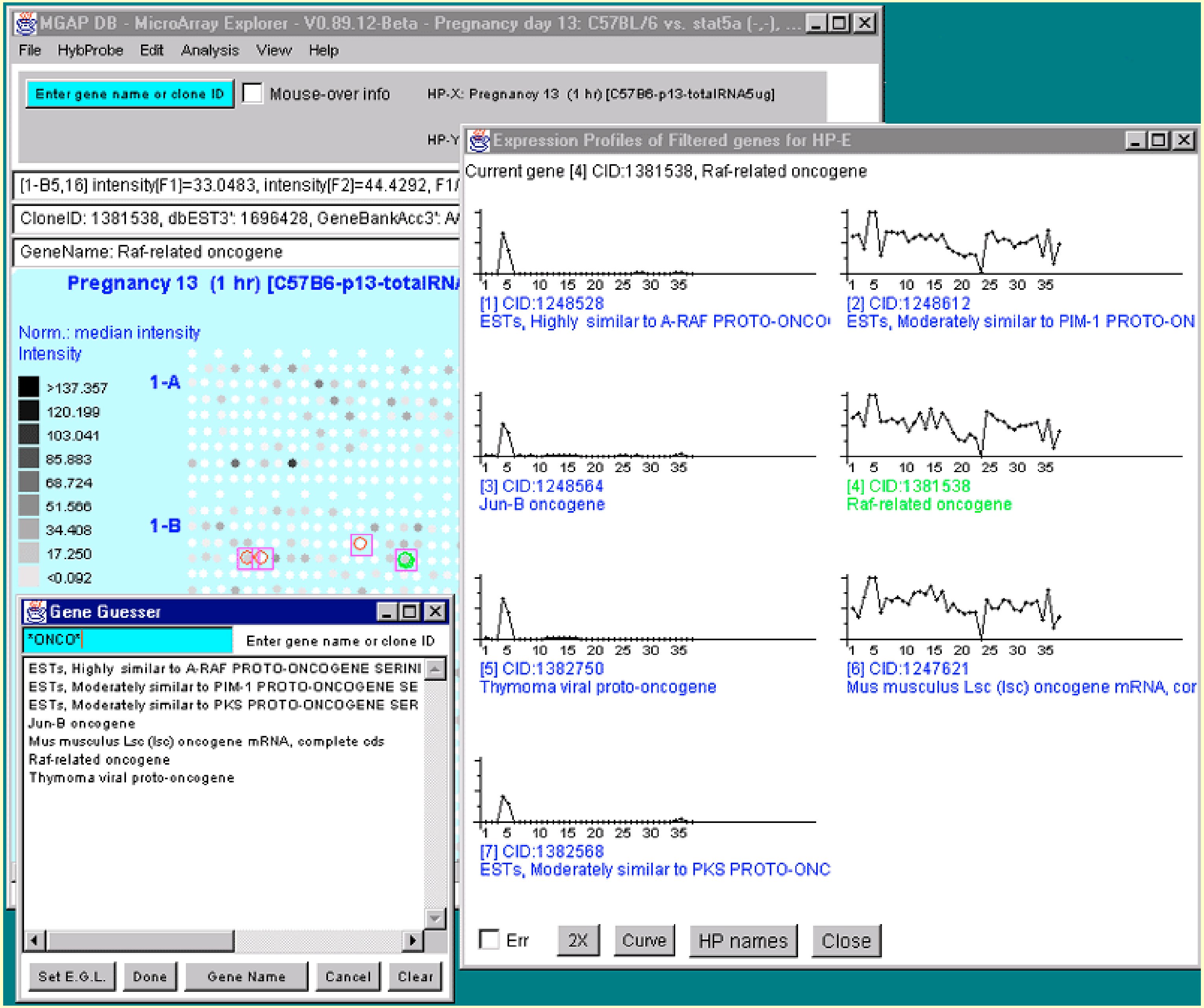
Current HP-E expression profile assi...

Current hybridization probe expression profile HP-E 'set' a

#1 Virgin 10 weeks old (3 hrs) [C57B6-virgin-3hrs-1]
#2 Virgin 10 weeks old (3 hrs) [C57B6-virgin-3hrs-2]
#3 Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]
#4 Pregnancy 13 (1 hr) [C57B6-p13.1]
#5 Pregnancy 13 (1 hr) [C57B6-p13.2poly-A]
#6 Pregnancy 16 (1hr) [C57-p16-2hrs-50]
#7 Lactation 1 (30 min) [C57B6-L1-30min]
#8 Lactation 1 (3 hrs) [C57B6-L1-3hrs]
#9 Lactation 1 (4 hrs) [C57B6-L1-4hrs]
#10 Lactation 1 (30 min) [C57B6-L1-total]
#11 Lactation 3 (1 hr) [C57B6-L3-1hr]
#12 Lactation 10 (29 hrs) [C57B6-L10-29hrs-1]
#13 Lactation 10 (29 hrs) [C57B6-L10-29hrs-2]
#14 Lactation 10 (29 hrs) [C57B6-L10-29hrs-2nd-1]
#15 Lactation 10 (29 hrs) [C57B6-L10-29hrs-2nd-2]
#16 Lactation 10 (6 hrs) [C57B6-L10-6hrs-1]
#17 Lactation 10 (hrs) [C57B6-L10-6hrs-2]
#18 Involution 4 (25 hrs) [C57B6-I4-25hrs]
#19 Pregnancy 13 (15 min) [Stat5a.--.p13-15min]
#20 Pregnancy 13 (15 min) [Stat5a.--.p13-15min2]
#21 Pregnancy 13 (1 hr) [Stat5a.--.p13-1hr2]
#22 Pregnancy 13 (30 min) [Stat5a.--.p13-30min]
#23 Pregnancy 13 (30 min) [Stat5a.--.p13-30min2]
#24 Lactation 1 (30 min) [Stat5a.--.L1-30min]
#25 Virgin (18 hrs) [CEBP-null-virgin-18hrs]
#26 Pregnancy 13 (18 hrs) [CEBP-null-p13-18hrs]
#27 Pregnancy 16 (18 hrs) [CEBP-null-p16-18hrs]
#28 Lactation 1 (1 hr) [CEBP-null-1hr]
#29 Pregnancy 18 (1 hr) [4A-1hr]
#30 Pregnancy 18 (5 hrs) [4B-1hr]
#31 Pregnancy 18 (1 hr) [4C-5hrs]
#32 * (1 hr) [BA-1hr]
#33 Virgin (1 hr) [BetaB-Hemo]
#34 Virgin (1 hr) [BetaB-null-1hr]
#35 * (1 hr) [HR1-1hr]
#36 * (5 hrs) [HR2-5hrs]
#37 Lactation 1 (1 hr) [Inh-31-null-1hr]
#38 * (1 hr) [WAP-Tag-tumor-1h]

5.b List of Expression Profile Plots

- Scrollable list of EP plots for onco and proto-oncogenes in EGL for MGAP database



HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug]

[1-B5,16] intensity[F1]=33.0483, intensity[F2]=44.4292, F1/

CloneID: 1381538, dbEST3': 1696428, GeneBankAcc3': A/

GeneName: Raf-related oncogene

Pregnancy 13 (1 hr) [C57B6-p13-totalRN/

ONCO

ESTs, Highly similar to A-RAF PROTO-ONCOGENE SERINI
ESTs, Moderately similar to PIM-1 PROTO-ONCOGENE SE
ESTs, Moderately similar to PKS PROTO-ONCOGENE SER
Jun-B oncogene
Mus musculus Lsc (lsc) oncogene mRNA, complete cds
Raf-related oncogene
Thymoma viral proto-oncogene

Current gene [4] CID:1381538, Raf-related oncogene

[1] CID:1248528
ESTs, Highly similar to A-RAF PROTO-ONCO

[2] CID:1248612
ESTs, Moderately similar to PIM-1 PROTO-ON

[3] CID:1248564
Jun-B oncogene

[4] CID:1381538
Raf-related oncogene

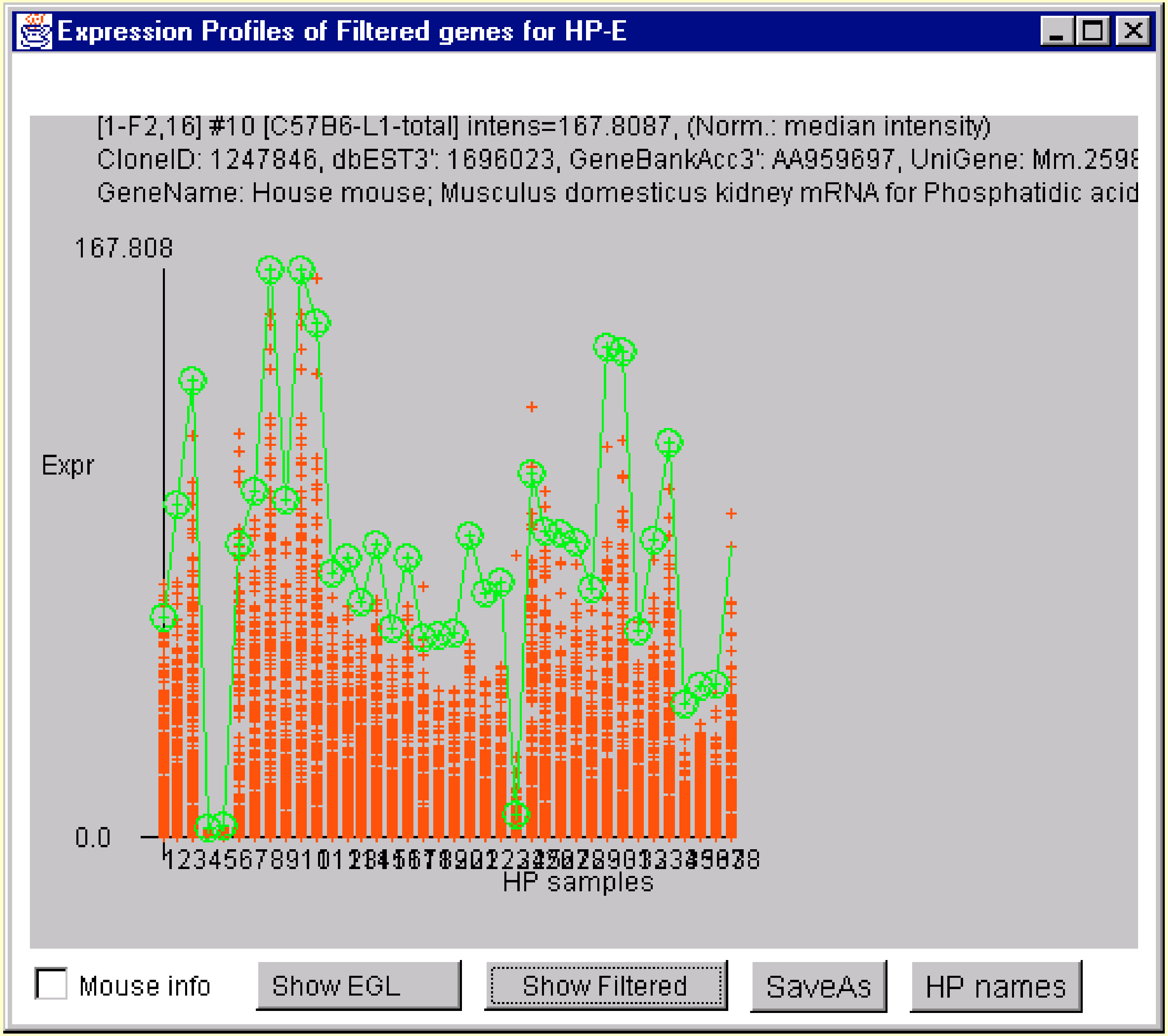
[5] CID:1382750
Thymoma viral proto-oncogene

[6] CID:1247621
Mus musculus Lsc (lsc) oncogene mRNA, cor

[7] CID:1382568
ESTs, Moderately similar to PKS PROTO-ONC

5.c Expression Profile Overlay Plots

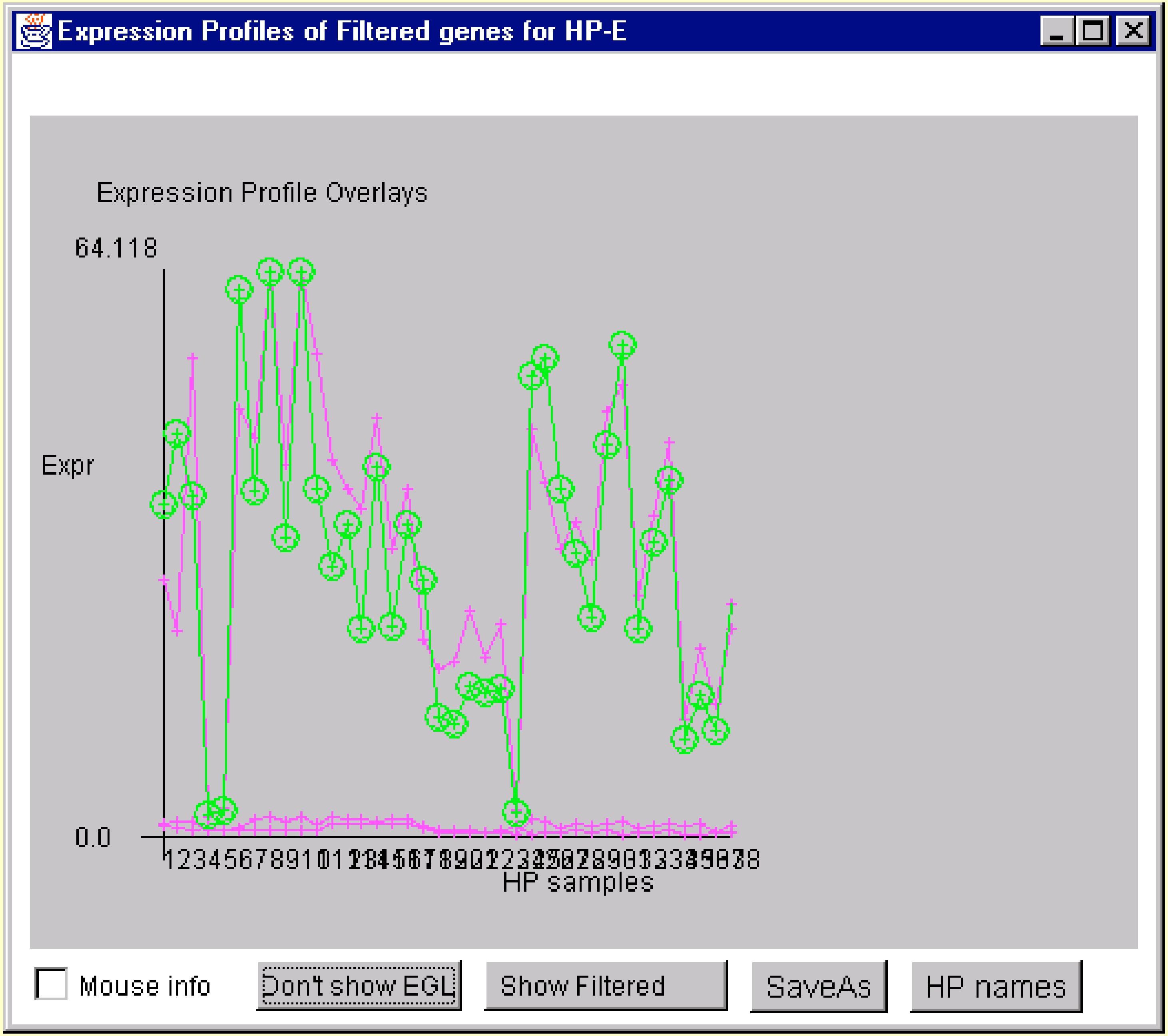
- Overlay EP plots of multiple genes showing current gene for MGAP database



[1-F2,16] #10 [C57B6-L1-total] intens=167.8087, (Norm.: median intensity)
CloneID: 1247846, dbEST3': 1696023, GeneBankAcc3': AA959697, UniGene: Mm.259
GeneName: House mouse; Musculus domesticus kidney mRNA for Phosphatidic acid

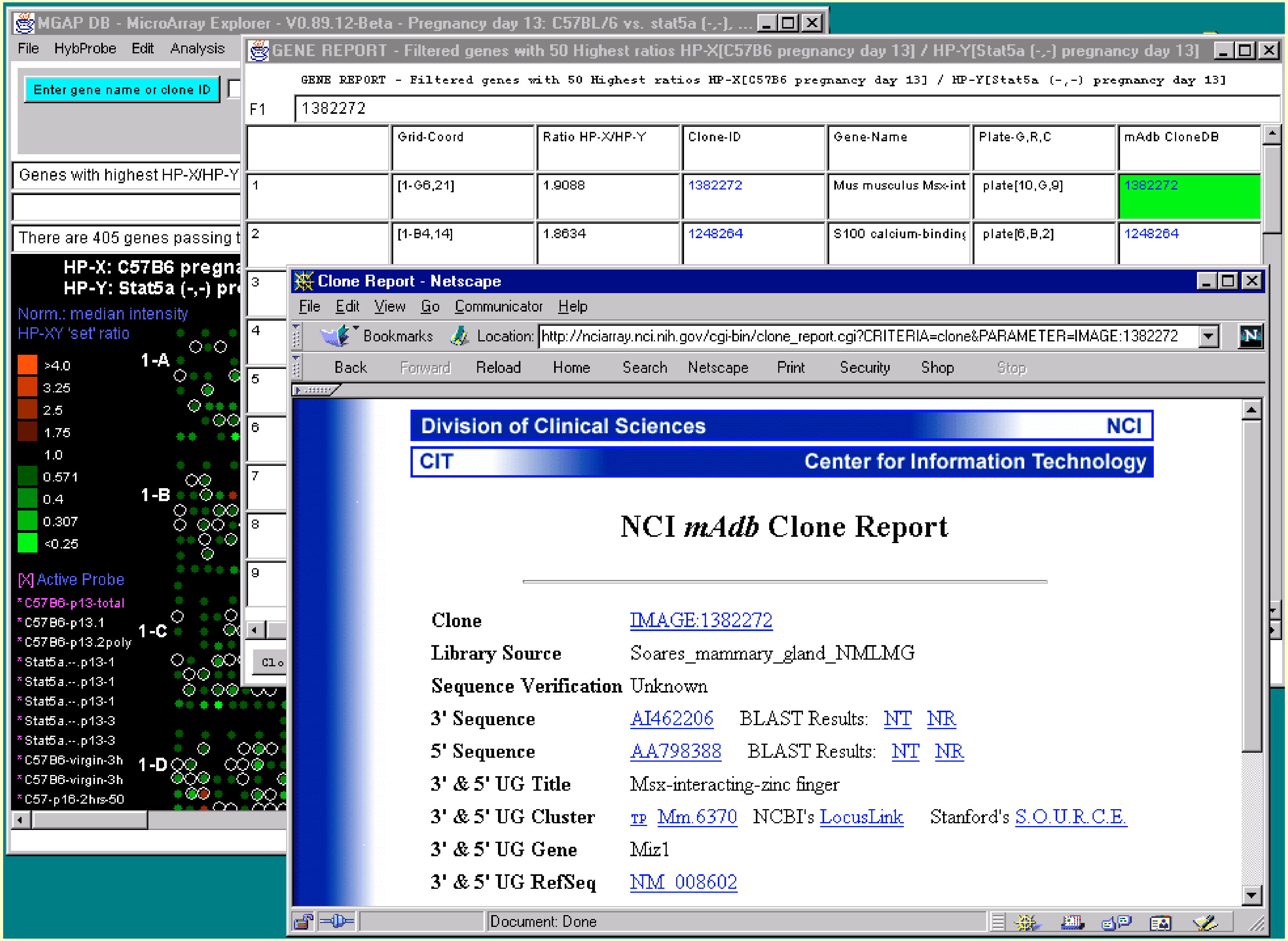
5.d Expression Profile Overlay Plots

- Overlay EP plots for onco and proto-oncogenes in EGL for MGAP database



6.a Scrollable Dynamic Gene Reports

- Scrollable gene report of highest ratio genes & NCI mAdb pop up Web browser page (foreground) of particular gene. Clicking on blue hypertext cell in gene report (middle) invokes pop up web page (NCI mAdb Clone Report shown here)



MGAP DB - MicroArray Explorer - V0.89.12-Beta - Pregnancy day 13: C57BL/6 vs. stat5a (-,-), ...

Genes with highest HP-X/HP-Y

There are 405 genes passing t

HP-X: C57B6 pregna

HP-Y: Stat5a (-,-) pre

Norm.: median intensity

HP-XY 'set' ratio

GENE REPORT - Filtered genes with 50 Highest ratios HP-X[C57B6 pregnancy day 13] / HP-Y[Stat5a (-,-) pregnancy day 13]

```
GENE REPORT - Filtered genes with 50 Highest ratios HP-X[C57B6 pregnancy day 13] / HP-Y[Stat5a (-,-) pregnancy day 13]
```

F1 1382272

	Grid-Coord	Ratio HP-X/HP-Y	Clone-ID	Gene-Name	Plate-G,R,C	mAdb CloneDB
1	[1-G6,21]	1.9088	1382272	Mus musculus Msx-int	plate[10,G,9]	1382272
2	[1-B4,14]	1.8634	1248264	S100 calcium-binding	plate[6,B,2]	1248264

Clone Report - Netscape

Location: http://nciarray.nci.nih.gov/cgi-bin/clone_report.cgi?CRITERIA=clone&PARAMETER=IMAGE:1382272

Division of Clinical Sciences NCI

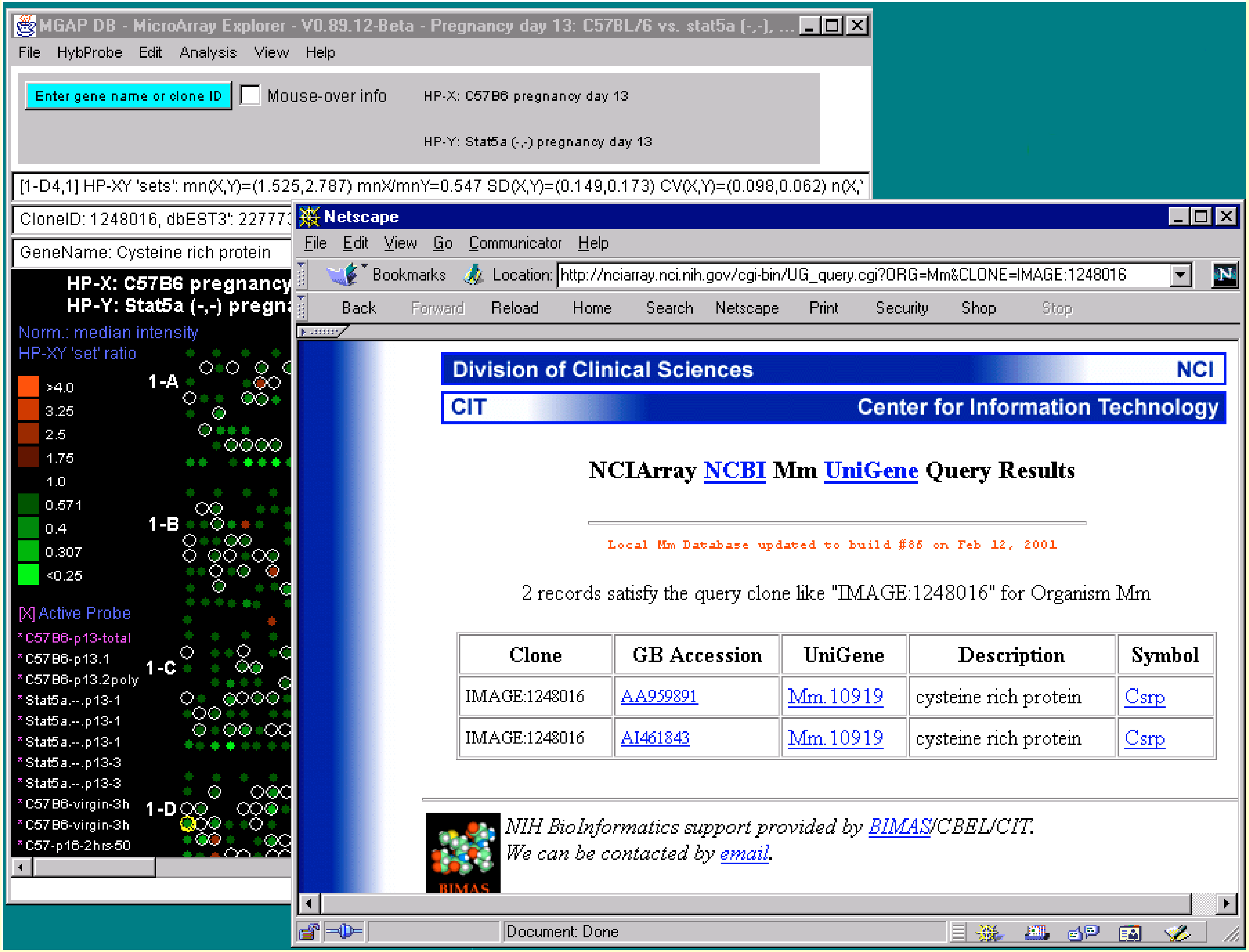
CIT Center for Information Technology

NCI *mAdb* Clone Report

Clone	IMAGE:1382272
Library Source	Soares_mammary_gland_NMLMG
Sequence Verification	Unknown
3' Sequence	AI462206 BLAST Results: NT NR
5' Sequence	AA798388 BLAST Results: NT NR
3' & 5' UG Title	Msx-interacting-zinc finger
3' & 5' UG Cluster	TP Mm.6370 NCBI's LocusLink Stanford's S.O.U.R.C.E.
3' & 5' UG Gene	Miz1
3' & 5' UG RefSeq	NM_008602

Document: Done

6.b Scrollable Dynamic Gene Reports - UniGene Report



MGAP DB - MicroArray Explorer - V0.89.12-Beta - Pregnancy day 13: C57BL/6 vs. stat5a (-,-), ...

HP-X: C57B6 pregnancy day 13

HP-Y: Stat5a (-,-) pregnancy day 13

[1-D4,1] HP-XY 'sets': mn(X,Y)=(1.525,2.787) mnX/mnY=0.547 SD(X,Y)=(0.149,0.173) CV(X,Y)=(0.098,0.062) n(X,'

CloneID: 1248016, dbEST3': 22777

GeneName: Cysteine rich protein

Location: http://nciarray.nci.nih.gov/cgi-bin/UG_query.cgi?ORG=Mm&CLONE=IMAGE:1248016

Division of Clinical Sciences **NCI**

CIT **Center for Information Technology**

NCIArray NCBI Mm UniGene Query Results

Local Mm Database updated to build #86 on Feb 12, 2001

2 records satisfy the query clone like "IMAGE:1248016" for Organism Mm

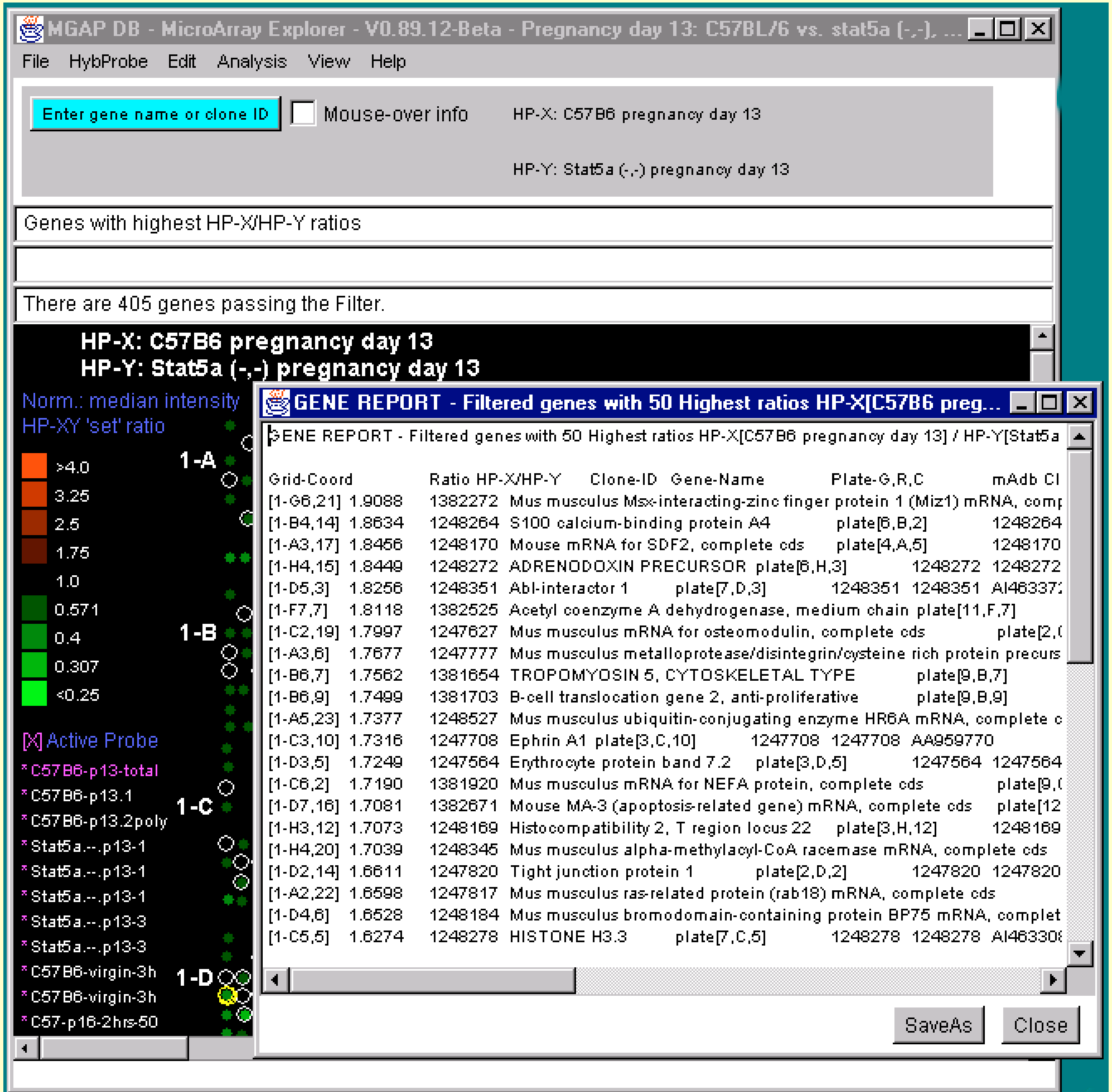
Clone	GB Accession	UniGene	Description	Symbol
IMAGE:1248016	AA959891	Mm.10919	cysteine rich protein	Csrp
IMAGE:1248016	AI461843	Mm.10919	cysteine rich protein	Csrp

NIH BioInformatics support provided by BIMAS/CBEL/CIT.
We can be contacted by email.

Document: Done

6.c Gene Reports are Exportable to Excel

- Tab-delimited gene reports are exportable to Excel using cut & paste or SaveAs



Genes with highest HP-X/HP-Y ratios

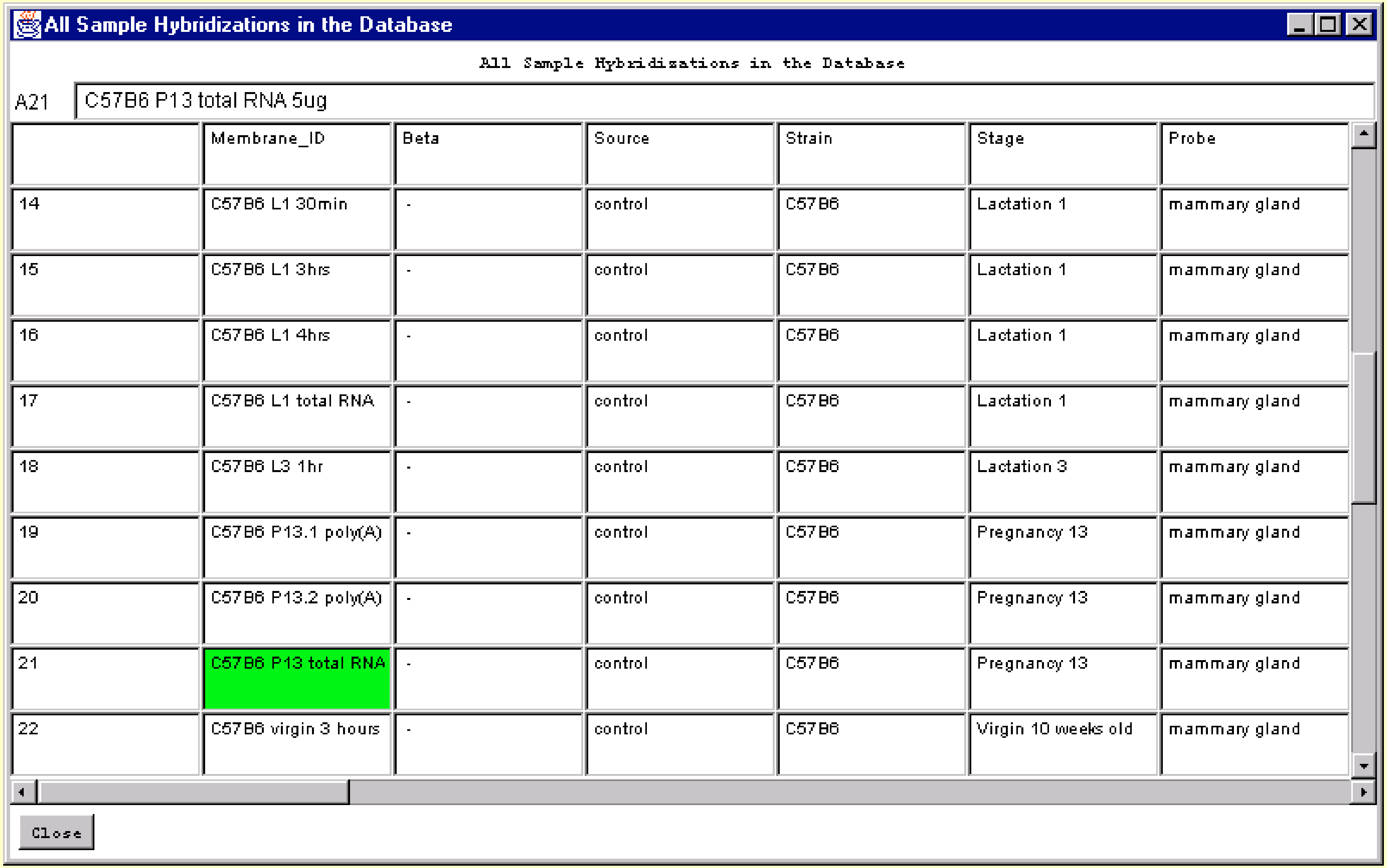
There are 405 genes passing the Filter.

```
GENE REPORT - Filtered genes with 50 Highest ratios HP-X[C57B6 pregnancy day 13] / HP-Y[Stat5a

Grid-Coord          Ratio HP-X/HP-Y     Clone-ID  Gene-Name            Plate-G,R,C          mAdb Cl
[1-G6,21] 1.9088    1382272  Mus musculus Msx-interacting-zinc finger protein 1 (Miz1) mRNA, com
[1-B4,14] 1.8634    1248264  S100 calcium-binding protein A4          plate[6,B,2]          1248264
[1-A3,17] 1.8456    1248170  Mouse mRNA for SDF2, complete cds        plate[4,A,5]          1248170
[1-H4,15] 1.8449    1248272  ADRENODOXIN PRECURSOR  plate[6,H,3]               1248272  1248272
[1-D5,3]  1.8256    1248351  Abl-interactor 1         plate[7,D,3]          1248351  1248351  AI46337
[1-F7,7]  1.8118    1382525  Acetyl coenzyme A dehydrogenase, medium chain  plate[11,F,7]
[1-C2,19] 1.7997    1247627  Mus musculus mRNA for osteomodulin, complete cds              plate[2,
[1-A3,6]  1.7677    1247777  Mus musculus metalloprotease/disintegrin/cysteine rich protein precurs
[1-B6,7]  1.7562    1381654  TROPOMYOSIN 5, CYTOSKELETAL TYPE                plate[9,B,7]
[1-B6,9]  1.7499    1381703  B-cell translocation gene 2, anti-proliferative  plate[9,B,9]
[1-A5,23] 1.7377    1248527  Mus musculus ubiquitin-conjugating enzyme HR6A mRNA, complete 
[1-C3,10] 1.7316    1247708  Ephrin A1  plate[3,C,10]          1247708  1247708  AA959770
[1-D3,5]  1.7249    1247564  Erythrocyte protein band 7.2     plate[3,D,5]          1247564  1247564
[1-C6,2]  1.7190    1381920  Mus musculus mRNA for NEFA protein, complete cds              plate[9,
[1-D7,16] 1.7081    1382671  Mouse MA-3 (apoptosis-related gene) mRNA, complete cds    plate[12
[1-H3,12] 1.7073    1248169  Histocompatibility 2, T region locus 22     plate[3,H,12]          1248169
[1-H4,20] 1.7039    1248345  Mus musculus alpha-methylacyl-CoA racemase mRNA, complete cds
[1-D2,14] 1.6611    1247820  Tight junction protein 1         plate[2,D,2]          1247820  1247820
[1-A2,22] 1.6598    1247817  Mus musculus ras-related protein (rab18) mRNA, complete cds
[1-D4,6]  1.6528    1248184  Mus musculus bromodomain-containing protein BP75 mRNA, complet
[1-C5,5]  1.6274    1248278  HISTONE H3.3         plate[7,C,5]          1248278  1248278  AI46330
```

6d. Sample Information Array Reports

- Details are available on all hybridized array samples



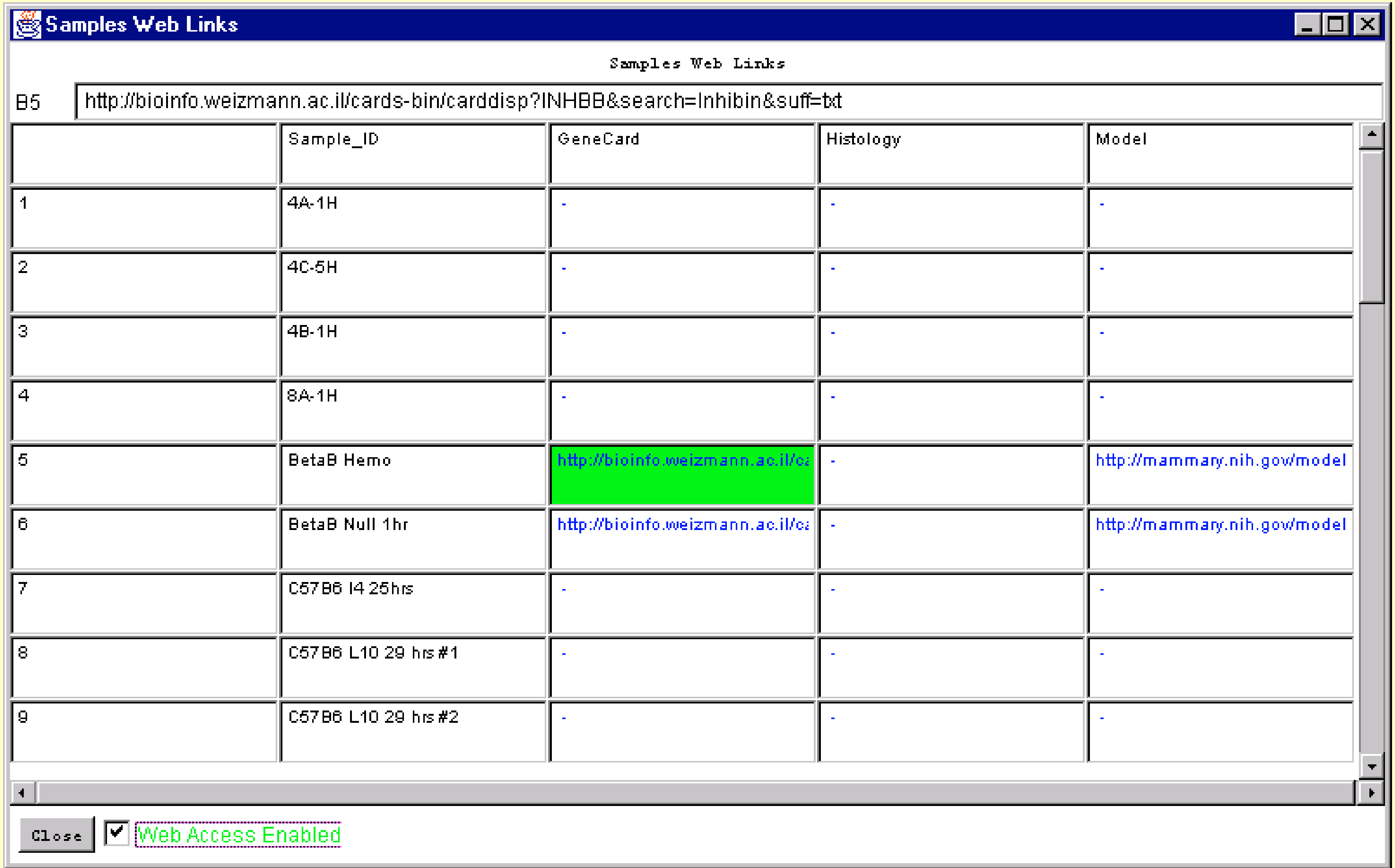
All Sample Hybridizations in the Database

A21 C57B6 P13 total RNA 5ug

	Membrane_ID	Beta	Source	Strain	Stage	Probe
14	C57B6 L1 30min	-	control	C57B6	Lactation 1	mammary gland
15	C57B6 L1 3hrs	-	control	C57B6	Lactation 1	mammary gland
16	C57B6 L1 4hrs	-	control	C57B6	Lactation 1	mammary gland
17	C57B6 L1 total RNA	-	control	C57B6	Lactation 1	mammary gland
18	C57B6 L3 1hr	-	control	C57B6	Lactation 3	mammary gland
19	C57B6 P13.1 poly(A)	-	control	C57B6	Pregnancy 13	mammary gland
20	C57B6 P13.2 poly(A)	-	control	C57B6	Pregnancy 13	mammary gland
21	C57B6 P13 total RNA	-	control	C57B6	Pregnancy 13	mammary gland
22	C57B6 virgin 3 hours	-	control	C57B6	Virgin 10 weeks old	mammary gland

6e. Sample Web links Array Reports

- Hyper-links to Web databases describing the hybridized samples popup Web browser



Samples Web Links

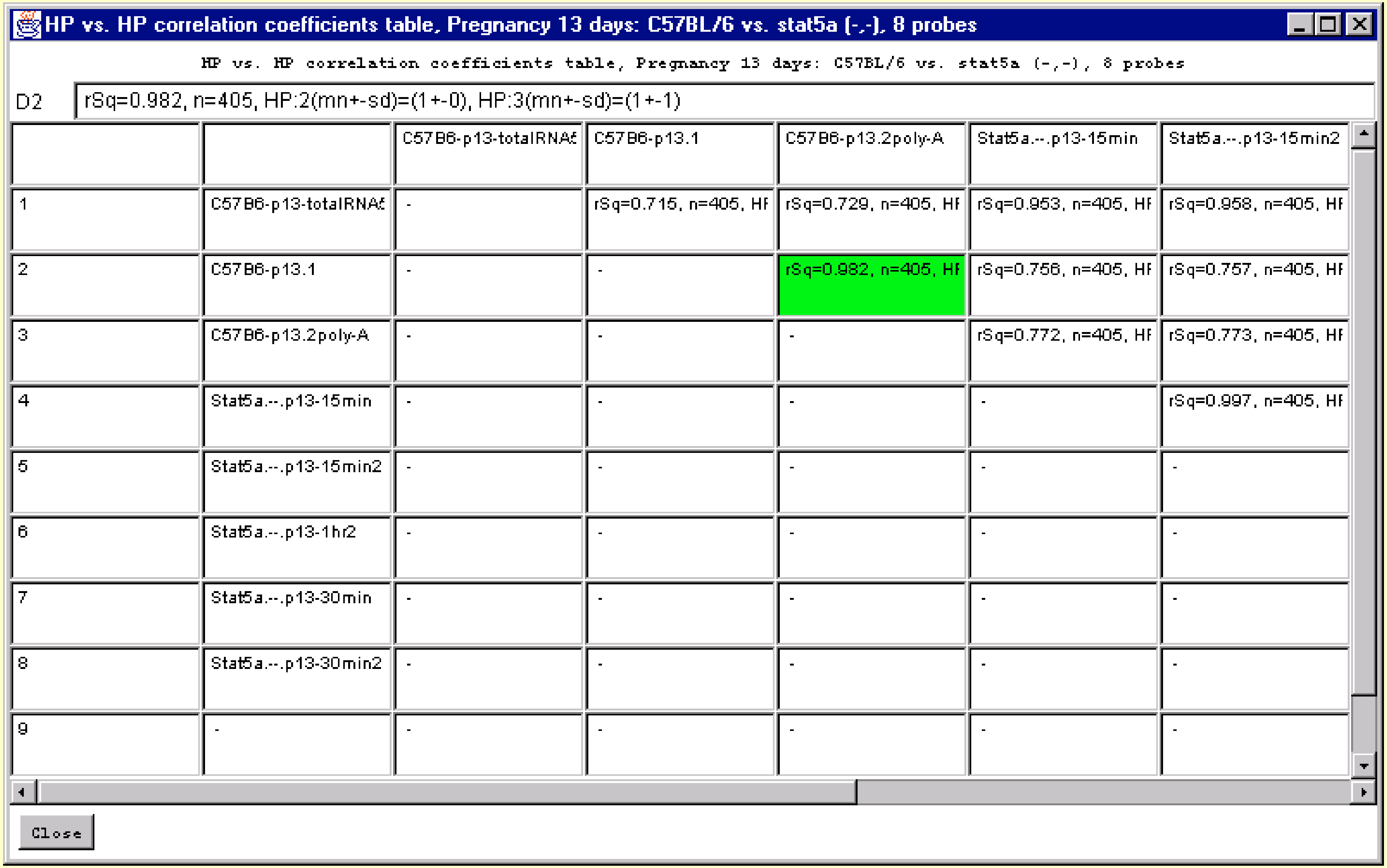
B5 http://bioinfo.weizmann.ac.il/cards-bin/carddisp?INHBB&search=Inhibin&suff=txt

	Sample_ID	GeneCard	Histology	Model
1	4A-1H	-	-	-
2	4C-5H	-	-	-
3	4B-1H	-	-	-
4	8A-1H	-	-	-
5	BetaB Hemo	http://bioinfo.weizmann.ac.il/ca	-	http://mammary.nih.gov/model
6	BetaB Null 1hr	http://bioinfo.weizmann.ac.il/ca	-	http://mammary.nih.gov/model
7	C57B6 I4 25hrs	-	-	-
8	C57B6 L10 29 hrs #1	-	-	-
9	C57B6 L10 29 hrs #2	-	-	-

☑ Web Access Enabled

6f. Sample Probe Correlation Reports

- Hybridized Probe vs Probe correlation coefficient reports for set of Filtered genes



HP vs. HP correlation coefficients table, Pregnancy 13 days: C57BL/6 vs. stat5a (-,-), 8 probes

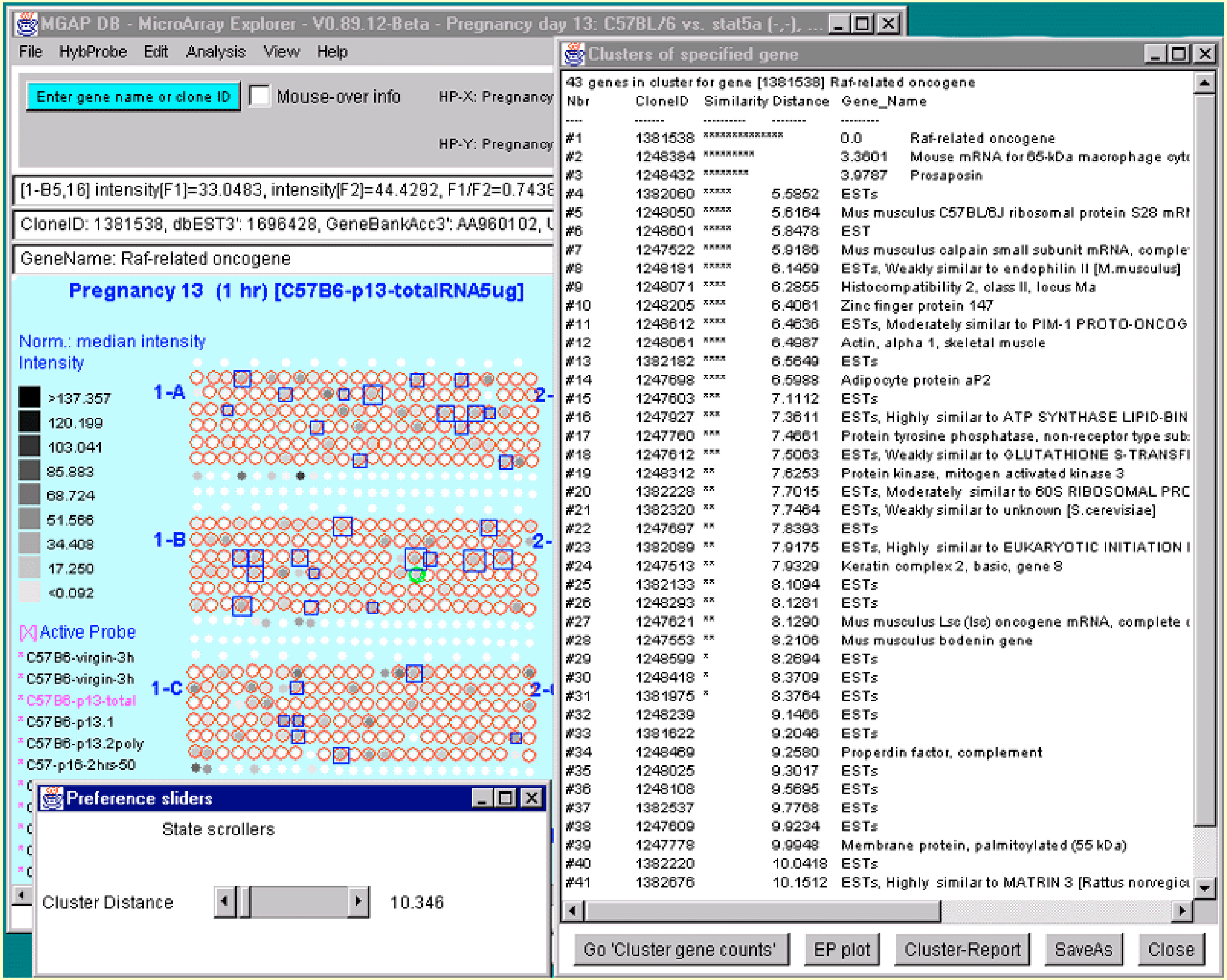
HP vs. HP correlation coefficients table, Pregnancy 13 days: C57BL/6 vs. stat5a (-,-), 8 probes

D2 rSq=0.982, n=405, HP:2(mn+-sd)=(1+-0), HP:3(mn+-sd)=(1+-1)

		C57B6-p13-totalRNA	C57B6-p13.1	C57B6-p13.2poly-A	Stat5a.--.p13-15min	Stat5a.--.p13-15min2
1	C57B6-p13-totalRNA	-	rSq=0.715, n=405, H	rSq=0.729, n=405, H	rSq=0.953, n=405, H	rSq=0.958, n=405, H
2	C57B6-p13.1	-	-	rSq=0.982, n=405, H	rSq=0.756, n=405, H	rSq=0.757, n=405, H
3	C57B6-p13.2poly-A	-	-	-	rSq=0.772, n=405, H	rSq=0.773, n=405, H
4	Stat5a.--.p13-15min	-	-	-	-	rSq=0.997, n=405, H
5	Stat5a.--.p13-15min2	-	-	-	-	-
6	Stat5a.--.p13-1hr2	-	-	-	-	-
7	Stat5a.--.p13-30min	-	-	-	-	-
8	Stat5a.--.p13-30min2	-	-	-	-	-
9	-	-	-	-	-	-

7a. Finding Genes With Similar Expression

- Genes that clustered to Raf-related oncogene with similar expression patterns



[1-B5,16] intensity[F1]=33.0483, intensity[F2]=44.4292, F1/F2=0.7438

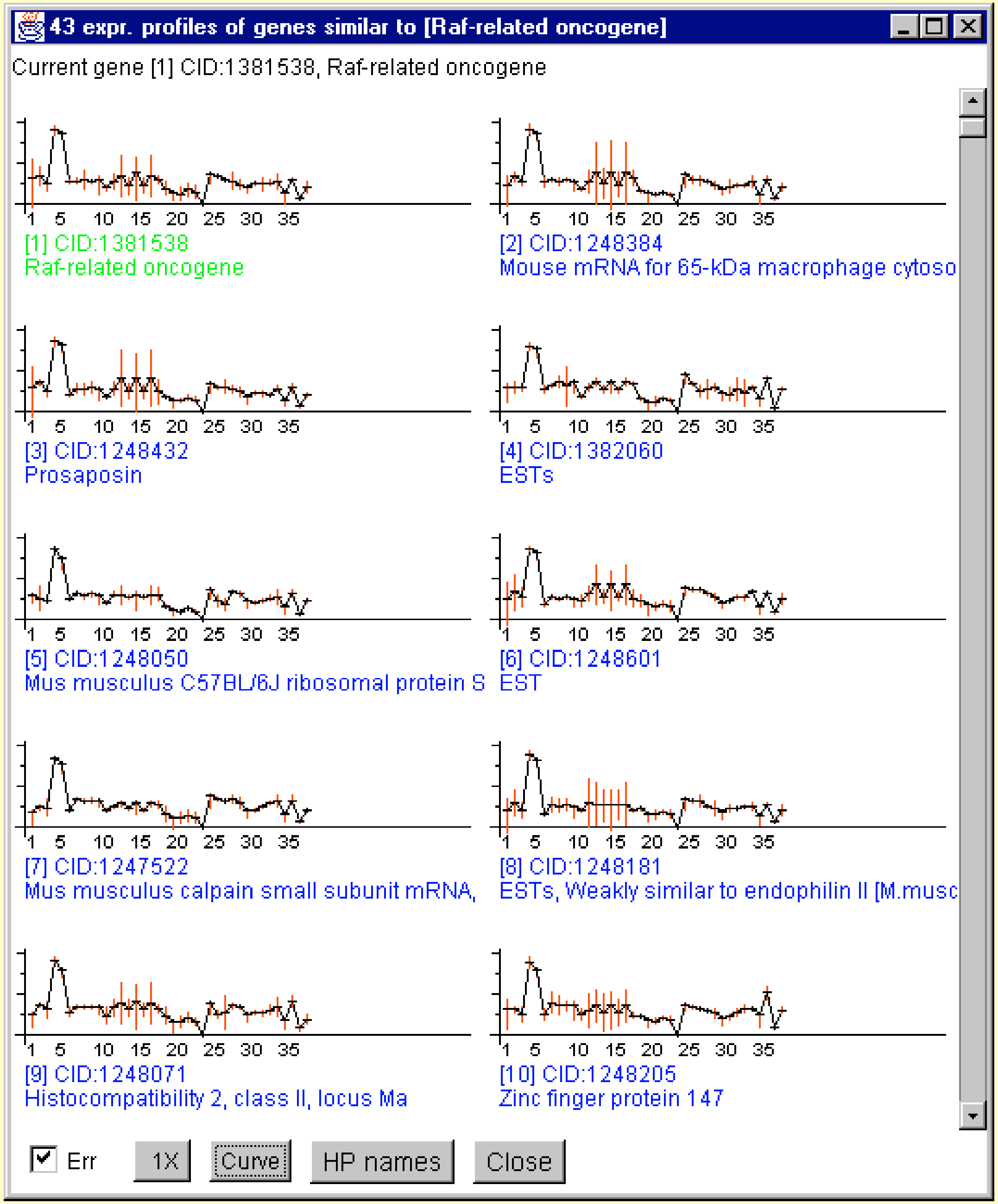
CloneID: 1381538, dbEST3': 1696428, GeneBankAcc3': AA960102, U

GeneName: Raf-related oncogene

```
43 genes in cluster for gene [1381538] Raf-related oncogene
Nbr    CloneID   Similarity Distance   Gene_Name
----   -------   ----------  --------  ---------
#1     1381538   **************        0.0       Raf-related oncogene
#2     1248384   **********            3.3601    Mouse mRNA for 65-kDa macrophage cyt
#3     1248432   *********             3.9787    Prosaposin
#4     1382060   *****       5.5852    ESTs
#5     1248050   *****       5.6164    Mus musculus C57BL/6J ribosomal protein S28 mR
#6     1248601   *****       5.8478    EST
#7     1247522   *****       5.9186    Mus musculus calpain small subunit mRNA, comple
#8     1248181   *****       6.1459    ESTs, Weakly similar to endophilin II [M.musculus]
#9     1248071   ****        6.2855    Histocompatibility 2, class II, locus Ma
#10    1248205   ****        6.4061    Zinc finger protein 147
#11    1248612   ****        6.4636    ESTs, Moderately similar to PIM-1 PROTO-ONCOG
#12    1248061   ****        6.4987    Actin, alpha 1, skeletal muscle
#13    1382182   ****        6.5649    ESTs
#14    1247698   ****        6.5988    Adipocyte protein aP2
#15    1247603   ***         7.1112    ESTs
#16    1247927   ***         7.3611    ESTs, Highly similar to ATP SYNTHASE LIPID-BIN
#17    1247760   ***         7.4661    Protein tyrosine phosphatase, non-receptor type sub
#18    1247612   ***         7.5063    ESTs, Weakly similar to GLUTATHIONE S-TRANSFI
#19    1248312   **          7.6253    Protein kinase, mitogen activated kinase 3
#20    1382228   **          7.7015    ESTs, Moderately similar to 60S RIBOSOMAL PRC
#21    1382320   **          7.7464    ESTs, Weakly similar to unknown [S.cerevisiae]
#22    1247697   **          7.8393    ESTs
#23    1382089   **          7.9175    ESTs, Highly similar to EUKARYOTIC INITIATION I
#24    1247513   **          7.9329    Keratin complex 2, basic, gene 8
#25    1382133   **          8.1094    ESTs
#26    1248293   **          8.1281    ESTs
#27    1247621   **          8.1290    Mus musculus Lsc (lsc) oncogene mRNA, complete
#28    1247553   **          8.2106    Mus musculus bodenin gene
#29    1248599   *           8.2694    ESTs
#30    1248418   *           8.3709    ESTs
#31    1381975   *           8.3764    ESTs
#32    1248239               9.1466    ESTs
#33    1381622               9.2046    ESTs
#34    1248469               9.2580    Properdin factor, complement
#35    1248025               9.3017    ESTs
#36    1248108               9.5695    ESTs
#37    1382537               9.7768    ESTs
#38    1247609               9.9234    ESTs
#39    1247778               9.9948    Membrane protein, palmitoylated (55 kDa)
#40    1382220               10.0418   ESTs
#41    1382676               10.1512   ESTs, Highly similar to MATRIN 3 [Rattus norvegicu
```

7b. EP Plots for Similar Genes

- Sorted list of EP plots of similar genes that clustered to Raf-related oncogene



43 expr. profiles of genes similar to [Raf-related oncogene]

Current gene [1] CID:1381538, Raf-related oncogene

[1] CID:1381538
Raf-related oncogene

[2] CID:1248384
Mouse mRNA for 65-kDa macrophage cytoso

[3] CID:1248432
Prosaposin

[4] CID:1382060
ESTs

[5] CID:1248050
Mus musculus C57BL/6J ribosomal protein S

[6] CID:1248601
EST

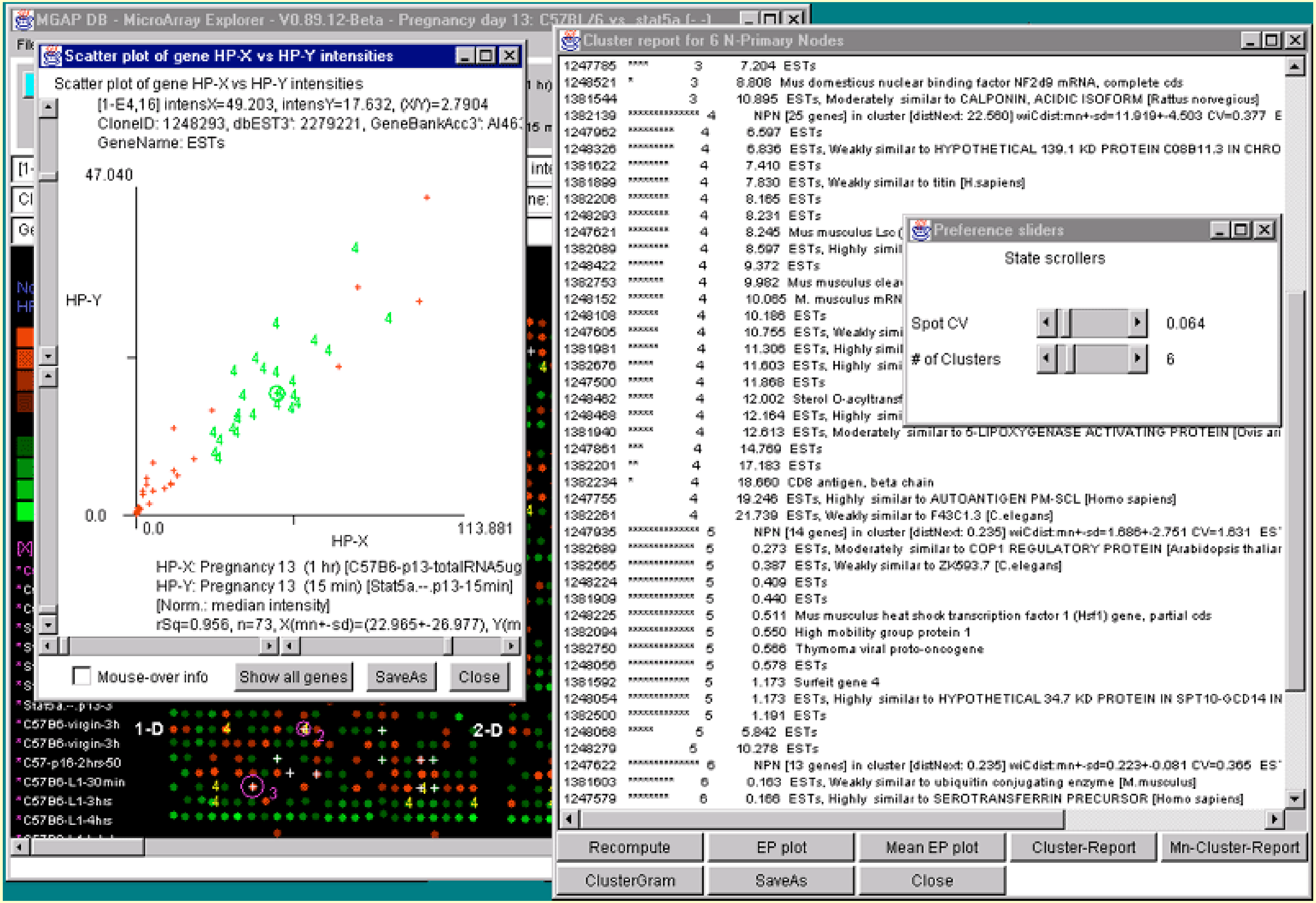
[7] CID:1247522
Mus musculus calpain small subunit mRNA,

[8] CID:1248181
ESTs, Weakly similar to endophilin II [M.musc

[9] CID:1248071
Histocompatibility 2, class II, locus Ma

[10] CID:1248205
Zinc finger protein 147

7c. Finding N-Clusters of Genes with Similar Expression Patterns (similar to K-means)



Scatter plot of gene HP-X vs HP-Y intensities

[1-E4,16] intensX=49.203, intensY=17.632, (X/Y)=2.7904
CloneID: 1248293, dbEST3': 2279221, GeneBankAcc3': AI46
GeneName: ESTs

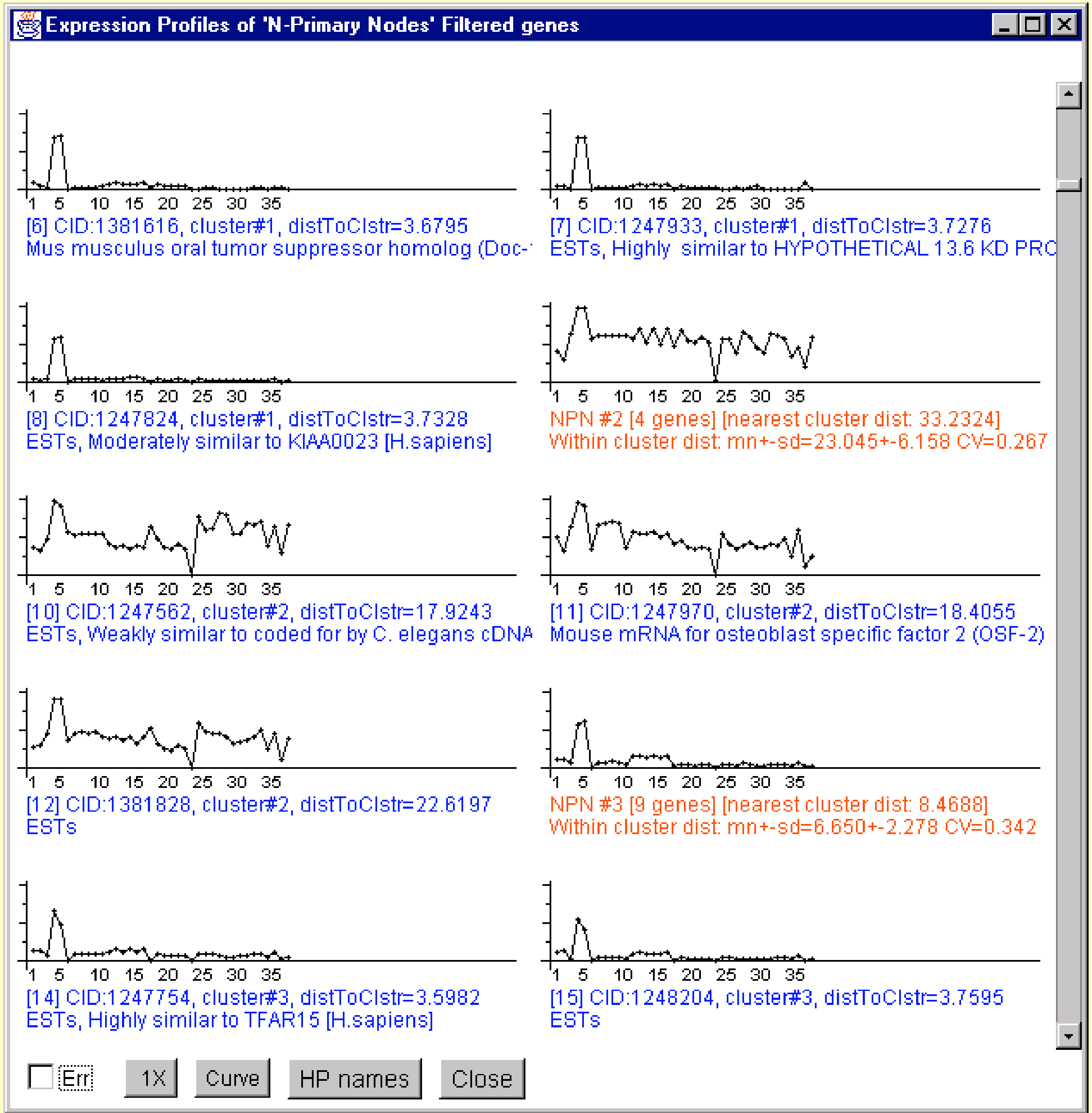
HP-X: Pregnancy 13 (1 hr) [C57B6-p13-totalRNA5ug
HP-Y: Pregnancy 13 (15 min) [Stat5a.--.p13-15min]
[Norm.: median intensity]
rSq=0.956, n=73, X(mn+-sd)=(22.965+-26.977), Y(m

Spot CV 0.064
# of Clusters 6

```
1247785 ****           3      7.204  ESTs
1248521 *              3      8.808  Mus domesticus nuclear binding factor NF2d9 mRNA, complete cds
1381544                3      10.895  ESTs, Moderately  similar to CALPONIN, ACIDIC ISOFORM [Rattus norvegicus]
1382139 ************** 4        NPN [25 genes] in cluster [distNext: 22.560] wiCdist:mn+-sd=11.919+-4.503 CV=0.377  E
1247962 **********     4       6.597  ESTs
1248326 **********     4       6.836  ESTs, Weakly similar to HYPOTHETICAL 139.1 KD PROTEIN C08B11.3 IN CHRO
1381622 ********       4       7.410  ESTs
1381699 ********       4       7.630  ESTs, Weakly similar to titin [H.sapiens]
1382206 ********       4       8.165  ESTs
1248293 ********       4       8.231  ESTs
1247621 ********       4       8.245  Mus musculus Lsc (
1382089 ********       4       8.597  ESTs, Highly  simil
1248422 *******        4       9.372  ESTs
1382753 *******        4       9.982  Mus musculus clea
1248152 *******        4       10.065  M. musculus mRN
1248108 ******         4       10.186  ESTs
1247605 ******         4       10.755  ESTs, Weakly simi
1381981 ******         4       11.306  ESTs, Highly simil
1382676 *****          4       11.603  ESTs, Highly  simi
1247500 *****          4       11.868  ESTs
1248462 *****          4       12.002  Sterol O-acyltransf
1248468 *****          4       12.164  ESTs, Highly  simi
1381940 *****          4       12.613  ESTs, Moderately  similar to 5-LIPOXYGENASE ACTIVATING PROTEIN [Ovis ari
1247861 ***            4      14.769  ESTs
1382201 **             4      17.183  ESTs
1382234 *              4      18.660  CD8 antigen, beta chain
1247755                4      19.246  ESTs, Highly  similar to AUTOANTIGEN PM-SCL [Homo sapiens]
1382261                4      21.739  ESTs, Weakly similar to F43C1.3 [C.elegans]
1247935 ************** 5        NPN [14 genes] in cluster [distNext: 0.235] wiCdist:mn+-sd=1.686+-2.761 CV=1.631  ES
1382689 *************  5        0.273  ESTs, Moderately  similar to COP1 REGULATORY PROTEIN [Arabidopsis thaliar
1382565 *************  5        0.387  ESTs, Weakly similar to ZK593.7 [C.elegans]
1248224 *************  5        0.409  ESTs
1381909 *************  5        0.440  ESTs
1248225 *************  5        0.511  Mus musculus heat shock transcription factor 1 (Hsf1) gene, partial cds
1382094 *************  5        0.550  High mobility group protein 1
1382750 *************  5        0.566  Thymoma viral proto-oncogene
1248056 *************  5        0.578  ESTs
1381592 ************   5        1.173  Surfeit gene 4
1248054 ************   5        1.173  ESTs, Highly  similar to HYPOTHETICAL 34.7 KD PROTEIN IN SPT10-GCD14 IN
1382500 ************   5        1.191  ESTs
1248068 *****          5       5.842  ESTs
1248279                5      10.278  ESTs
1247622 ************** 6        NPN [13 genes] in cluster [distNext: 0.235] wiCdist:mn+-sd=0.223+-0.081 CV=0.365  ES
1381603 **********     6       0.163  ESTs, Weakly similar to ubiquitin conjugating enzyme [M.musculus]
1247579 ********       6       0.166  ESTs, Highly  similar to SEROTRANSFERRIN PRECURSOR [Homo sapiens]
```

7.d Expression Profiles of Clusters

- Scrollable list of EP plots showing genes from clusters #1, #2, #3 (from figure II.12c)



Expression Profiles of 'N-Primary Nodes' Filtered genes

[6] CID:1381616, cluster#1, distToClstr=3.6795
Mus musculus oral tumor suppressor homolog (Doc-

[7] CID:1247933, cluster#1, distToClstr=3.7276
ESTs, Highly similar to HYPOTHETICAL 13.6 KD PR

[8] CID:1247824, cluster#1, distToClstr=3.7328
ESTs, Moderately similar to KIAA0023 [H.sapiens]

NPN #2 [4 genes] [nearest cluster dist: 33.2324]
Within cluster dist: mn+-sd=23.045+-6.158 CV=0.267

[10] CID:1247562, cluster#2, distToClstr=17.9243
ESTs, Weakly similar to coded for by C. elegans cDNA

[11] CID:1247970, cluster#2, distToClstr=18.4055
Mouse mRNA for osteoblast specific factor 2 (OSF-2)

[12] CID:1381828, cluster#2, distToClstr=22.6197
ESTs

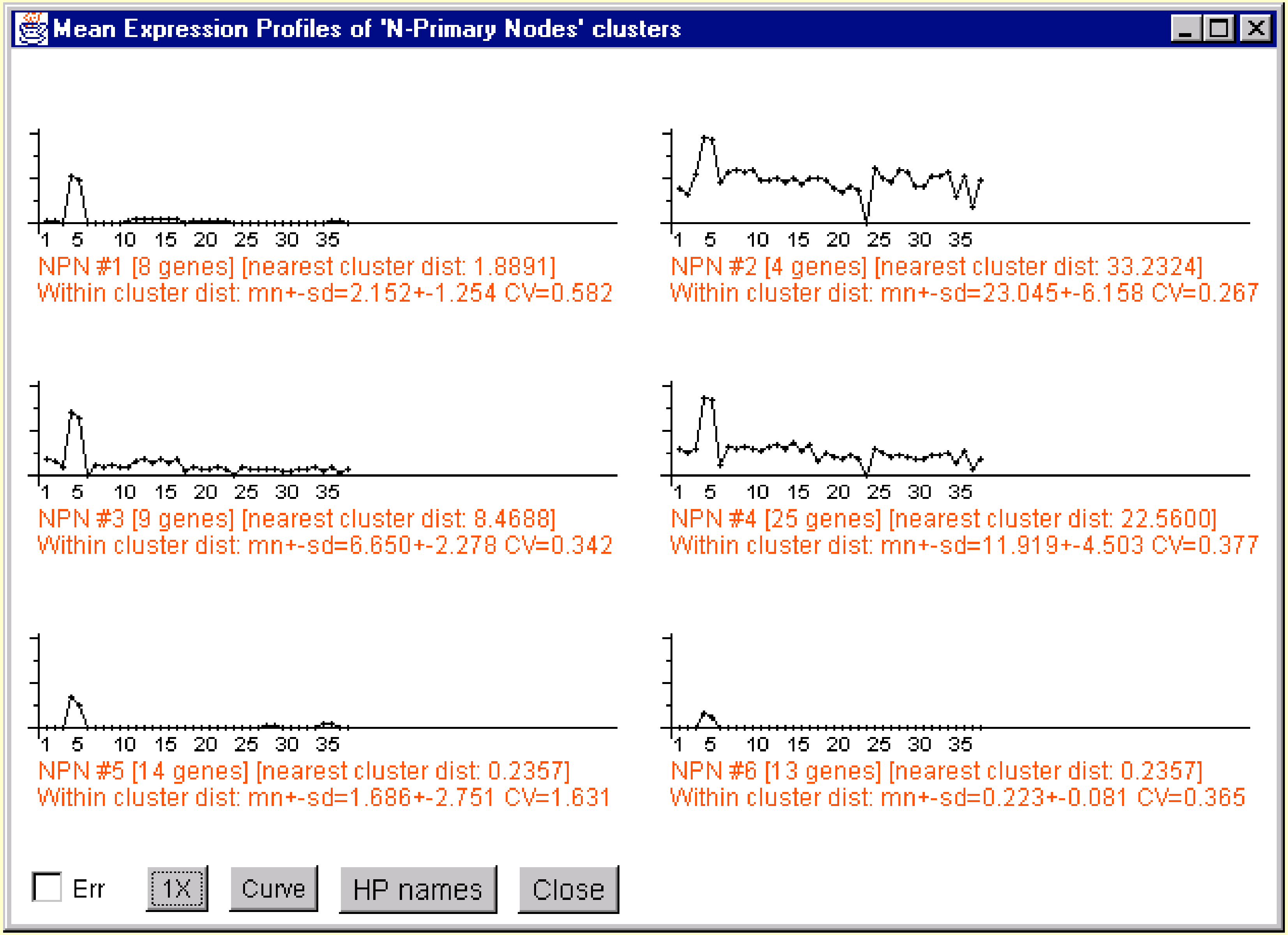
NPN #3 [9 genes] [nearest cluster dist: 8.4688]
Within cluster dist: mn+-sd=6.650+-2.278 CV=0.342

[14] CID:1247754, cluster#3, distToClstr=3.5982
ESTs, Highly similar to TFAR15 [H.sapiens]

[15] CID:1248204, cluster#3, distToClstr=3.7595
ESTs

7.e Mean Expression Profile Plots of Clusters

- Mean clusters and their statistics (from figure II.12c). Error bars are standard-deviation of genes’ intensities in each cluster



NPN #1 [8 genes] [nearest cluster dist: 1.8891]
Within cluster dist: mn+-sd=2.152+-1.254 CV=0.582

NPN #2 [4 genes] [nearest cluster dist: 33.2324]
Within cluster dist: mn+-sd=23.045+-6.158 CV=0.267

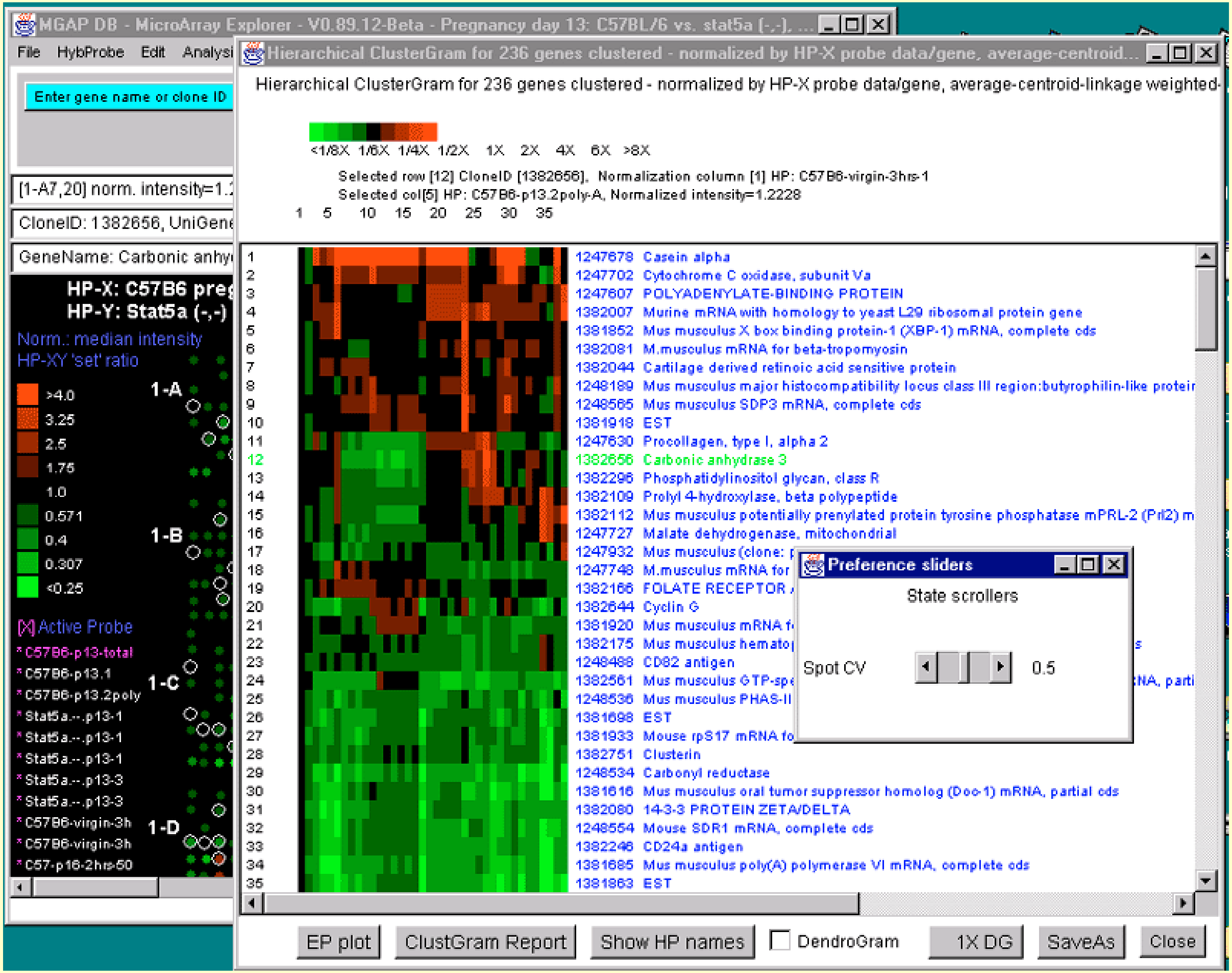
NPN #3 [9 genes] [nearest cluster dist: 8.4688]
Within cluster dist: mn+-sd=6.650+-2.278 CV=0.342

NPN #4 [25 genes] [nearest cluster dist: 22.5600]
Within cluster dist: mn+-sd=11.919+-4.503 CV=0.377

NPN #5 [14 genes] [nearest cluster dist: 0.2357]
Within cluster dist: mn+-sd=1.686+-2.751 CV=1.631

NPN #6 [13 genes] [nearest cluster dist: 0.2357]
Within cluster dist: mn+-sd=0.223+-0.081 CV=0.365

7.f Hierarchical Clustering ClusterGrams of Expression Profiles



[1-A7,20] norm. intensity=1.

CloneID: 1382656, UniGene

GeneName: Carbonic anhy

HP-X: C57B6 preg
HP-Y: Stat5a (-,-)

Norm.: median intensity
HP-XY 'set' ratio

🟥	>4.0
▨	3.25
🟥	2.5
🟥	1.75
	1.0
🟩	0.571
🟩	0.4
🟩	0.307
🟩	<0.25

[X] Active Probe
- * C57B6-p13-total
- * C57B6-p13.1
- * C57B6-p13.2poly
- * Stat5a.--.p13-1
- * Stat5a.--.p13-1
- * Stat5a.--.p13-1
- * Stat5a.--.p13-3
- * Stat5a.--.p13-3
- * C57B6-virgin-3h
- * C57B6-virgin-3h
- * C57-p16-2hrs-50

Hierarchical ClusterGram for 236 genes clustered - normalized by HP-X probe data/gene, average-centroid-linkage weighted-

Selected row [12] CloneID [1382656], Normalization column [1] HP: C57B6-virgin-3hrs-1
Selected col[5] HP: C57B6-p13.2poly-A, Normalized intensity=1.2228

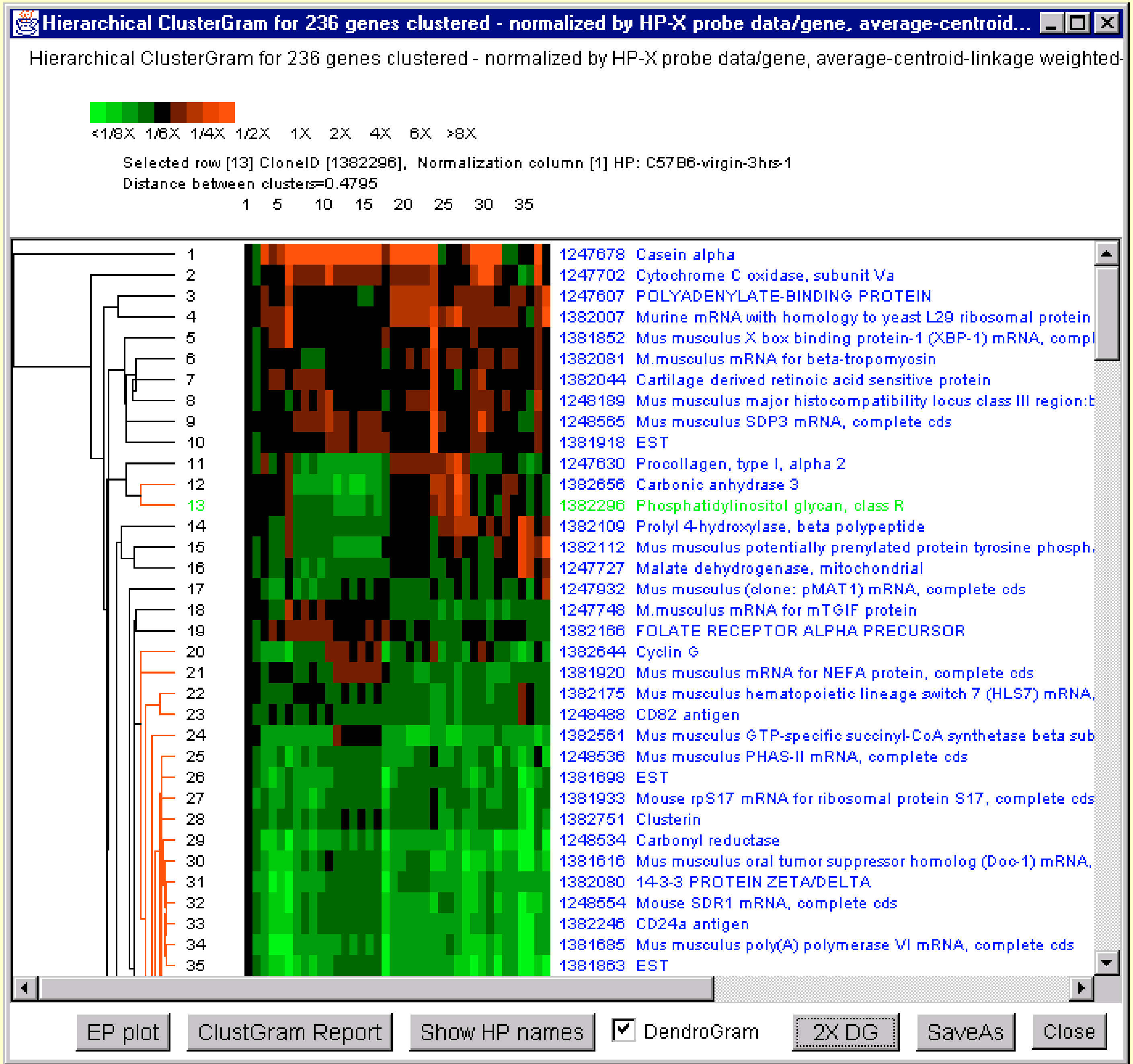
```
1  1247678  Casein alpha
2  1247702  Cytochrome C oxidase, subunit Va
3  1247607  POLYADENYLATE-BINDING PROTEIN
4  1382007  Murine mRNA with homology to yeast L29 ribosomal protein gene
5  1381852  Mus musculus X box binding protein-1 (XBP-1) mRNA, complete cds
6  1382081  M.musculus mRNA for beta-tropomyosin
7  1382044  Cartilage derived retinoic acid sensitive protein
8  1248189  Mus musculus major histocompatibility locus class III region:butyrophilin-like protei
9  1248565  Mus musculus SDP3 mRNA, complete cds
10 1381918  EST
11 1247630  Procollagen, type I, alpha 2
12 1382656  Carbonic anhydrase 3
13 1382296  Phosphatidylinositol glycan, class R
14 1382109  Prolyl 4-hydroxylase, beta polypeptide
15 1382112  Mus musculus potentially prenylated protein tyrosine phosphatase mPRL-2 (Prl2) m
16 1247727  Malate dehydrogenase, mitochondrial
17 1247932  Mus musculus (clone: 
18 1247748  M.musculus mRNA for
19 1382166  FOLATE RECEPTOR 
20 1382644  Cyclin G
21 1381920  Mus musculus mRNA f
22 1382175  Mus musculus hemato                                s
23 1248488  CD82 antigen
24 1382561  Mus musculus GTP-spe                               NA, parti
25 1248536  Mus musculus PHAS-II
26 1381698  EST
27 1381933  Mouse rpS17 mRNA fo
28 1382751  Clusterin
29 1248534  Carbonyl reductase
30 1381616  Mus musculus oral tumor suppressor homolog (Doc-1) mRNA, partial cds
31 1382080  14-3-3 PROTEIN ZETA/DELTA
32 1248554  Mouse SDR1 mRNA, complete cds
33 1382246  CD24a antigen
34 1381685  Mus musculus poly(A) polymerase VI mRNA, complete cds
35 1381863  EST
```

State scrollers

Spot CV 0.5

7.g Hierarchical Clustering Dendrogram

- Clusters less than cluster distance from each other are shown in red (from figure II.12f)



Hierarchical ClusterGram for 236 genes clustered - normalized by HP-X probe data/gene, average-centroid...

Hierarchical ClusterGram for 236 genes clustered - normalized by HP-X probe data/gene, average-centroid-linkage weighted-

Selected row [13] CloneID [1382296], Normalization column [1] HP: C57B6-virgin-3hrs-1
Distance between clusters=0.4795

1	1247678	Casein alpha
2	1247702	Cytochrome C oxidase, subunit Va
3	1247607	POLYADENYLATE-BINDING PROTEIN
4	1382007	Murine mRNA with homology to yeast L29 ribosomal protein
5	1381852	Mus musculus X box binding protein-1 (XBP-1) mRNA, compl
6	1382081	M.musculus mRNA for beta-tropomyosin
7	1382044	Cartilage derived retinoic acid sensitive protein
8	1248189	Mus musculus major histocompatibility locus class III region:b
9	1248565	Mus musculus SDP3 mRNA, complete cds
10	1381918	EST
11	1247630	Procollagen, type I, alpha 2
12	1382656	Carbonic anhydrase 3
13	1382296	Phosphatidylinositol glycan, class R
14	1382109	Prolyl 4-hydroxylase, beta polypeptide
15	1382112	Mus musculus potentially prenylated protein tyrosine phosph.
16	1247727	Malate dehydrogenase, mitochondrial
17	1247932	Mus musculus (clone: pMAT1) mRNA, complete cds
18	1247748	M.musculus mRNA for mTGIF protein
19	1382166	FOLATE RECEPTOR ALPHA PRECURSOR
20	1382644	Cyclin G
21	1381920	Mus musculus mRNA for NEFA protein, complete cds
22	1382175	Mus musculus hematopoietic lineage switch 7 (HLS7) mRNA,
23	1248488	CD82 antigen
24	1382561	Mus musculus GTP-specific succinyl-CoA synthetase beta sub
25	1248536	Mus musculus PHAS-II mRNA, complete cds
26	1381698	EST
27	1381933	Mouse rpS17 mRNA for ribosomal protein S17, complete cds
28	1382751	Clusterin
29	1248534	Carbonyl reductase
30	1381616	Mus musculus oral tumor suppressor homolog (Doc-1) mRNA,
31	1382080	14-3-3 PROTEIN ZETA/DELTA
32	1248554	Mouse SDR1 mRNA, complete cds
33	1382246	CD24a antigen
34	1381685	Mus musculus poly(A) polymerase VI mRNA, complete cds
35	1381863	EST

Summary of MAExplorer

- MAExplorer is used as a stand-alone application or as applet over the Web
- Accepts different array geometries, spot supports, ^{33}P or Cy3/Cy5 labeling, scanners
- Analyzes multiple probes, X-Y replicate sets, expression profiles, replicate spots
- Provides direct manipulation of array pseudo images, scatter-plots, histograms, clustergrams, dendrograms, silhouette plots, spreadsheets
- Data filters genes by gene subsets, spot intensities and ratios, and statistical tests, etc.
- Set operations on gene subsets help manage search results
- Uses active Web links to genomic, histology and model Web databases
- Generates reports as Web-accessible spreadsheets or exportable to Excel
- Users may save their data-mining session state locally for later use or sharing
- Building tools to import commercial and academic quantified micro array data
- MAExplorer used to identify genes in MGAP DB preferentially expressed during lactation. Results verified using northern blots (NIDDK), *Nucleic Acids Res.* **28**:4452-4459 (2000).
- Online documentation (manual, tutorials, examples, etc.) is available on Web site

Some MAExplorer URL References

- **Home Page (includes the following and other links)**
  http://www.lecb.ncifcrf.gov/MAExplorer/

- **Reference Manual (including tutorials, and use with other arrays sections)**
  http://www.lecb.ncifcrf.gov/MAExplorer/hmaeHelp.html (online)
  http://www.lecb.ncifcrf.gov/MAExplorer/MaeRefMan.zip (download)

- **Overview of MAExplorer**
  http://www.lecb.ncifcrf.gov/MAExplorer/PDF/Overview-MAE.pdf

- **Introduction to data mining with MAExplorer**
  http://www.lecb.ncifcrf.gov/MAExplorer/IntroDataMiningWithMAExplorer.pdf

- **Using with mAdb with MAExplorer**
  http://www.lecb.ncifcrf.gov/MAExplorer/Using-mAdb-with-MAExplorer.pdf

- ***Nucleic Acids Res*. (2000) 28:4452 paper**
  http://www.lecb.ncifcrf.gov/MAExplorer/lemkin-NAR-2000-Vol28-pp4452.pdf

- **Download MAExplorer (includes 38 samples from MGAP DB)**
  http://www.lecb.ncifcrf.gov/MAExplorer/hmaeInstall.html