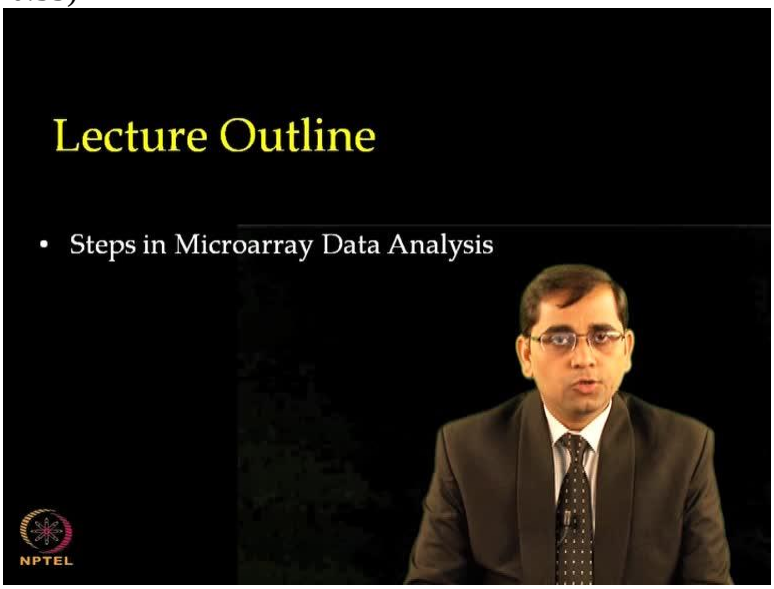
Interactomics: Protein Arrays & Label Free Biosensors
Professor Sanjeeva Srivastava
MOOC NPTEL Course
Indian Institute of Technology Bombay
Module 7
Lecture No 36
MicroArray Data Analysis: Part II

Welcome to the mooc interactomics course. In our previous lecture we discussed the acuity software interface and became familiar to the control parameters as well as data analysis options.

(Refer Slide Time: 0:35)



In today's lecture we will talk about data analysis and delve more into the details of all the theoretical aspects and the constrains that we have discussed over the last few lectures. It gives me a must pleasure to invite me Mr. Pankaj from Spinco biotech again who will now talk to us about data analysis interface on acuity software.

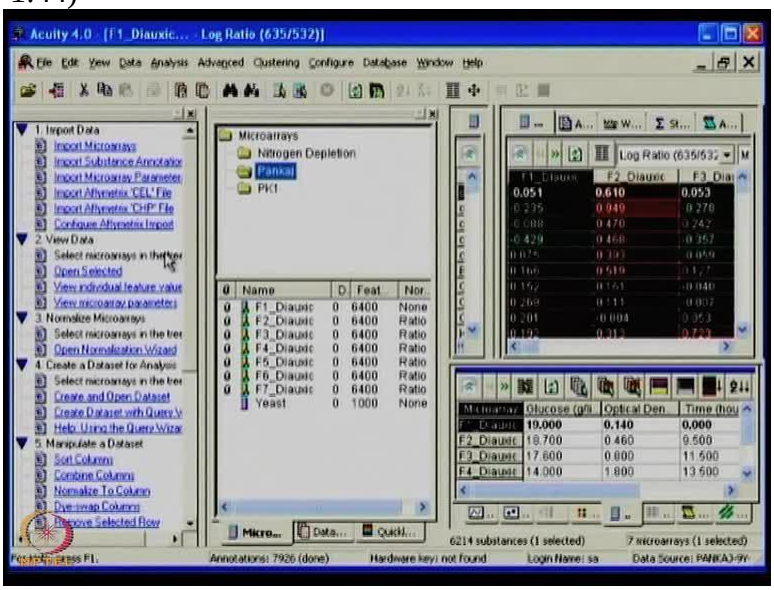
Prof Sanjeeva Srivastava: So, Pankaj in the last lecture when we talked about GenePix pro uhh then you showed one yeast slide how to scan that yeast slide uhh by using the GenePix pro software uhh I guess now its will be good if you can use same slide what we scanned in the last lecture and see how we can analyze that here.

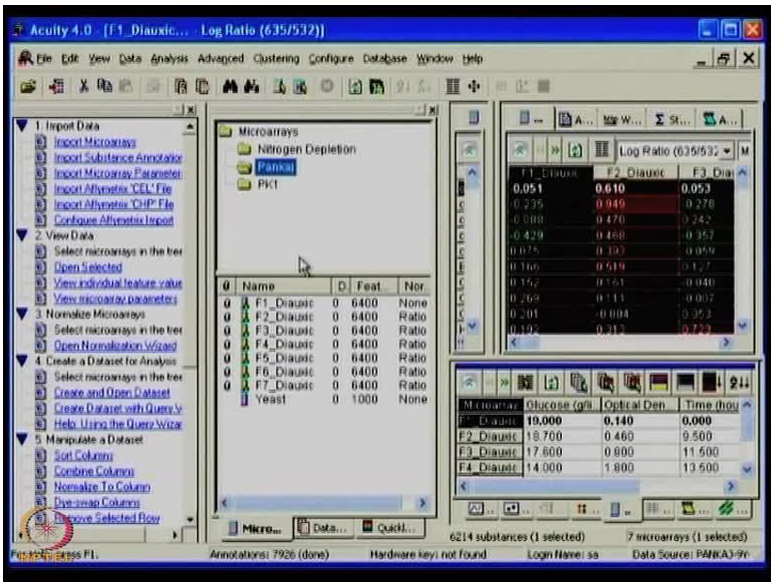
So, this slide actually was used for looking at the glucose response in various time point from 0 to 20.5 hours in yeast and it would be interesting to see what type of trends we observe in various time point with the glucose utilization here.

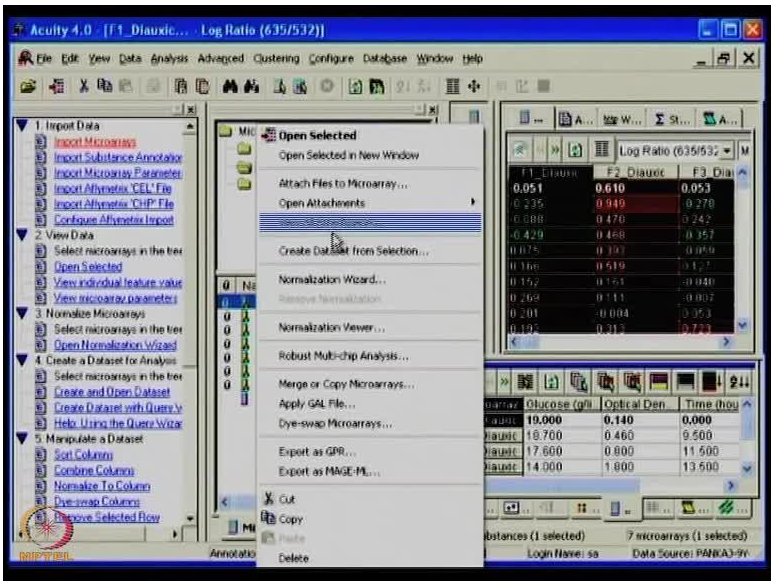
Mr Pankaj Khanna: Sure.

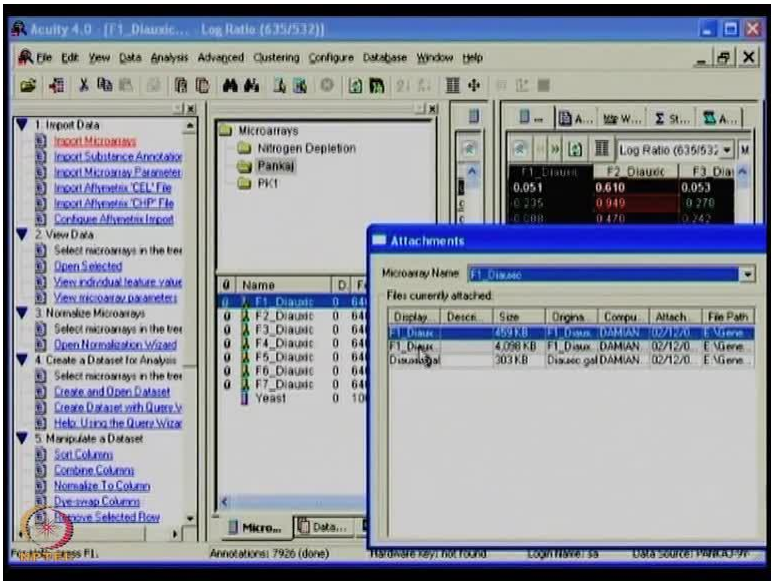
Prof Sanjeeva Srivastava: So, please use that slide and we can look the demo here.

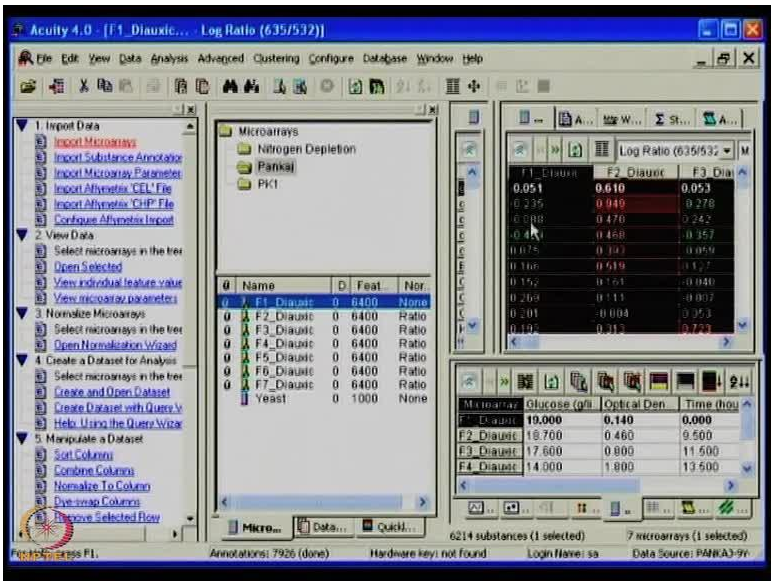
(Refer Slide Time: 1:44)

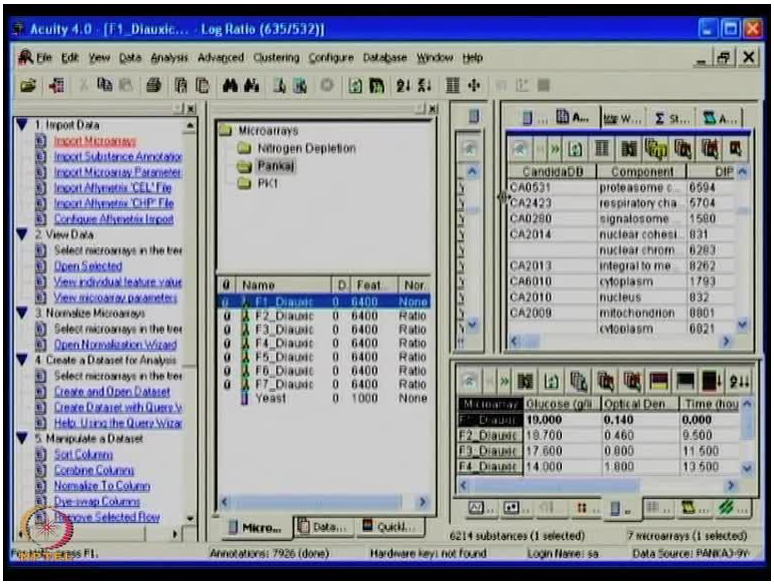


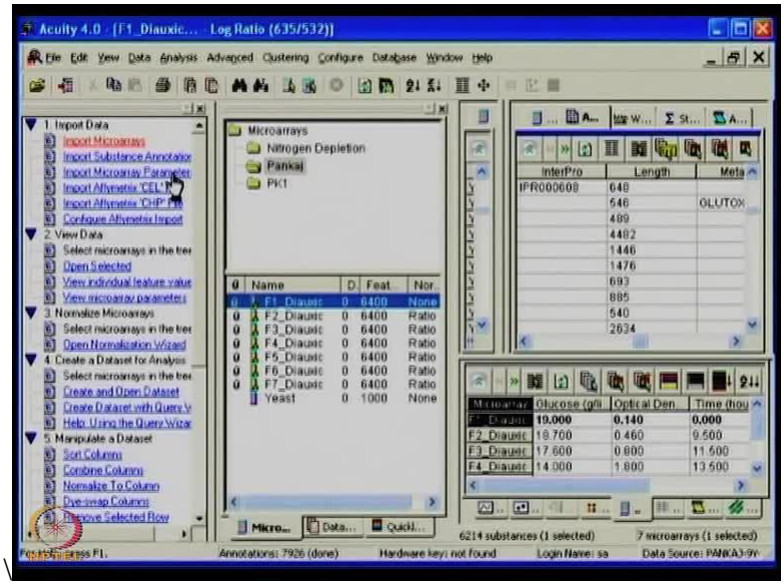


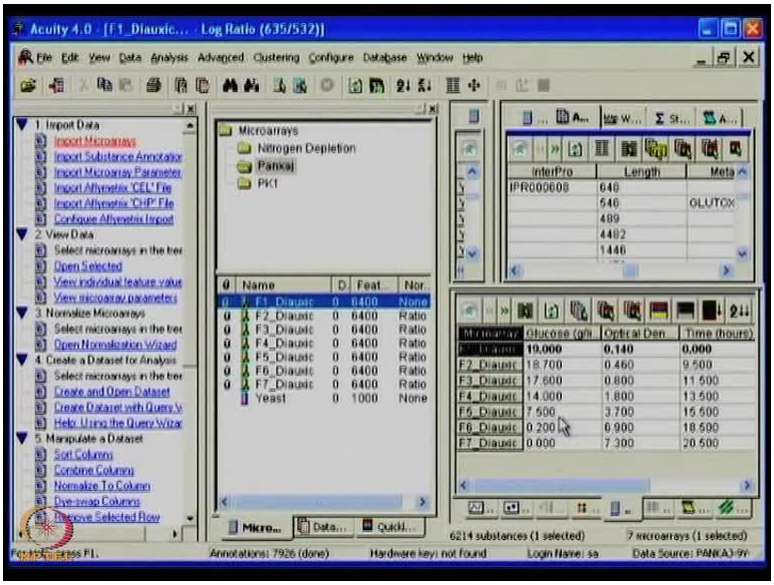

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameters
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Open Selected
Open Selected in New Window
Attach Files to Microarray...
Open Attachments
Create Dataset from Selection...
Normalization Wizard...
Normalization Viewer...
Robust Multi-chip Analysis...
Merge or Copy Microarrays...
Apply GAL File...
Dye-swap Microarrays...
Export as GPR...
Export as MAGE-ML...
Cut
Copy
Paste
Delete
Log Ratio (635/532
F1_Diauxic
F2_Diauxic
F3_Diau
0.051
0.610
0.053
Glucose (g/l
Optical Den
Time (hou
19.000
0.140
0.000
18.700
0.460
9.500
17.600
0.800
11.500
14.000
1.800
13.500
7 microarrays (1 selected)
Login Name: sa
Data Source: PANKAJ-9F

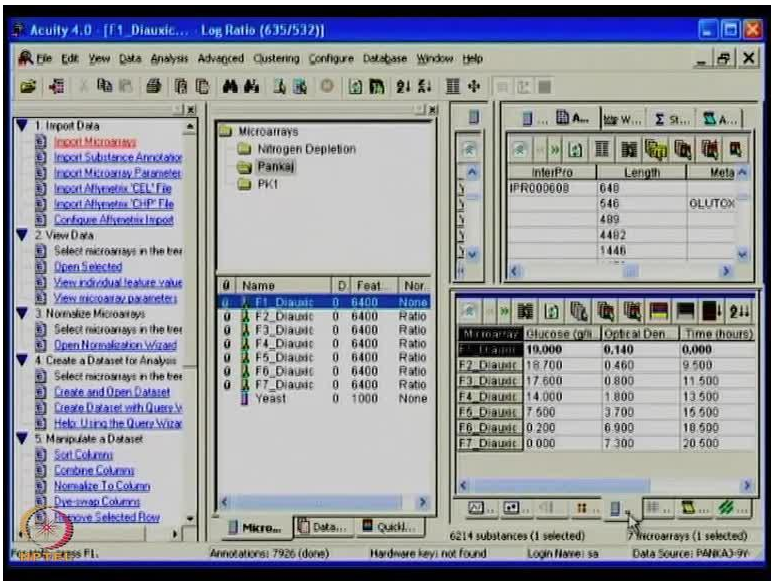

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameters
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name
F1_Diauxic
F2_Diauxic
F3_Diauxic
F4_Diauxic
F5_Diauxic
F6_Diauxic
F7_Diauxic
Yeast
Micro...
Data...
Attachments
Microarray Name: F1_Diauxic
Files currently attached:
Display
Descri...
Size
Origins
Compu...
Attach...
File Path
F1_Diaux
4,098 KB
F1_Diaux
DAMIAN
02/12/0...
E:\Gene...
Diauxic.gal
303 KB
Diauxic.gal
DAMIAN
02/12/0...
E:\Gene...
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-9F


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizar
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor...
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro...
Data...
Quick...
Log Ratio (635/532)
F1_Diauxic F2_Diauxic F3_Diauxic
0.051 0.610 0.053
Microarray Glucose (g/l Optical Den... Time (hou...
F1_Diauxic 19.000 0.140 0.000
F2_Diauxic 18.700 0.460 9.500
F3_Diauxic 17.600 0.800 11.500
F4_Diauxic 14.000 1.800 13.500
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizar
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor...
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro...
Data...
Quick...
CandidaDB Component DIP
CA0531 proteasome c... 6594
CA2423 respiratory cha... 5704
CA0280 signalosome 1580
CA2014 nuclear cohesi... 831
nuclear chrom... 6283
CA2013 integral to me... 8262
CA6010 cytoplasm 1793
CA2010 nucleus 832
CA2009 mitochondrion 8801
cytoplasm 6021
Microarray Glucose (g/l Optical Den... Time (hou...
F1_Diauxic 19.000 0.140 0.000
F2_Diauxic 18.700 0.460 9.500
F3_Diauxic 17.600 0.800 11.500
F4_Diauxic 14.000 1.800 13.500
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1 Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2 View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3 Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4 Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5 Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
InterPro Length Meta
IPR000608 648
546 GLUTOX
489
4402
1446
1476
693
885
540
2634
Glucose (g/l Optical Den Time (hou
F1_Diauxic 19.000 0.140 0.000
F2_Diauxic 18.700 0.460 9.500
F3_Diauxic 17.600 0.800 11.500
F4_Diauxic 14.000 1.800 13.500
Micro... Data... Quick...
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-9Y


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1 Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2 View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3 Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4 Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5 Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
InterPro Length Meta
IPR000608 648
546 GLUTOX
489
4402
1446
Glucose (g/l Optical Den Time (hours)
F1_Diauxic 19.000 0.140 0.000
F2_Diauxic 18.700 0.460 9.500
F3_Diauxic 17.600 0.800 11.500
F4_Diauxic 14.000 1.800 13.500
F5_Diauxic 7.500 3.700 15.500
F6_Diauxic 0.200 6.900 18.500
F7_Diauxic 0.000 7.300 20.500
Micro... Data... Quick...
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-9Y



Mr Pankaj Khanna: Let us walk through the data. So as we describe so we see a first import data tags. So basically this allows one to take the data from the microarray database and store in the form of GPR file and allow that to be understand by the acuity software. We also have opportunity in the form of text file input were you will define what is available for what. So in this fashion I here I have defined in the micro-arrays the folder where it says about the training and it says what all different slides are available to us to design.

So, here we see seven different kinds of time points collected and these are individual slides which one has run with board si3 and si5. So, this is ratio based image which we are going to see in and we have also got an yeast step another file which has come from the text to show you even that can also be imported.

Prof Sanjeeva Srivastava: So, first you have to individually scan each of the images, make one folder where you group all of this and then used that whole data set for the combined analysis.

Mr Pankaj Khanna: Right. So you can have all the GPR file stored or one by one from the genepix pro and just import the data in the form of import microarray data file. So it a just goes on looking for the GPR file and this now can be imported and ready for your analysis.

Prof Sanjeeva Srivastava: okay

Mr Pankaj Khanna: So once the data is in this will be displayed in the down in the form of folder which you kept in for the analysis.

Prof Sanjeeva Srivastava: ya.

Mr Pankaj Khanna: So here the folders are as we discussed that this can be used for a data (())(3:21) so, this little bit kind of image here that tells you that there is some files are attached.

Prof Sanjeeva Srivastava: okay.

Mr Pankaj Khanna: You can view them, you can see them that what all somebody has attached. So if I see view all attachments I will be able to see one has attached to it the important files coming from the GPS in the form so also a GAL file and the image.

So it give me an complete opportunity look at. It is good to emphasise here if you have a ZIP file this particular one can also behave a partial visualization tool as in case of Genepix.

Prof Sanjeeva Srivastava: Okay

Mr Pankaj Khanna: In case you want to look how this spot has behaved. So in this fashion it it can be it any any file can be attached to it.

Prof Sanjeeva Srivastava: And since you have the GAL file you also know the gene array list so at any time point if you identify in a spot which is looking interesting, you know what that gene is…

Mr Pankaj Khanna: True.

Prof Sanjeeva Srivastava: …by aligning with the GAL map.

Mr Pankaj Khanna: Correct. And another is good point you raised because very important next step is a substance annotation. The substance here I mean is a each spot which could be the feature which is again could be a RNA or a gene or a protein. So that is why we called as a substance annotation, very few people extensively traced them.

So here I will show you in the form of tab data in the annotation one can look at what all different information can be seen for each tab wise. So you have the same annotation tab being given for the substance ID and then because it is each database of candidized being attached here component different function even at the level of enzyme commission numbers there are different annotations which people try to get in which also you can import in the form of text delimited, rename to dot std file which allows one to take all the annotation information possible.

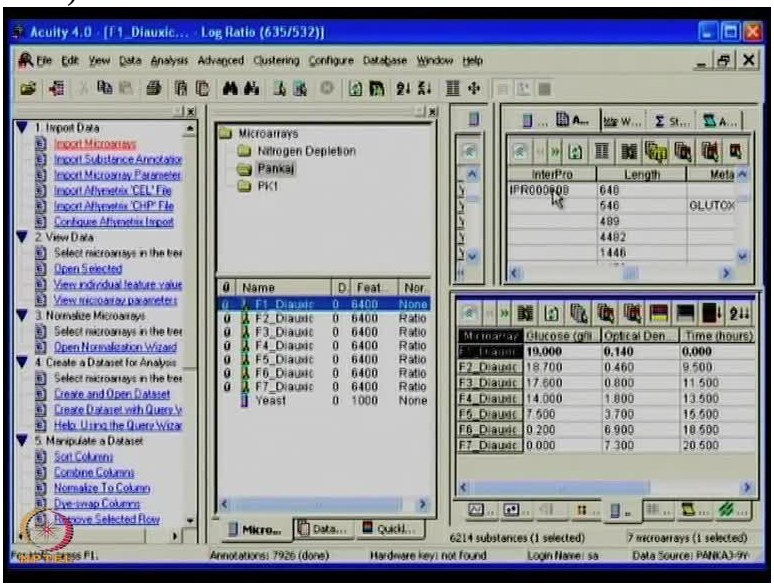
So another very important thing the parameter file, so as you have said already that it is very very important for one scientist to look all the parameters and group accordingly.

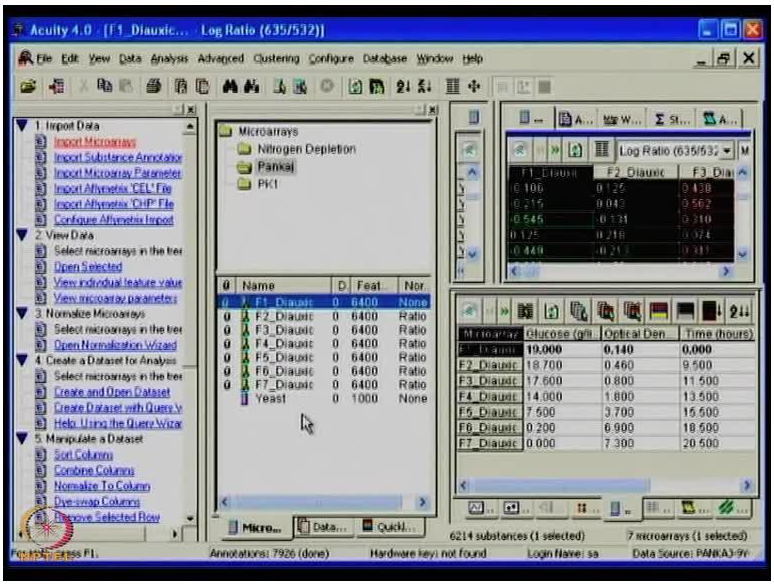
Mr Pankaj Khanna: Right.

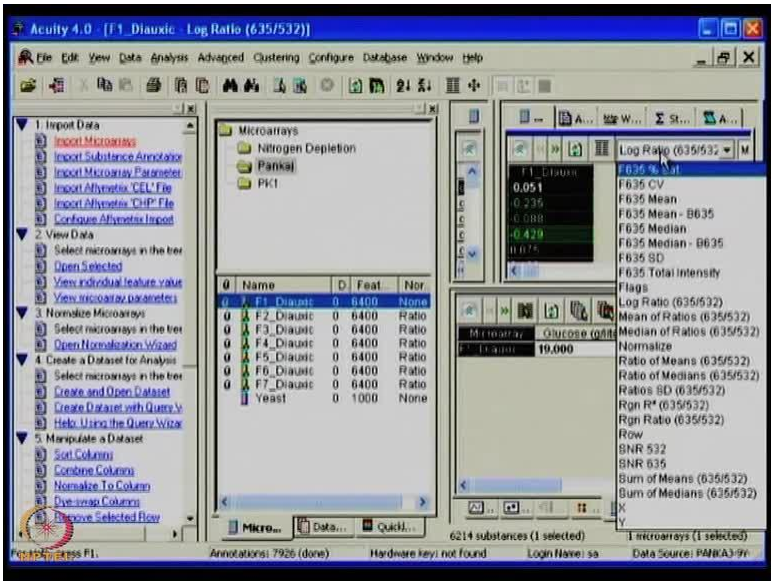
Prof Sanjeeva Srivastava: This can be make again in text delimited form and can be renamed to dot mdt file and these again can be imported to look at all the parameters are visible in the form in the down window here to look at. In in few seconds more you will understand what each window means but as it this is case you can just switch over to different types and you can just go a parameter file and look at what all details I have.

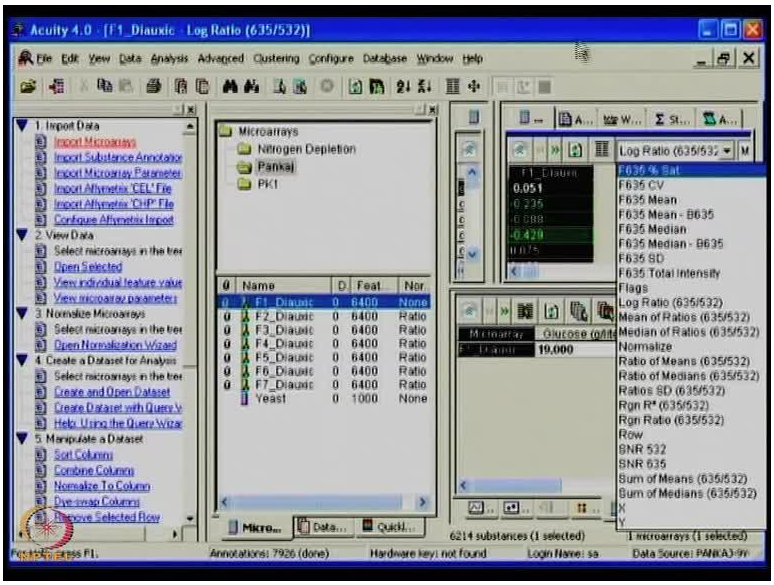
Prof Sanjeeva Srivastava: So maybe you just briefly explain each of the tabs so that uhh students are clear about a what is the they can infer from each of these window.

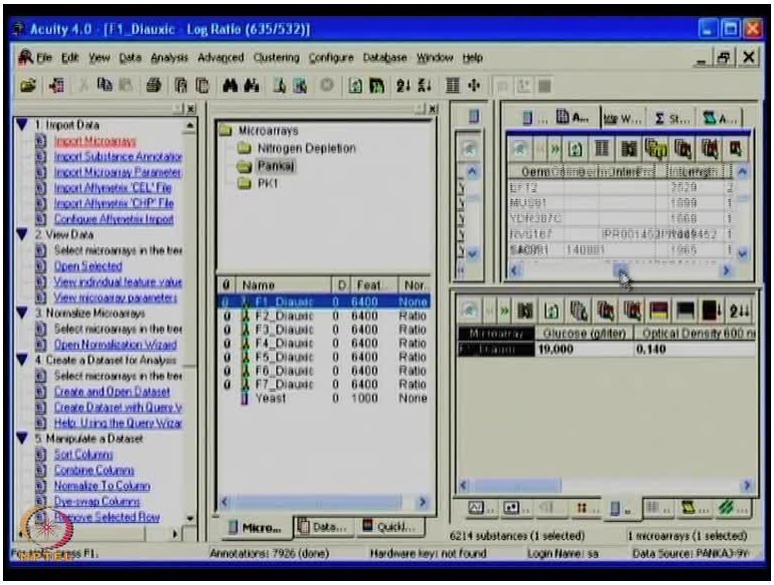
(Refer Slide Time: 6:08)

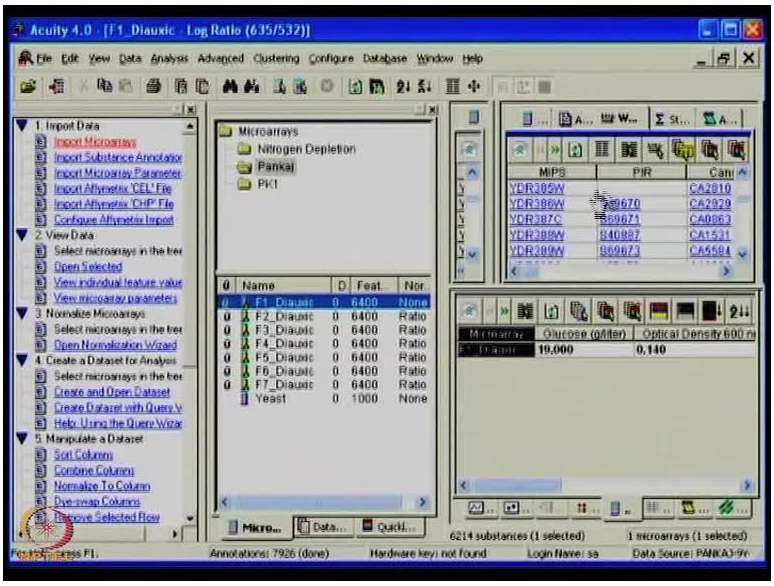


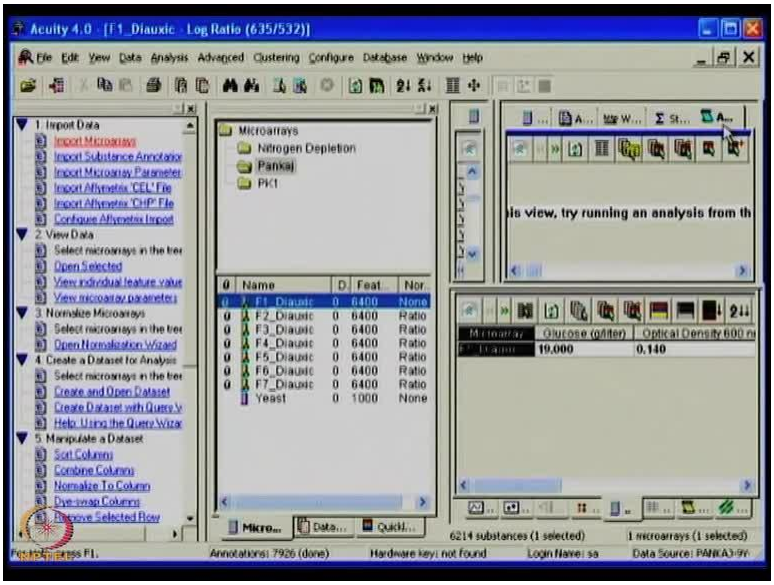

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quickl...
Log Ratio (635/532)
F1_Diauxic F2_Diauxic F3_Diauxic
0.106 0.125 0.438
0.215 0.043 0.562
0.545 -0.131 0.310
0.125 0.218 0.074
-0.449 -0.213 0.181
Microarray Glucose (g/l) Optical Den Time (hours)
F1_Diauxic 19.000 0.140 0.000
F2_Diauxic 18.700 0.460 9.500
F3_Diauxic 17.600 0.800 11.500
F4_Diauxic 14.000 1.800 13.500
F5_Diauxic 7.500 3.700 15.500
F6_Diauxic 0.200 6.900 18.500
F7_Diauxic 0.000 7.300 20.500
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

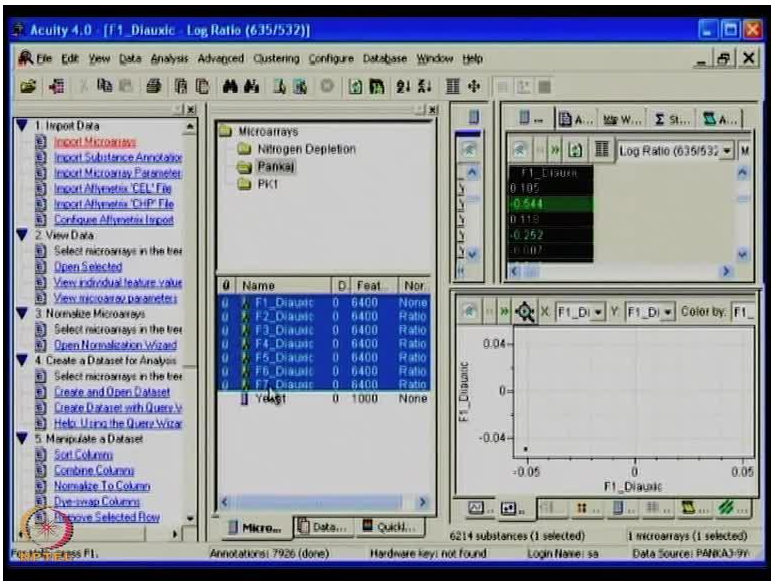

Acuity 4.0 - [F1_Diauxic - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quickl...
Log Ratio (635/532)
F1_Diauxic
0.051
0.235
0.088
-0.429
0.075
F635 % Sat
F635 CV
F635 Mean
F635 Mean - B635
F635 Median
F635 Median - B635
F635 SD
F635 Total Intensity
Flags
Log Ratio (635/532)
Mean of Ratios (635/532)
Median of Ratios (635/532)
Normalize
Ratio of Means (635/532)
Ratio of Medians (635/532)
Ratios SD (635/532)
Rgn R² (635/532)
Rgn Ratio (635/532)
Row
SNR 532
SNR 635
Sum of Means (635/532)
Sum of Medians (635/532)
X
Y
Microarray Glucose (g/l)
F1_Diauxic 19.000
6214 substances (1 selected)
1 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

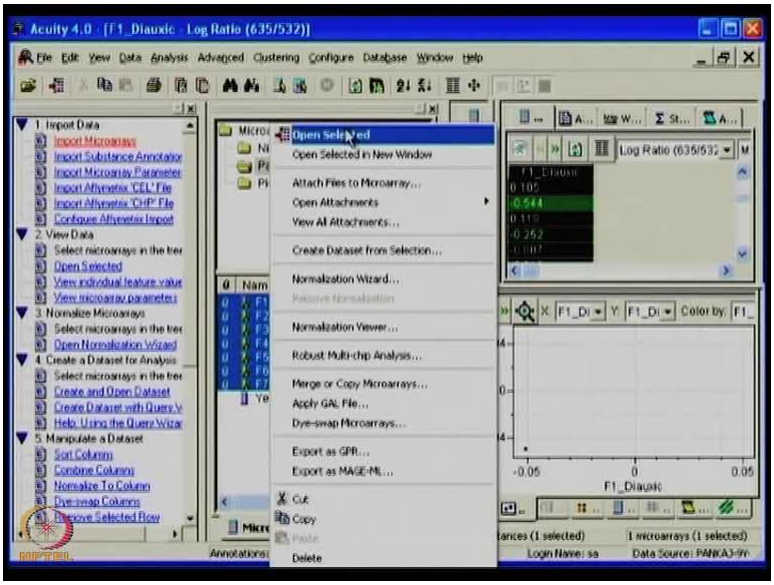

Acuity 4.0 - [F1_Diauxic - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quick...
Log Ratio (635/532)
F635 % Sat
F635 CV
F635 Mean
F635 Mean - B635
F635 Median
F635 Median - B635
F635 SD
F635 Total Intensity
Flags
Log Ratio (635/532)
Mean of Ratios (635/532)
Median of Ratios (635/532)
Normalize
Ratio of Means (635/532)
Ratio of Medians (635/532)
Ratios SD (635/532)
Rgn R² (635/532)
Rgn Ratio (635/532)
Row
SNR 532
SNR 635
Sum of Means (635/532)
Sum of Medians (635/532)
X
Y
F1_Diaux
0.051
-0.235
-0.588
-0.420
0.075
Microarray
Glucose (g/liter)
19,000
6214 substances (1 selected)
1 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

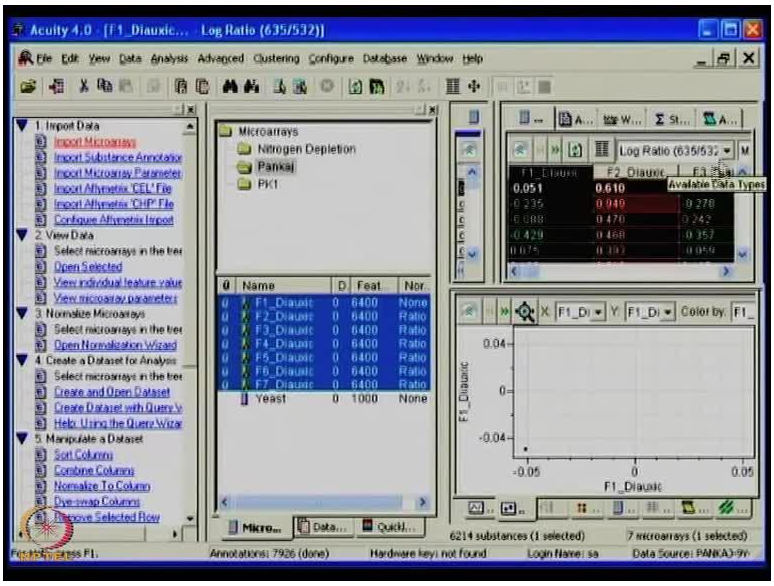

Acuity 4.0 - [F1_Diauxic - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quick...
Microarray
Glucose (g/liter)
Optical Density 600 nm
19,000
0.140
6214 substances (1 selected)
1 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

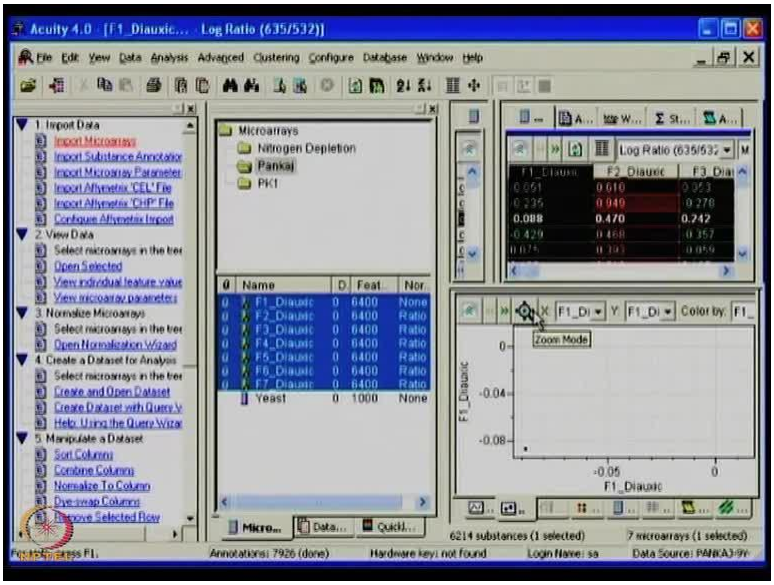

Acuity 4.0 - [F1_Diauxic - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
MIPS PIR
YDR385W CA2910
YDR386W CA2929
YDR387C CA0963
YDR388W CA1521
YDR389W CA5584
Glucose (g/liter) Optical Density 600 n
19.000 0.140
Micro... Data... Quickl...
6214 substances (1 selected) 1 microarrays (1 selected)
Annotations: 7926 (done) Hardware key: not found Login Name: sa

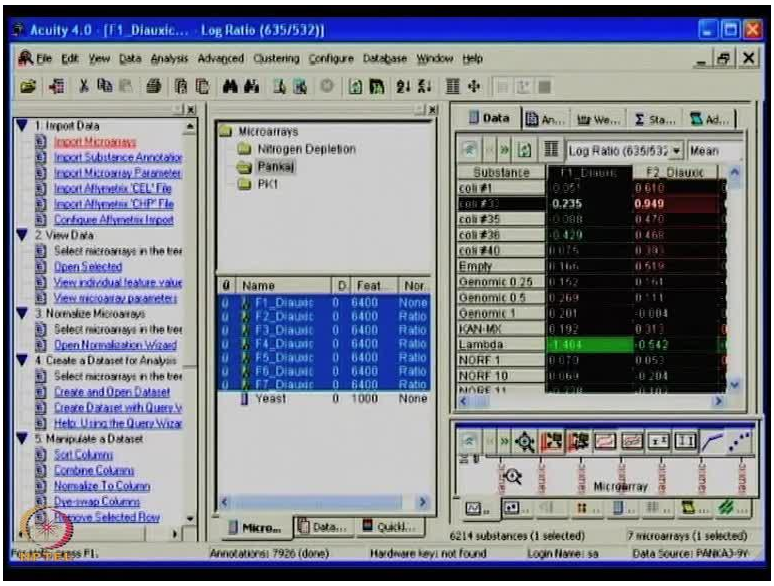

Acuity 4.0 - [F1_Diauxic - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
is view, try running an analysis from th
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Glucose (g/liter) Optical Density 600 n
19.000 0.140
Micro... Data... Quickl...
6214 substances (1 selected) 1 microarrays (1 selected)
Annotations: 7926 (done) Hardware key: not found Login Name: sa

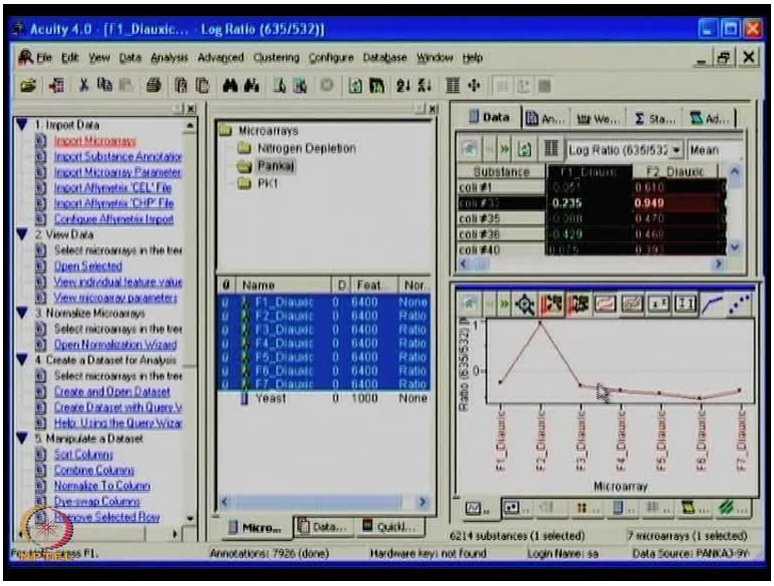

Acuity 4.0 - [F1_Diauxic - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizar
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quick...
Log Ratio (635/532
F1_Diauxic
0.105
0.544
0.119
-0.252
-0.007
X: F1_Di Y: F1_Di Color by: F1_
0.04
0
-0.04
F1_Diauxic
-0.05 0 0.05
F1_Diauxic
6214 substances (1 selected)
1 microarrays (1 selected)
For Help, press F1.
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

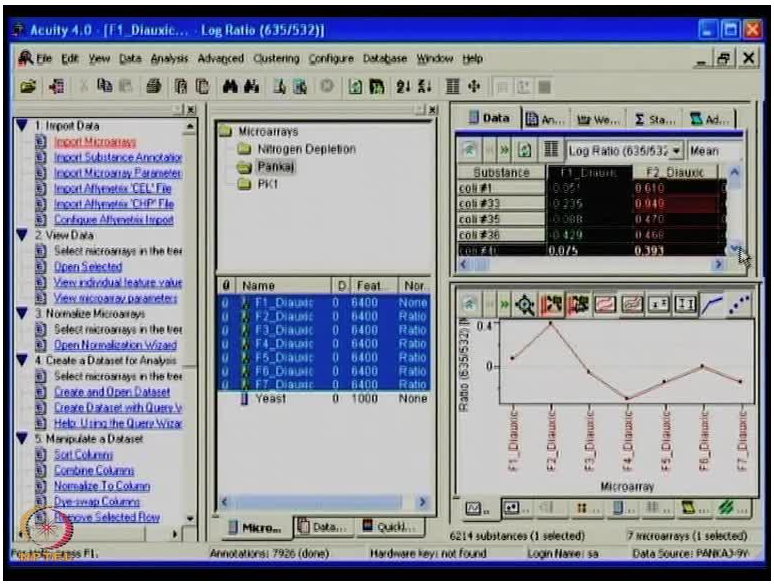

Acuity 4.0 - [F1_Diauxic - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Open Selected
Open Selected in New Window
Attach Files to Microarray...
Open Attachments
View All Attachments...
Create Dataset from Selection...
Normalization Wizard...
Remove Normalization
Normalization Viewer...
Robust Multi-chip Analysis...
Merge or Copy Microarrays...
Apply GAL File...
Dye-swap Microarrays...
Export as GPR...
Export as MAGE-ML...
Cut
Copy
Paste
Delete
Log Ratio (635/532
F1_Diauxic
0.544
-0.252
X: F1_Di Y: F1_Di Color by: F1_
-0.05 0 0.05
F1_Diauxic
1 microarrays (1 selected)
Login Name: sa
Data Source: PANKAJ-99


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizar
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quick...
Log Ratio (635/532)
F1_Diauxic F2_Diauxic F3_Diauxic
0.051 0.610 0.278
Available Data Types
X: F1_Di Y: F1_Di Color by: F1_
0.04
0
-0.04
-0.05 0 0.05
F1_Diauxic
6214 substances (1 selected)
7 microarrays (1 selected)
For Help, press F1.
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

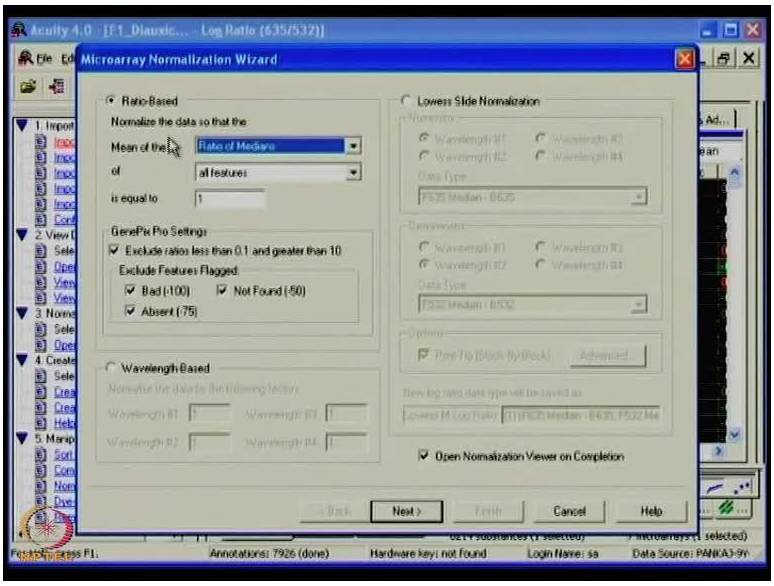

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizar
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quick...
Log Ratio (635/532)
F1_Diauxic F2_Diauxic F3_Diauxic
0.088 0.470 0.242
X: F1_Di Y: F1_Di Color by: F1_
Zoom Mode
0
-0.04
-0.08
-0.05 0
F1_Diauxic
6214 substances (1 selected)
7 microarrays (1 selected)
For Help, press F1.
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

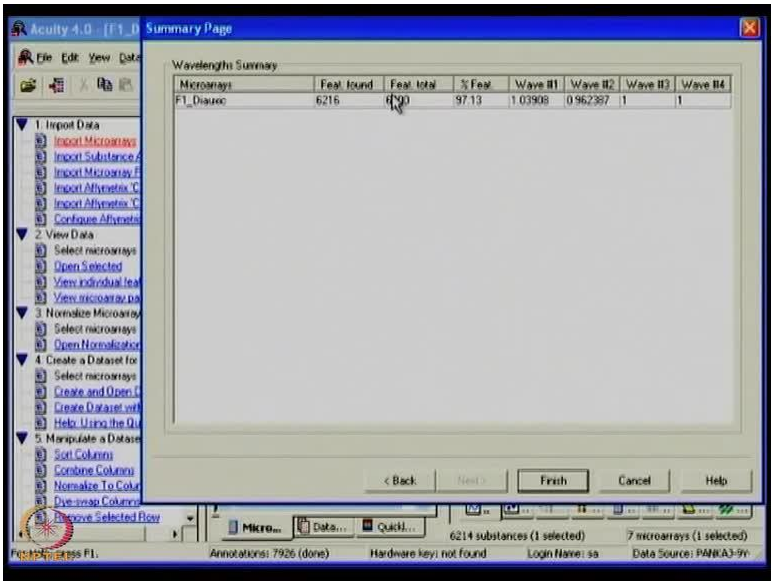

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor...
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro...
Data...
Quickl...
Data
An...
We...
Sta...
Ad...
Log Ratio (635/532)
Mean
Substance F1_Diauxic F2_Diauxic
coli #1
coli #33 0.235 0.949
coli #35
coli #36 0.429 0.468
coli #40
Empty
Genomic 0.25
Genomic 0.5
Genomic 1
KAN-MX
Lambda 1.464 -0.542
NORF 1
NORF 10
Microarray
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

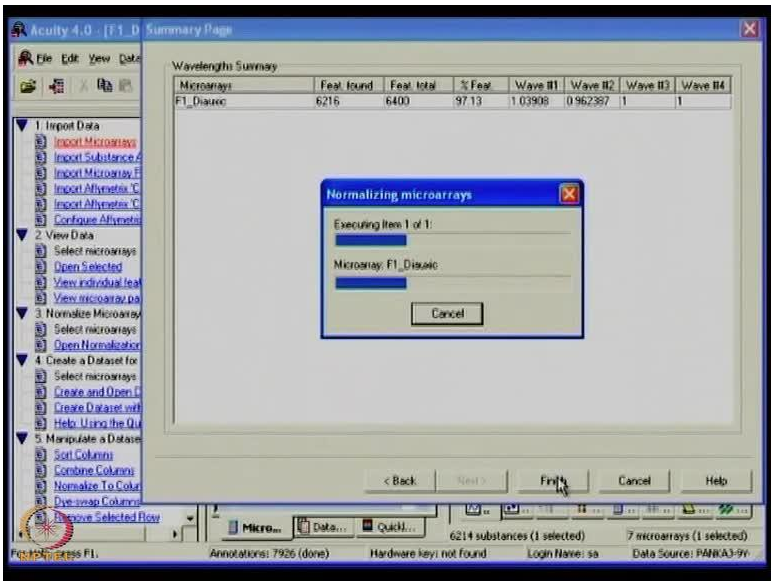

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor...
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro...
Data...
Quickl...
Data
An...
We...
Sta...
Ad...
Log Ratio (635/532)
Mean
Substance F1_Diauxic F2_Diauxic
coli #1
coli #33 0.235 0.949
coli #35
coli #36 0.429 0.468
coli #40
Ratio (635/532) [L
1
0
F1_Diauxic
F2_Diauxic
F3_Diauxic
F4_Diauxic
F5_Diauxic
F6_Diauxic
F7_Diauxic
Microarray
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

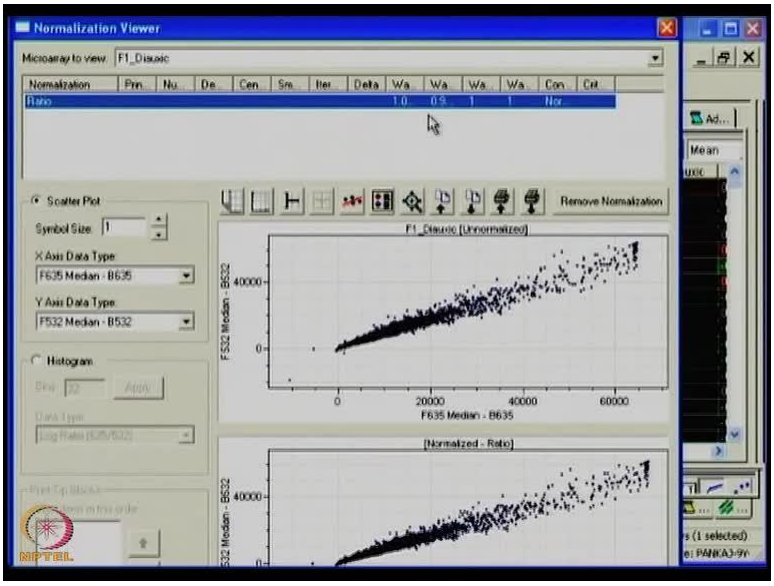

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameters
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quickl...
Data An... We... Sta... Ad...
Log Ratio (635/532) Mean
Substance F1_Diauxic F2_Diauxic
coli #1 0.610
coli #33 0.046
coli #35 0.470
coli #36 0.468
0.075 0.393
Ratio (635/532)
0.4
0
F1_Diauxic F2_Diauxic F3_Diauxic F4_Diauxic F5_Diauxic F6_Diauxic F7_Diauxic
Microarray
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-9


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameters
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name D Feat... Nor
F1_Diauxic 0 6400 None
F2_Diauxic 0 6400 Ratio
F3_Diauxic 0 6400 Ratio
F4_Diauxic 0 6400 Ratio
F5_Diauxic 0 6400 Ratio
F6_Diauxic 0 6400 Ratio
F7_Diauxic 0 6400 Ratio
Yeast 0 1000 None
Micro... Data... Quickl...
Data An... We... Sta... Ad...
Log Ratio (635/532) Mean
Substance F1_Diauxic F2_Diauxic
Empty 0.166 0.519
Genomic 0.25 0.152 0.061
Genomic 0.5 0.269 0.111
Genomic 1 0.201 0.004
KAN-MX 0.192 0.113
Lambda -1.464 -0.542
NORF 1 0.079 0.053
NORF 10 0.069 -0.204
NORF 11 0.220 -0.102
NORF 12 0.047 0.105
NORF 13 0.179 -0.329
NORF 14 0.080 0.196
NORF 15 0.290 -0.227
NORF 16 0.213 0.139
NORF 17 0.168 -0.356
NORF 18 0.089 -0.119
NORF 19 0.119 -0.378
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-9


Microarray Normalization Wizard
Ratio-Based
Normalize the data so that the
Mean of the
Ratio of Medians
of
all features
is equal to
1
GenePix Pro Settings:
Exclude ratios less than 0.1 and greater than 10
Exclude Features Flagged:
Bad (-100)
Not Found (-50)
Absent (-75)
Wavelength Based
Lowess Slide Normalization
Open Normalization Viewer on Completion
< Back
Next >
Finish
Cancel
Help


Summary Page
Wavelength Summary
Microarrays
Feat. found
Feat. total
% Feat
Wave #1
Wave #2
Wave #3
Wave #4
F1_Diauxic
6216
6400
97.13
1.03908
0.962387
1
1
< Back
Finish
Cancel
Help
6214 substances (1 selected)
7 microarrays (1 selected)





Mr Pankaj Khanna: Sure. So here in the working area which we have defined it is again splitted into 2 so which allows one one on the top to the level of different data visualization methods so what all features are there and what all different arrays say for example like only open one array it shows me only one array and it tells me what I am looking at. I am looking at log ratio data.

Prof Sanjeeva Srivastava: ya.

Mr Pankaj Khanna: In a similar fashion I can look at any other one because it is a GPR import whatever the genepix all the data you have to watch for. You can look at the background signal

individual intensities but majorly use is the log ratios for specially expression analysis but if you use protein or single wavelength base you can at look at only one wavelength base one.

So you can always control what you are watching. Apart from these the other tab include the annotation which gives you that what all different one is tracing at the level of annotation themes in the form of different uhh databases, information on the genes or how the protein is behaving or even the localization.

So where it is being localized, all that can be traced and the other one to that it also gives little bit of other details in the form of statistics, warehouses and few of the auto scripting capability which advanced user sometimes want to use. But this one particular statistic one allow you to see what all you want to see which we have seen in detail. In bottom what you essentially see is how the data is looked at.

Many a times what happens let us say quickly for example I want to go to one data and I would like to see how the first base is looking at. So what I am going to do is look at each particular spot and look at the profile of it. So because it is only one it is showing you one dot point.

Prof Sanjeeva Srivastava: Right.

Mr Pankaj Khanna: so if I in keep on including my more and more arrays the data plots start increasing and immediately one feature profiling how it has perform can immediately see and.

Prof Sanjeeva Srivastava: Right can we select now couple of arrays and is aligns…?

Mr Pankaj Khanna: So the way to select here is just hold the shift button and if you want to select only one one more or if you select all to the last all can get selected at shift. Then right click and click on open selected what it does is it allow you to open all the images here.

Prof Sanjeeva Srivastava: right

Mr Pankaj Khanna : So it is given you all this whatever calculated one you want to display and if you click on that each profile now can be seen here and with refresh button if you keep it will be able to see all. So in the down if I just click on the profile button based on what I have selected I can look at the different profiles how it has behaved. So in this fashion immediately I know my

parameter file that each one is what is it and I see that ok this is all normal average, in one of the case it went up. So similarly different features can be individually analyse and then checked at how the behaviour is happen.

Prof Sanjeeva Srivastava: So we can actually look at the trend for the same gene during the whole time course analysis.

Mr Pankaj Khanna: Correct correct. So you can also trace little bit of working on the data. Before going into this let me explain an important factor here what does each images mea. Actually if you carefully look may not be very clear that this is little purplish in colour and the other down ones are little reddish in colour.

Prof Sanjeeva Srivastava: ya.

Mr Pankaj Khanna: The purple one means that the data is not normalised, the red colour means that the data is normalised and the little dark green colour what you are seeing tells me that I have a JPEG image which I can see down her. So it allows me a connectivity of what is happening just by visualisation. If you wont say that the data is not normalised it is a easy process of doing it so you once imported the data on that you can just click right click and look at the normalisation result.

This normalization result allow one to choose different kind of normalization process which we have discussed earlier that it could be a ratio based or lowest normalisation based. It is continuous and it is discontinues type. So one can select but one has to remember the way one has been normalised, all my time point has to be normalized in same way.

Prof Sanjeeva Srivastava: Same way.

Mr Pankaj Khanna: so you cannot cross differences in the form of a different normalization and compare them. So it you are looking at little bit of different biology.

Prof Sanjeeva Srivastava: because until unless you have normalize in the same scheme you cannot compare those across like because you are going to compare different time points here.

Mr Pankaj Khanna: Correct so, as other ones are being analysed in the form of ratio based I am going select the one you have an opportunity to select different types I am going to select ratio of medians which is the most preferred and if you see again just the next button with all the flagging which is setif they are flagged please do not allow for the calculation.

Prof Sanjeeva Srivastava: okay.

Mr Pankaj Khanna: so right. And then you just click next it allows that ok it is available for me it is being done and I can say finish which allows that fine I have finished my normalization. So, if you carefully look back the spot there the purplish will change to red so which allows one to understand that yes all my images are now being uhh kind of normalized in a similar fashion.

Prof Sanjeeva Srivastava: So I guess we are dealing here with a lot of dataset so it takes some time for processing the whole thing.

Mr Pankaj Khanna: so it describes how many flags were there? So 6400 at a time were analysed and it sure see after doing the normalization and before normalisation how the data looks like and after normalisation how the data is looking like. So you can look an different ways so what we have done is we have corrected at the level of background and I am trying to display across how these an XY is being scatter together before and after normalization.

(Refer Slide Time: 12:09)
12:09f12:45

Once you are satisfied with this now if I look back now this has changed to same colour of reddish to from the purple which tells me okay this is also...

Prof Sanjeeva Srivastava: fall in normalized.

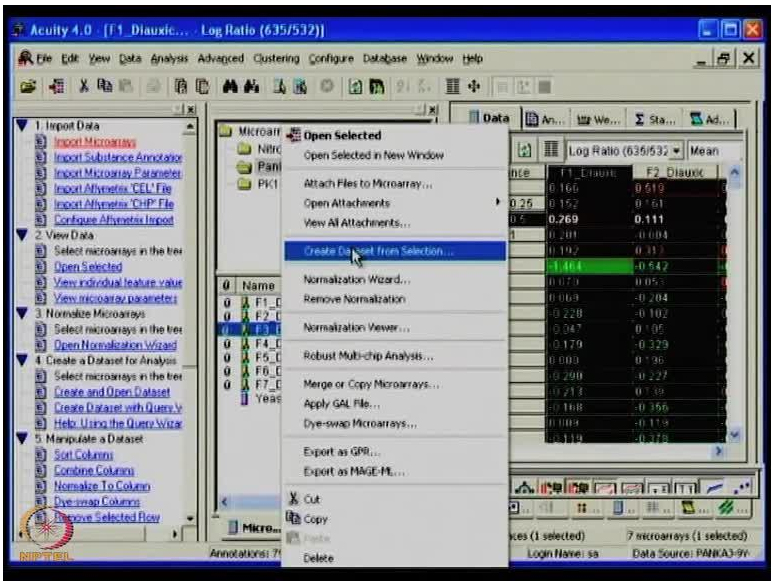
Mr Pankaj Khanna: Right. So you can also look at the say if I want to reconfirm which way I have done the normalization I can always go back and look at normalization viewer which allows one to say it is ratio based and you can go look back what kind of this one is being done for using the normalization process. So I can cross check once again how one is going about so once you have all the data been normalized after the import and you have all the places in the

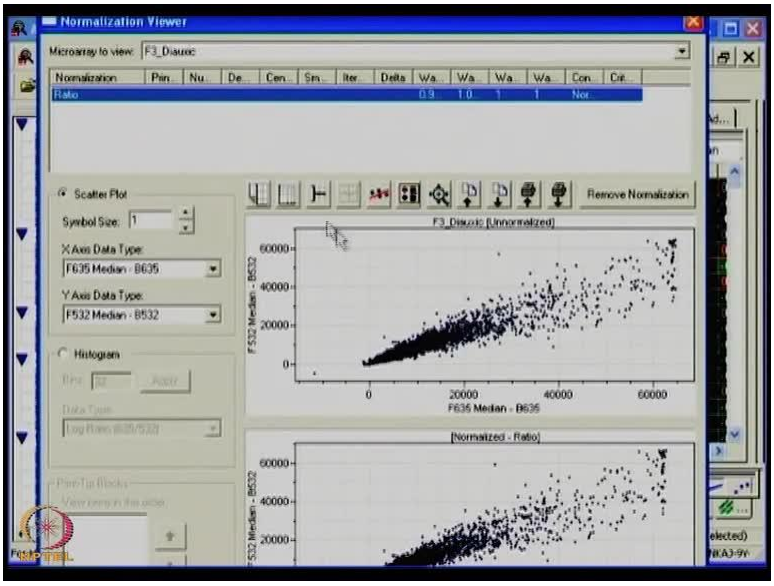
form of an annotation and the parameter file ready for you these are few ways which with which) you can look at the data. The next

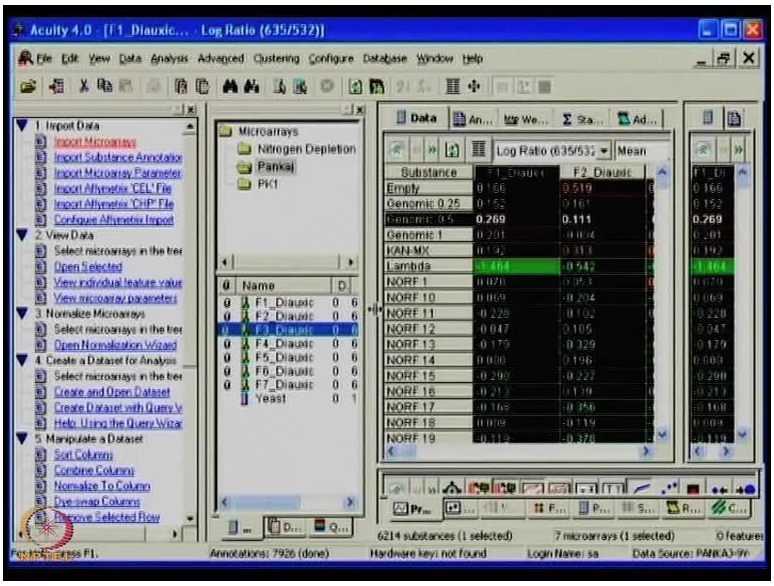
Prof Sanjeeva Srivastava: So I think before moving forward its important to ensure that normalization was done properly and one need to look at each slide carefully that.

Mr Pankaj Khanna: Yes. So as we discussed that it is very essential to have same normalization process done for all and it is not a thumb rule that which one is more preferable. One can choose anything but make sure all your different slides are being handled in the similar way. And you can do to different ways, get the data and do analyse in the different normalization processes also.

(Refer Slide Time: 13:28)


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizar
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Open Selected
Open Selected in New Window
Attach Files to Microarray...
Open Attachments
View All Attachments...
Create Dataset from Selection...
Normalization Wizard...
Remove Normalization
Normalization Viewer...
Robust Multi-chip Analysis...
Merge or Copy Microarrays...
Apply GAL File...
Dye-swap Microarrays...
Export as GPR...
Export as MAGE-ML...
Cut
Copy
Delete
Log Ratio (635/532)
Mean
F1_Diauxic
F2_Diauxic
Name
Data
Login Name: sa
Data Source: PANKA3-9V


Normalization Viewer
Microarray to view: F3_Diauxic
Normalization
Ratio
Scatter Plot
Symbol Size:
X Axis Data Type:
F635 Median - B635
Y Axis Data Type:
F532 Median - B532
Histogram
Remove Normalization
F3_Diauxic [Unnormalized]
F532 Median - B532
F635 Median - B635
[Normalized - Ratio]
0
20000
40000
60000



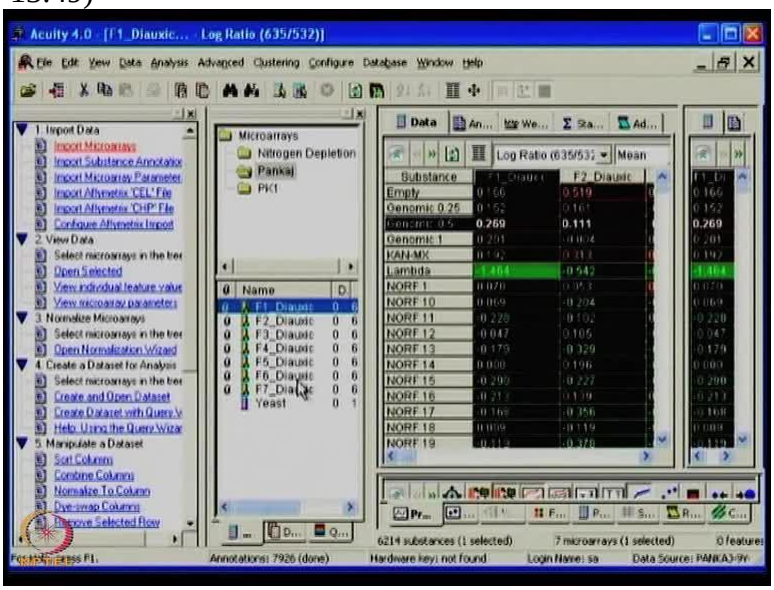
So, one has an opportunity to even correct that so because you have raised the point say for example I want to remove this normalization and put some other normalization I can just click here remove normalisation it remove normalisation allows Go back to the raw data…

Mr Pankaj Khanna: And again you can re normalize.

Prof Sanjeeva Srivastava: … Renormalize in a different size, may be you want do now lowers normalisation and check back all in that format.

Prof Sanjeeva Srivastava: okay.

(Refer Slide Time: 13:49)



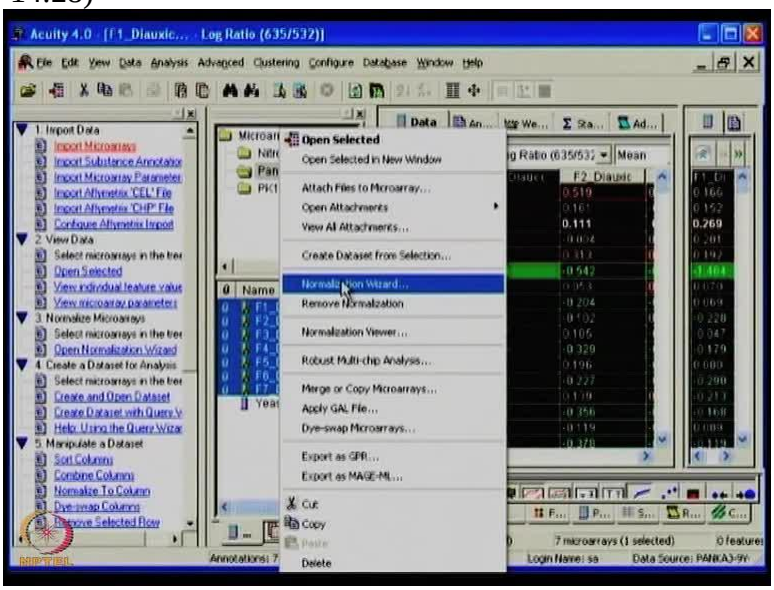
Mr Pankaj Khanna: And you can also select multiple in the similar way and in one short itself we can do normalisation is single one so this is what I prefer so usually you don not mess around with the different kind of normalisation either select or are remove all because I have imported one to show you how the process is being done here.

Prof Sanjeeva Srivastava: okay.

Mr Pankaj Khanna: And then immediately one would like to see how my data looks like the one way to look is the number's which is little cumbersome other way people like is colour. If you carefully observe the colouring scheme is going here red, black and green.

Prof Sanjeeva Srivastava: Ya may be you can tell that is a conventional things for the microarray, people always represent these colours so what each of this colour code means.

(Refer Slide Time: 14:28)



Mr Pankaj Khanna: Correct. So the black colour is towards one. The meaning is when I am looking at the data you expect that when green and red channel both are giving same colour, same intensity it becomes blackish in colour. If they are up regulated people put them towards the red colour and if it is down regulated - sign will be given and that will become down regulated.

The idea with this is which laser is being used what. So in context we have shown you are using which kind of ratio means to check so usually it is case over control what people report for.

Prof Sanjeeva Srivastava: ya.

Mr Pankaj Khanna: right? so in this fashion, conventionally you can see the colours but here there are many ones there are many which we have opened now six. So I want to see in nutshell what is happening acuity allows you to do it by seeing you can do an auto feed colour so quickly the numbers have gone only colours are shown to tell you how each particular substance or gene has behaved across your samples. You can look this is being dash I have just spilt table so that I can look back at the annotation also so I can have just a spilt table available, put an annotation file here so that I keep looking at what I am interested in.

Prof Sanjeeva Srivastava: okay

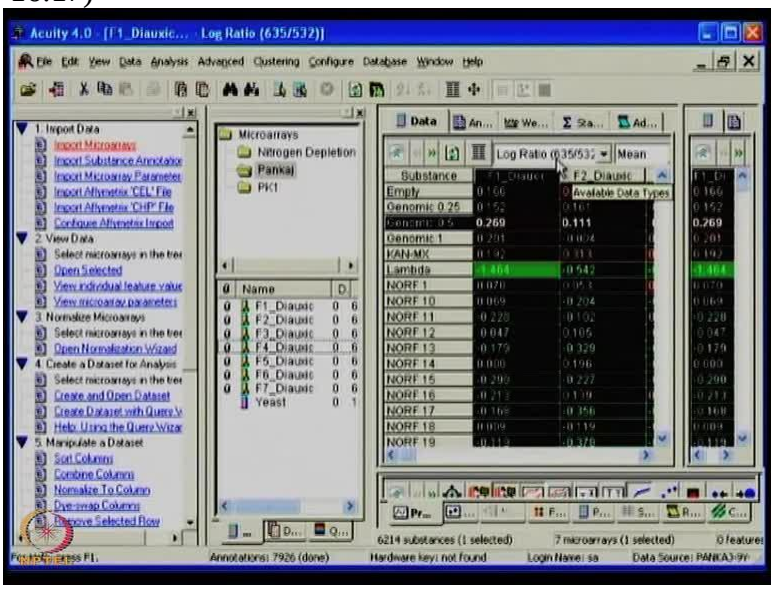
Mr Pankaj Khanna: So if there is enough opportunity for you to play around so how you want to look and customise your view.

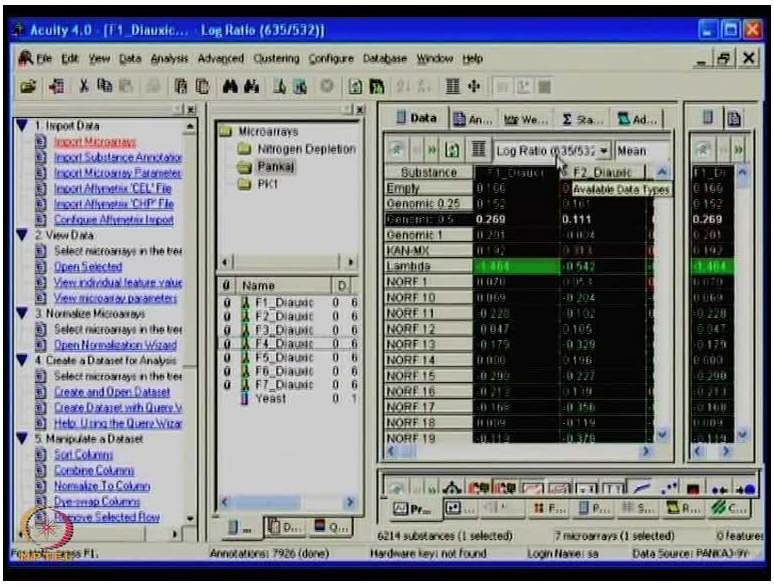
Prof Sanjeeva Srivastava: So, I think this type of heat maps right way gives you a feel about what type of genes across each time point has shown the variation or modulation and in the expression profile uhh looking at the colour itself like for the first one uhh I can just easily say ok it is going down as we are moving across the towards 20th hour.

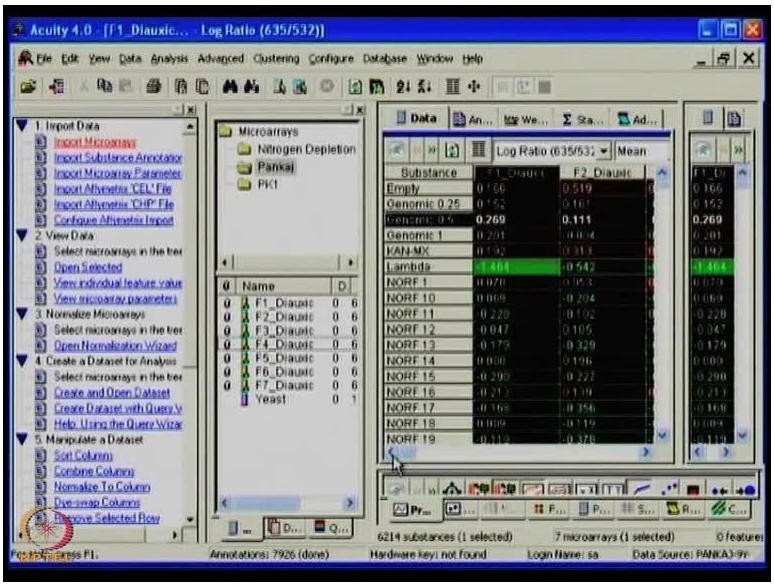
Mr Pankaj Khanna: True.

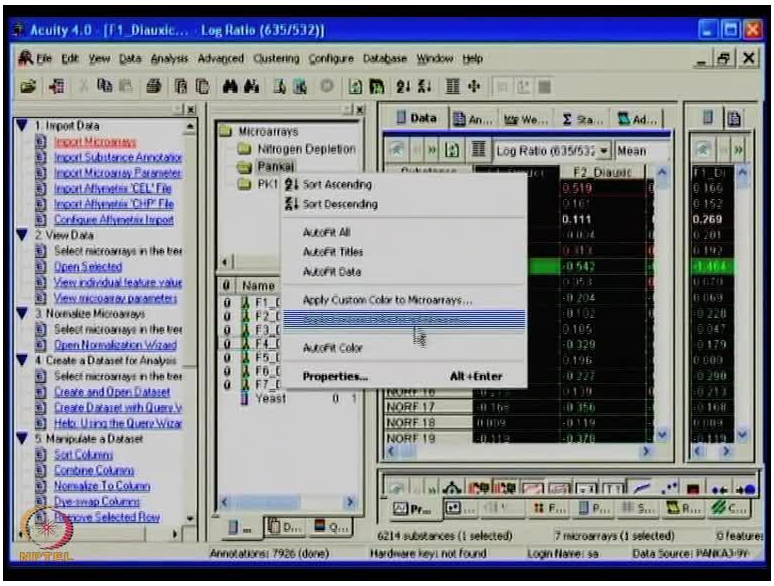
Prof Sanjeeva Srivastava: uhh And so I think by looking at type of data one can visually actually get the feel about the expression changes across the different time point.

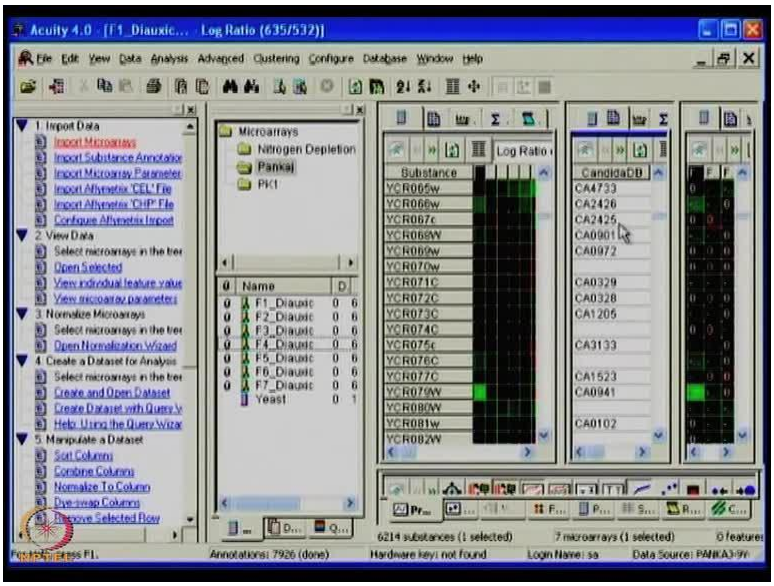
(Refer Slide Time: 16:17)

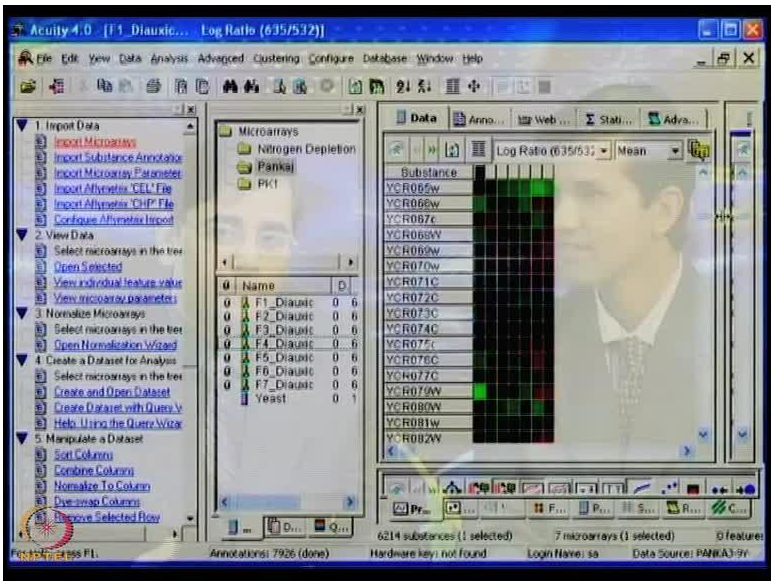


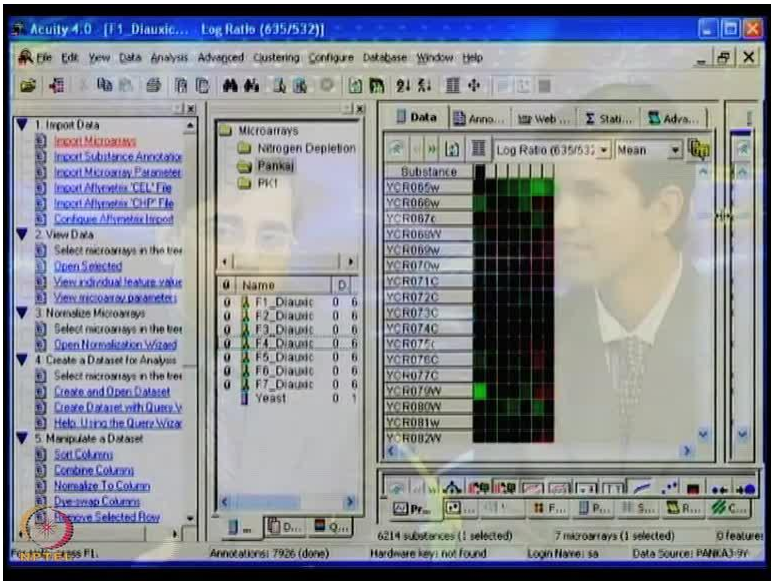

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name
F1_Diauxic 0 6
F2_Diauxic 0 6
F3_Diauxic 0 6
F4_Diauxic 0 6
F5_Diauxic 0 6
F6_Diauxic 0 6
F7_Diauxic 0 6
Yeast 0 1
Data
Log Ratio (635/532)
Mean
Substance
F2_Diauxic
Empty
Genomic 0.25
Genomic 0.5
Genomic 1
KAN-MX
Lambda
NORF 1
NORF 10
NORF 11
NORF 12
NORF 13
NORF 14
NORF 15
NORF 16
NORF 17
NORF 18
NORF 19
Available Data Types
0.269
0.111
-1.464
-0.542
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa

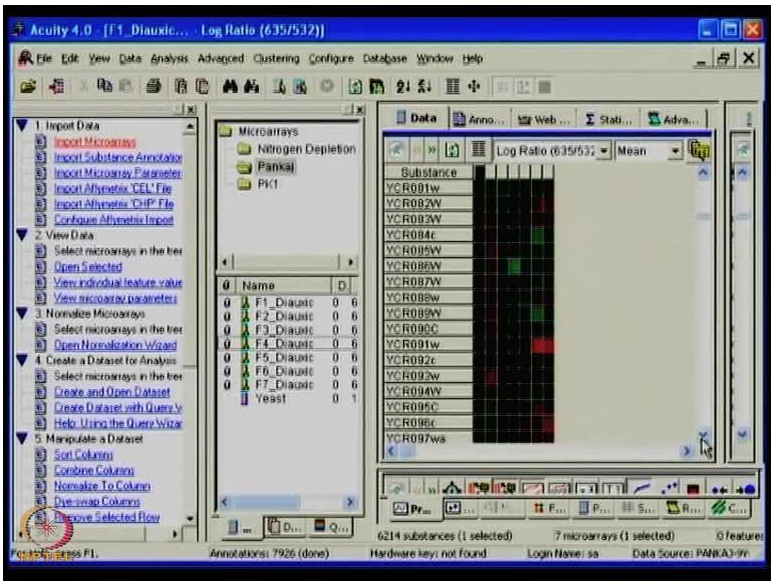

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name
F1_Diauxic 0 6
F2_Diauxic 0 6
F3_Diauxic 0 6
F4_Diauxic 0 6
F5_Diauxic 0 6
F6_Diauxic 0 6
F7_Diauxic 0 6
Yeast 0 1
Data
Log Ratio (635/532)
Mean
Substance
F2_Diauxic
Empty
Genomic 0.25
Genomic 0.5
Genomic 1
KAN-MX
Lambda
NORF 1
NORF 10
NORF 11
NORF 12
NORF 13
NORF 14
NORF 15
NORF 16
NORF 17
NORF 18
NORF 19
0.519
0.269
0.111
-1.464
-0.542
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa

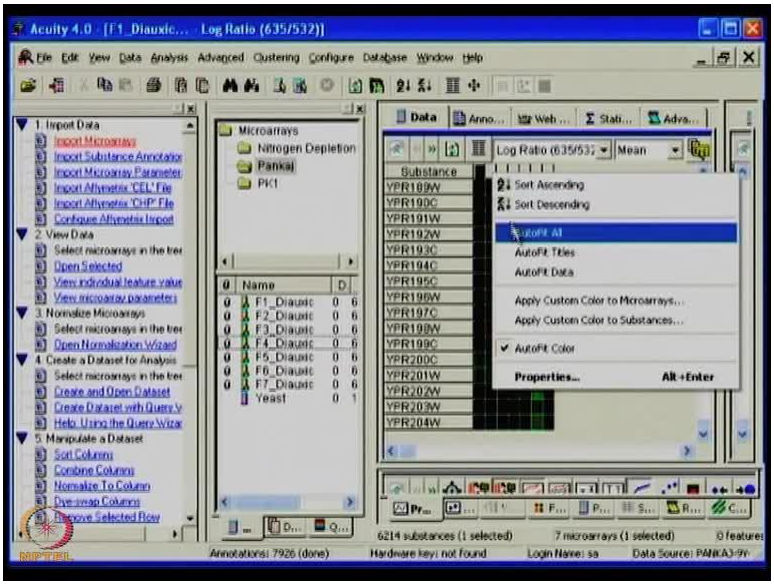

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Sort Ascending
Sort Descending
AutoFit All
AutoFit Titles
AutoFit Data
Apply Custom Color to Microarrays...
AutoFit Color
Properties... Alt+Enter
Name
F1_D
F2_D
F3_D
F4_D
F5_D
F6_D
F7_D
Yeast 0 1
Data
Log Ratio (635/532
Mean
F2_Diauxic
0.519
0.111
0.269
-0.542
-0.204
0.102
0.105
-0.329
-0.196
-0.227
-0.139
NORF 17 -0.165 -0.356
NORF 18 -0.009 -0.119
NORF 19 -0.113 -0.378
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

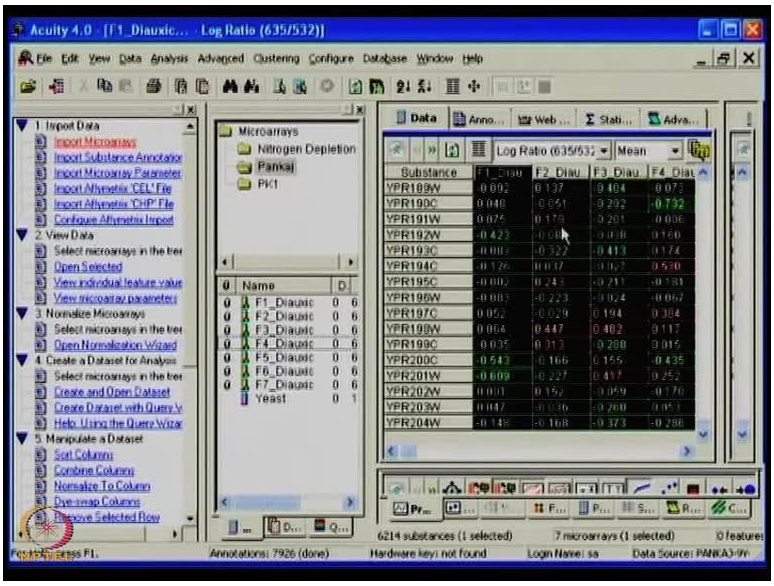

Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name
F1_Diauxic 0 6
F2_Diauxic 0 6
F3_Diauxic 0 6
F4_Diauxic 0 6
F5_Diauxic 0 6
F6_Diauxic 0 6
F7_Diauxic 0 6
Yeast 0 1
Log Ratio
Substance
YCR065w
YCR066w
YCR067c
YCR068W
YCR069w
YCR070w
YCR071C
YCR072C
YCR073C
YCR074C
YCR075c
YCR076C
YCR077C
YCR079W
YCR080W
YCR081w
YCR082W
CandidaDB
CA4733
CA2426
CA2425
CA0901
CA0972
CA0329
CA0328
CA1205
CA3133
CA1523
CA0941
CA0102
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

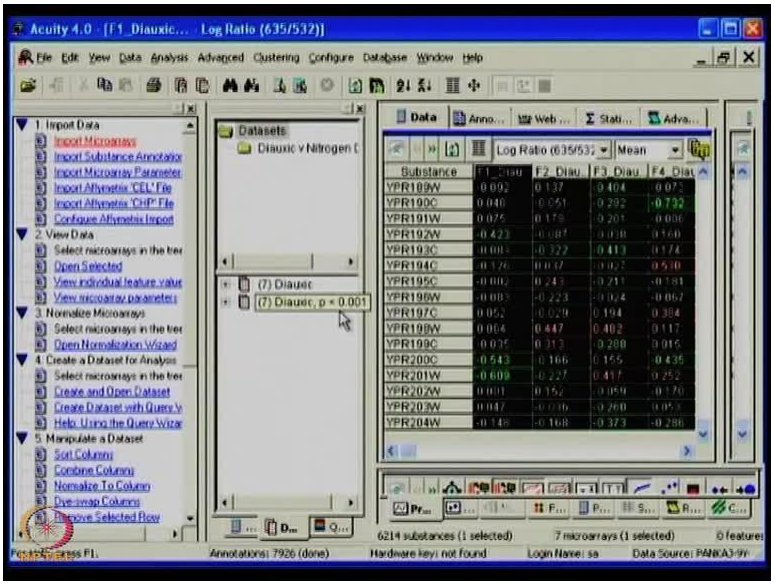

Acuity 4.0 - [F1_Diauxic... Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name
F1_Diauxic 0 6
F2_Diauxic 0 6
F3_Diauxic 0 6
F4_Diauxic 0 6
F5_Diauxic 0 6
F6_Diauxic 0 6
F7_Diauxic 0 6
Yeast 0 1
Data
Log Ratio (635/532)
Mean
Substance
YCR065W
YCR066W
YCR067C
YCR068W
YCR069W
YCR070W
YCR071C
YCR072C
YCR073C
YCR074C
YCR076C
YCR077C
YCR079W
YCR080W
YCR081W
YCR082W
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa


Acuity 4.0 - [F1_Diauxic... Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name
F1_Diauxic 0 6
F2_Diauxic 0 6
F3_Diauxic 0 6
F4_Diauxic 0 6
F5_Diauxic 0 6
F6_Diauxic 0 6
F7_Diauxic 0 6
Yeast 0 1
Data
Log Ratio (635/532)
Mean
Substance
YCR065W
YCR066W
YCR067C
YCR068W
YCR069W
YCR070W
YCR071C
YCR072C
YCR073C
YCR074C
YCR076C
YCR077C
YCR079W
YCR080W
YCR081W
YCR082W
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name
F1_Diauxic
F2_Diauxic
F3_Diauxic
F4_Diauxic
F5_Diauxic
F6_Diauxic
F7_Diauxic
Yeast
Data
Anno...
Web ...
Stati...
Adva...
Log Ratio (635/532
Mean
Substance
YCR081w
YCR082W
YCR083W
YCR084c
YCR085W
YCR086W
YCR087W
YCR088w
YCR089W
YCR090C
YCR091w
YCR092c
YCR093w
YCR094W
YCR095C
YCR096c
YCR097wa
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Sort Ascending
Sort Descending
AutoFit All
AutoFit Titles
AutoFit Data
Apply Custom Color to Microarrays...
Apply Custom Color to Substances...
AutoFit Color
Properties... Alt+Enter
YPR189W
YPR190C
YPR191W
YPR192W
YPR193C
YPR194C
YPR195C
YPR196W
YPR197C
YPR198W
YPR199C
YPR200C
YPR201W
YPR202W
YPR203W
YPR204W
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa





Mr Pankaj Khanna: Very true. So this gives you immediate visualisation tool to understand what is happening and get a rough idea and this is my new just the neat raw data. I have not perform just the normalization and we just seeing how they are behaving so it gives you a rough profile okay I have some biology which is going for this particular design of experiment. So with here on if I want to go back to numbers or auto feed my data I can select appropriate one auto feed all data so it says it just feeded based on that so it again shows you the number back.

Prof Sanjeeva Srivastava: okay

Mr Pankaj Khanna: So we have got the data imported now, we have done the normalizations, we are trying to see how they have behaved. Very important thing which sometimes people like case in the form of like able to move the data sorting up and down but before doing that acuity tells you that first you make a data set.

Prof Sanjeeva Srivastava: okay.

Mr Pankaj Khanna: The meaning of data set is this is just looking at the raw data and I want to extract the data and allow to keep in a data set here towards the down.

Prof Sanjeeva Srivastava: okay.

Mr Pankaj Khanna: so there you are available with all the different kind of things. So now there are two ways of doing things in the data set. So one thing is take all the features and sometimes people say I want to have my criteria is defined such that my visualisation makes more sense to me.

Prof Sanjeeva Srivastava: So maybe one can actually be very stringent at this stage itself and say okay I want to only the very very biological significant ones.

Mr Pankaj Khanna: True.

Prof Sanjeeva Srivastava: So define the p values.

Mr Pankaj Khanna: True.

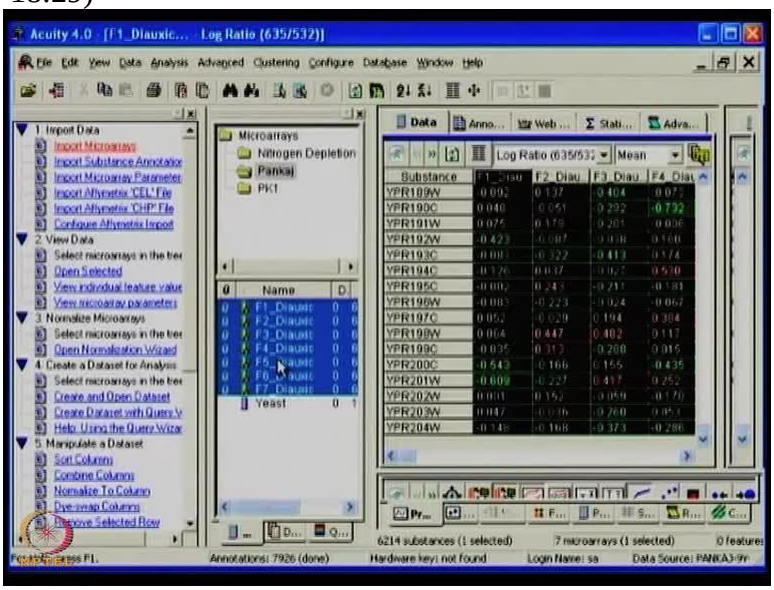
Prof Sanjeeva Srivastava: And then just a sort the data based on that.

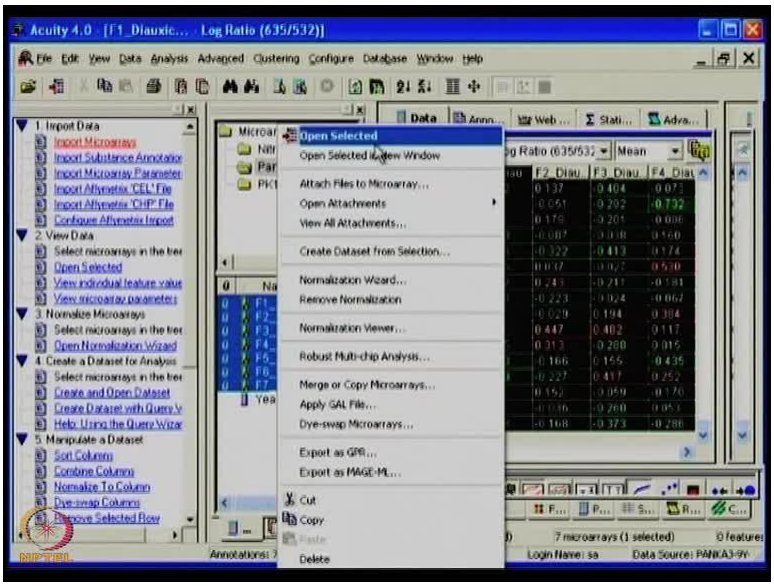
Mr Pankaj Khanna: True. So people can do okay which are changes up and down with a range of so and so which is twofold up regulated to fold down regulated so usually differential expressions data is being logged the meaning of log to the base 2 is essentially log to the base 2 value 1 is equal two fold change. So you are talking of log to the base 2 means you are taking four fold change which really become significant you can filter based on different parameters and generate the data sets.

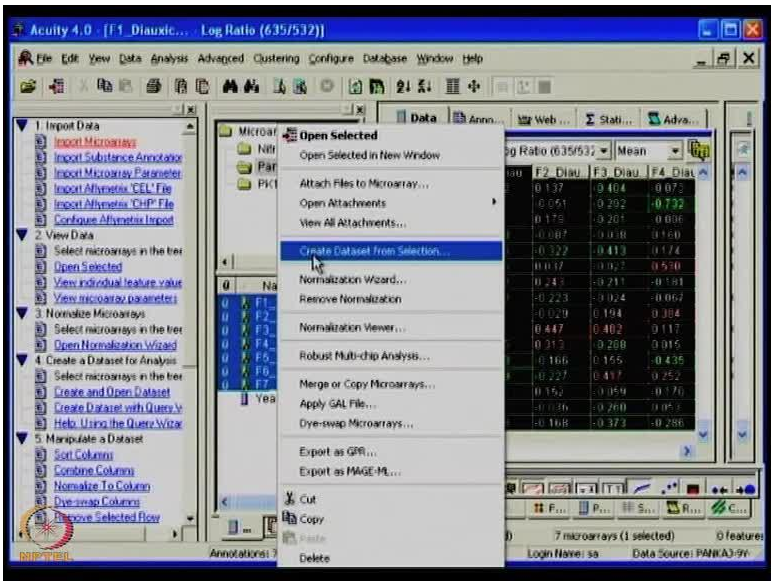
So essentially you are reducing the numbers so you make more sense in the form of visualization otherwise also you take all the data and you can do it. Let us quickly see how we can do that particular job.

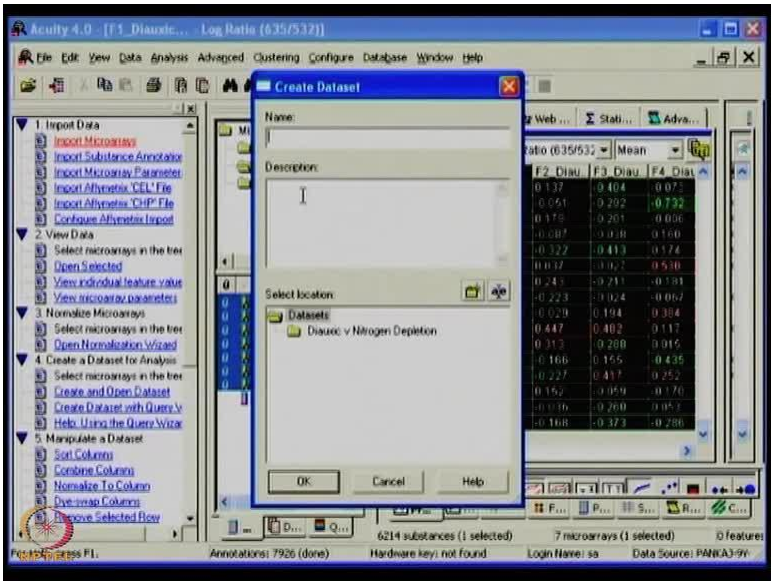
Prof Sanjeeva Srivastava: sure.

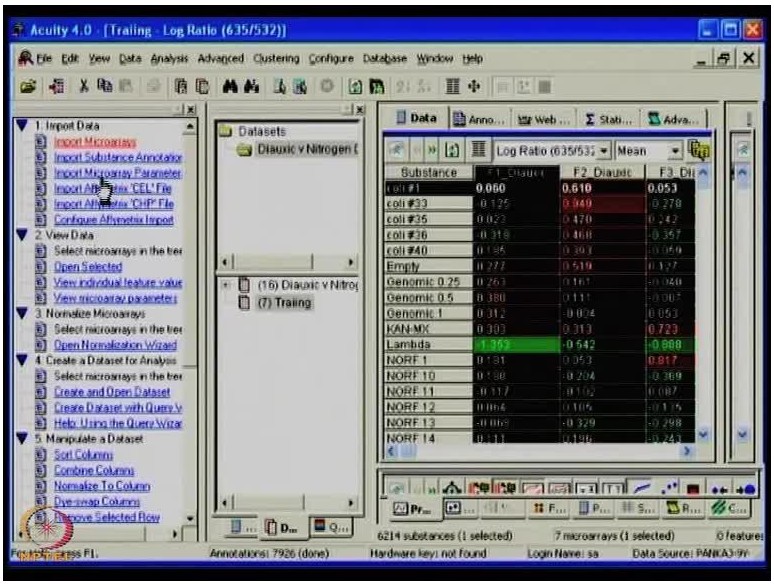
(Refer Slide Time: 18:29)

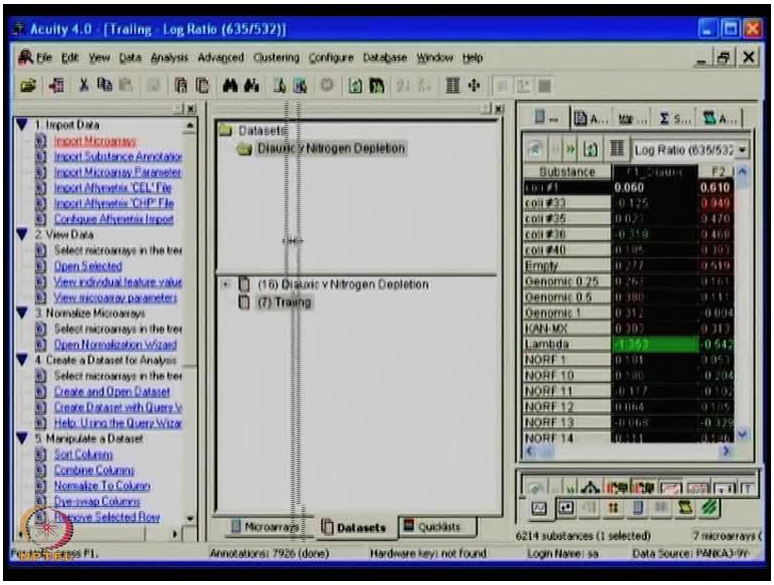





Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Open Selected
Open Selected in New Window
Attach Files to Microarray...
Open Attachments
View All Attachments...
Create Dataset From Selection...
Normalization Wizard...
Remove Normalization
Normalization Viewer...
Robust Multi-chip Analysis...
Merge or Copy Microarrays...
Apply GAL File...
Dye-swap Microarrays...
Export as GPR...
Export as MAGE-ML...
Cut
Copy
Delete
Mean
7 microarrays (1 selected)
Login Name: sa


Acuity 4.0 - [F1_Diauxic... - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Create Dataset
Name:
Description:
Select location:
Datasets
Diauxic v Nitrogen Depletion
OK
Cancel
Help
1. Import Data
2. View Data
3. Normalize Microarrays
4. Create a Dataset for Analysis
5. Manipulate a Dataset
Mean
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa


Acuity 4.0 - [Trailing - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Datasets
Diauxic v Nitrogen
(16) Diauxic v Nitro
(7) Trailing
Data
Log Ratio (635/532)
Mean
Substance
F1_Diauxic
F2_Diauxic
F3_Di
coli #1 0.060 0.610 0.053
coli #33 0.125 0.949 0.278
coli #35 0.023 0.470 0.242
coli #36 -0.318 0.468 -0.357
coli #40 0.195 0.393 -0.059
Empty 0.272 0.519 0.127
Genomic 0.25 0.263 0.161 -0.040
Genomic 0.5 0.380 0.111 -0.007
Genomic 1 0.312 0.004 0.053
KAN-MX 0.393 0.313 0.723
Lambda 1.363 -0.542 -0.868
NORF 1 0.181 0.053 0.817
NORF 10 0.180 -0.204 -0.369
NORF 11 -0.117 0.102 0.087
NORF 12 0.064 0.105 -0.135
NORF 13 -0.065 -0.329 -0.298
NORF 14 0.131 0.196 -0.243
6214 substances (1 selected)
7 microarrays (1 selected)
0 features
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99


Acuity 4.0 - [Trailing - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotation
Import Microarray Parameter
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature value
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Datasets
Diauxic v Nitrogen Depletion
(16) Diauxic v Nitrogen Depletion
(7) Trailing
Microarrays
Datasets
Quicklists
Log Ratio (635/532)
Substance
F2
coli #1 0.060 0.610
coli #33 0.125 0.949
coli #35 0.023 0.470
coli #36 -0.318 0.468
coli #40 0.195 0.393
Empty 0.272 0.519
Genomic 0.25 0.263 0.161
Genomic 0.5 0.380 0.111
Genomic 1 0.312 0.004
KAN-MX 0.393 0.313
Lambda 1.363 -0.542
NORF 1 0.181 0.053
NORF 10 0.180 -0.204
NORF 11 -0.117 0.102
NORF 12 0.064 0.105
NORF 13 -0.065 -0.329
NORF 14 0.131 0.196
6214 substances (1 selected)
7 microarrays
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-99

Mr Pankaj Khanna: So I can select my data again holding the shift I can click and I want it will tell me what you want do. It says you can have different kind of opportunities but you can create datasets from selections. So, this allows that whatever I have selected do create a data set from that. Once I click that it says where you want to keep in the folder. So beforehand I can generate my own folders I can define my studies.

Prof Sanjeeva Srivastava: Right.

Mr Pankaj Khanna: And I can place them say for example I am going to give it a name call training and it creates all the data from that so here you go from the seven microarrays you have got and you have got all the data.

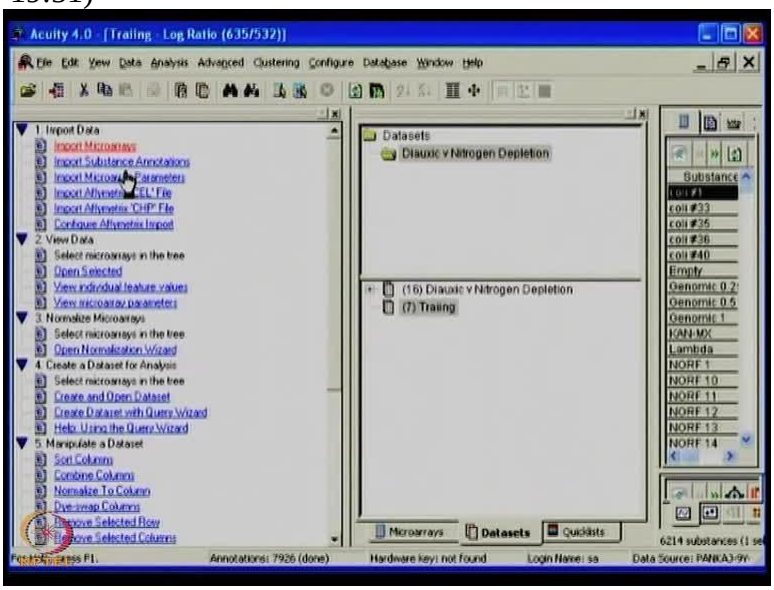
Now this is the one to get all complete data. You can also have other ways as you suggested P value importance coming or log fold change value is coming in. You can define a criteria of doing that. The way to do that again look at a common task pane. we have done the three steps and now after doing normalisation I can click on create an open data sets.

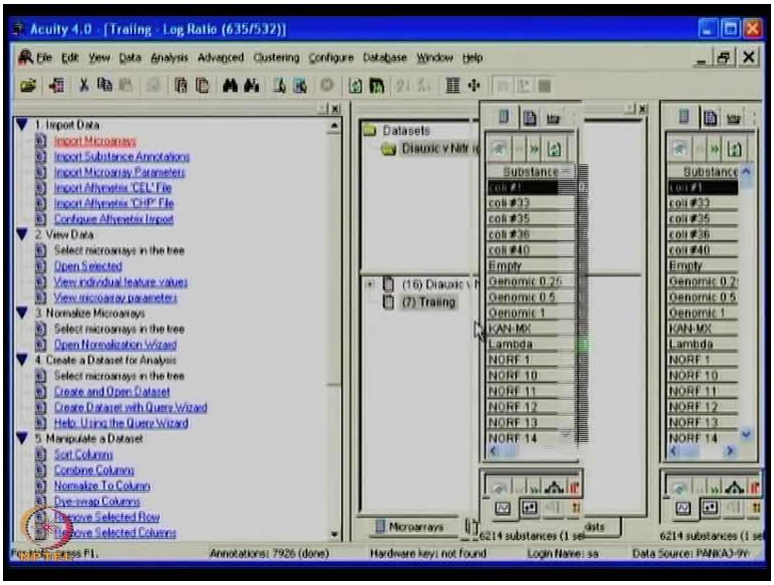
Prof Sanjeeva Srivastava: Before actually you move just I think good idea you just refresh again uhh sort of the task step wise.

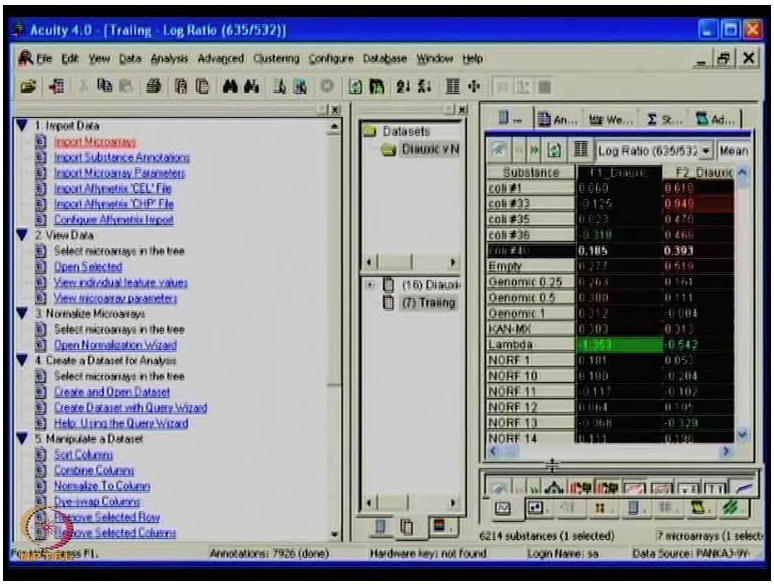
Mr Pankaj Khanna: Sure.

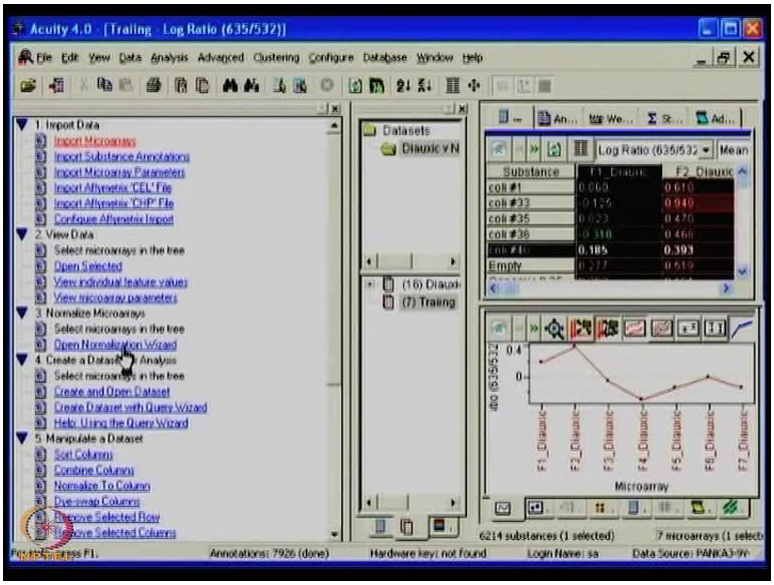
Prof Sanjeeva Srivastava: So first was a import data.

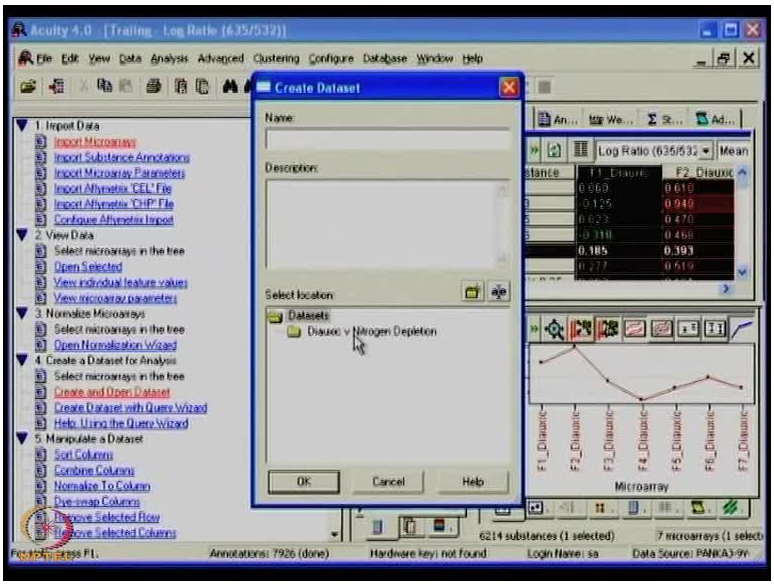
(Refer Slide Time: 19:51)

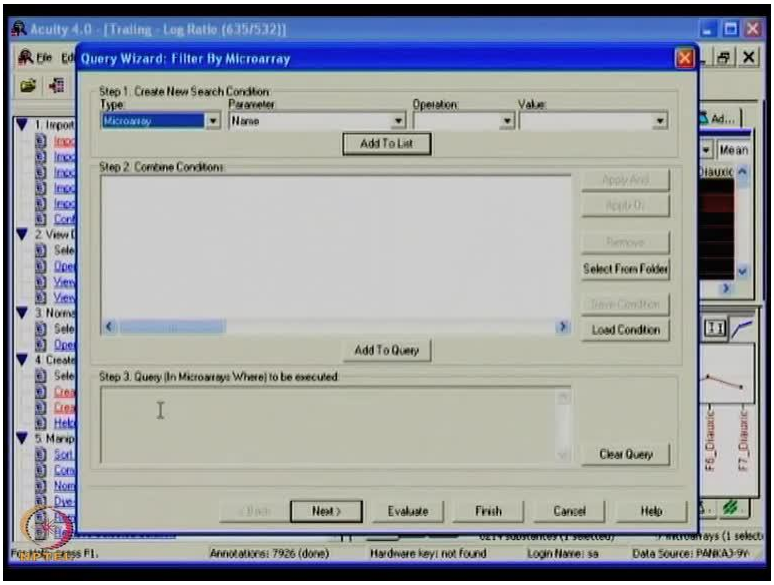


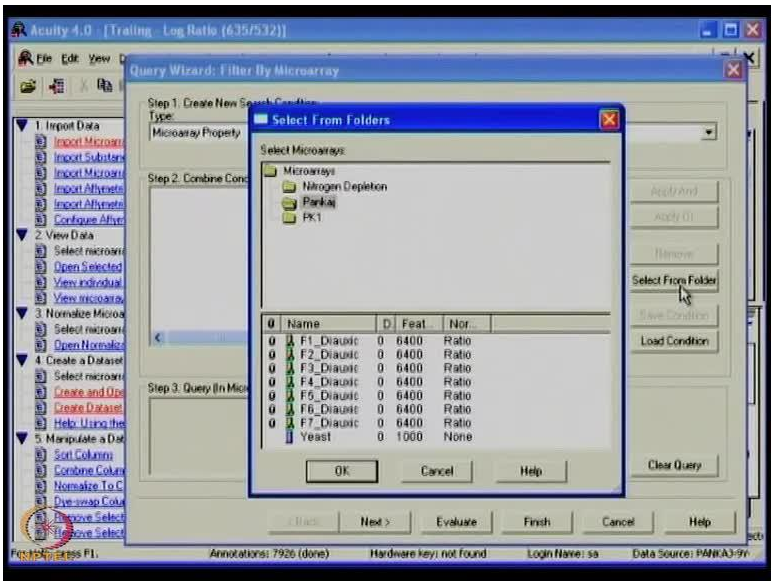



Acuity 4.0 - [Traling - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotations
Import Microarray Parameters
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature values
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query Wizard
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Remove Selected Columns
Datasets
Diauxic v N
(16) Diauxi
(7) Traling
Log Ratio (635/532
Mean
Substance
F1_Diauxic
F2_Diauxic
coli #1
coli #33
coli #35
coli #36
coli #40
0.185
0.393
Empty
Genomic 0.25
Genomic 0.5
Genomic 1
KAN-MX
Lambda
NORF 1
NORF 10
NORF 11
NORF 12
NORF 13
NORF 14
6214 substances (1 selected)
7 microarrays (1 select
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA3-9


Acuity 4.0 - [Traling - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotations
Import Microarray Parameters
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature values
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query Wizard
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Remove Selected Columns
Datasets
Diauxic v N
(16) Diauxi
(7) Traling
Log Ratio (635/532
Mean
Substance
F1_Diauxic
F2_Diauxic
coli #1
coli #33
coli #35
coli #36
coli #40
0.185
0.393
Empty
0.4
0
F1_Diauxic
F2_Diauxic
F3_Diauxic
F4_Diauxic
F5_Diauxic
F6_Diauxic
F7_Diauxic
Microarray
6214 substances (1 selected)
7 microarrays (1 select
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA3-9


Acuity 4.0 - [Trailing - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Create Dataset
Name:
Description:
Select location:
Datasets
Diauxic v Nitrogen Depletion
OK
Cancel
Help
1. Import Data
Import Microarrays
Import Substance Annotations
Import Microarray Parameters
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature values
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query Wizard
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Remove Selected Columns
Log Ratio (635/532)
Mean
F1_Diauxic
F2_Diauxic
0.185
0.393
Microarray
6214 substances (1 selected)
7 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA3-9W


Acuity 4.0 - [Trailing - Log Ratio (635/532)]
Query Wizard: Filter By Microarray
Step 1. Create New Search Condition:
Type:
Microarray
Parameter:
Name
Operation:
Value:
Add To List
Step 2. Combine Conditions:
Apply And
Apply Or
Remove
Select From Folder
Save Condition
Load Condition
Add To Query
Step 3. Query (In Microarrays Where) to be executed:
Clear Query
< Back
Next >
Evaluate
Finish
Cancel
Help
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA3-9W



Mr Pankaj Khanna: correct so if you see first was a import data then importing the substance annotation file which is dot sdt file which gives all the annotations which you have made tab delimited and name it to dot stt then the third one is microarray parameters where it is being stored in the form of (())(20:09) named to dot mdt file.

So which allows one to trace all the details. The other three you are seeing is based on the FA matrix which is actually a cell file. So there image is actually a dat file which they converted to intensity level cell file that can also be handle but just it does RNA analysis where they are moved little bit further for the different kind of analysis but it does give opportunity even to look at FA matrix outcome data.

And then you have a viewing data so we have seen you can auto feed data, we can look at the colour being coded, we can view the data in a different form. You can individually look them look at them or look with the numbers and comprise them and see all the arrays how the behaviour is happening, we can also do look at the profiling at the down based on what we have seen earlier, we can look at any of the things click on the profile and you will be able to see how this has behaved acrossed.

So apart from that next step involves the normalization result so I need to make sure that all my chips are being normalized similarly.

Prof Sanjeeva Srivastava: Right

Mr Pankaj Khanna: Or essentially when I import the data I prefer it non normalised and all I select and select one base of normalisation. I can create same level of experiment with different wave in different normalisation per se and do the further analysis down there. So once we have finished the normalisation method we can look at a query which is creating a data sets.

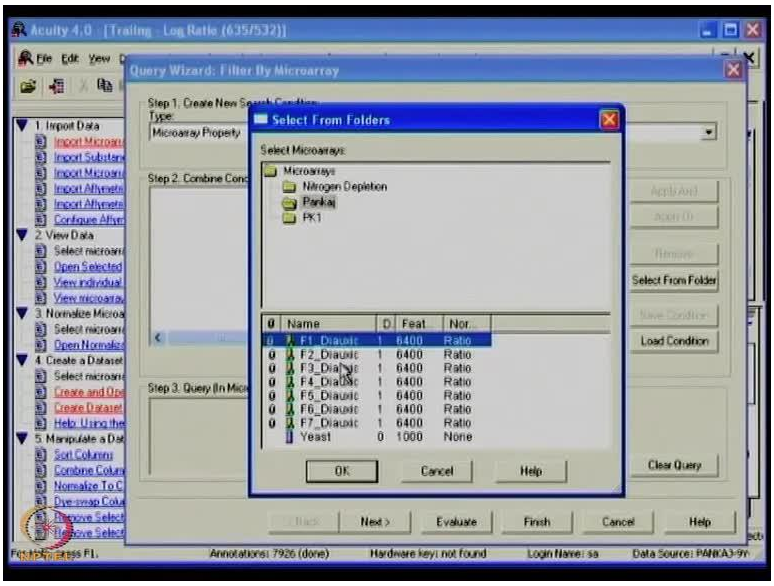
So as we described one simple way is right at the data, select the chips and then create a data or you can come here and say create at a data set from the open data set. So how to go for this one? If I click on create an open data sets right

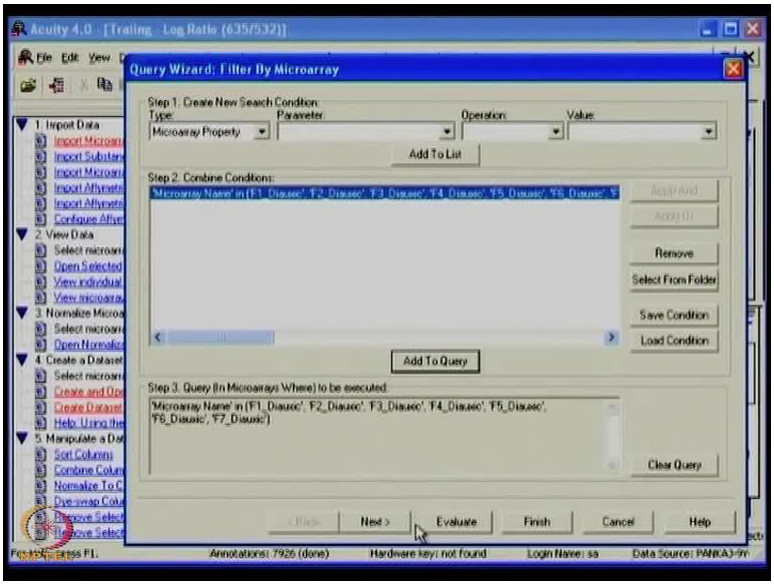
Prof Sanjeeva Srivastava:Ya.

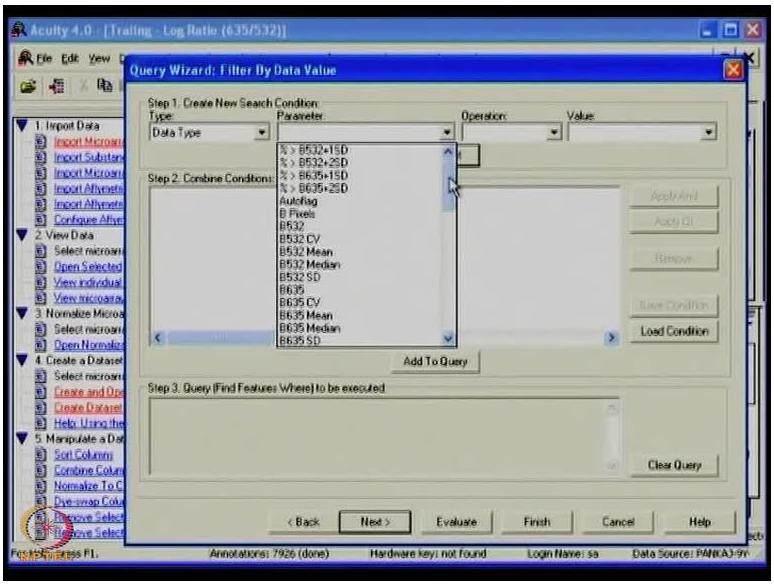
Mr Pankaj Khanna: So you can see similar way it has popped up the window where it shows in the dataset the (())(21:55) folder and then childs of it that what all can be generated now and I can create a data set here in this form which we can do directly there. Very other important factor is creating a data set from the query. The meaning is that you can define the different criterias right from the design of the experiment to little bit of more details of statistics to import what I want to import inside. Say for example quickly here I can define which mic experiment I want to do type but based on the parameters as such I can say that okay I want to import the data in the form of only one particular parameter. This is a little little tricky because if you have same name multiple times the data will be imported twice.

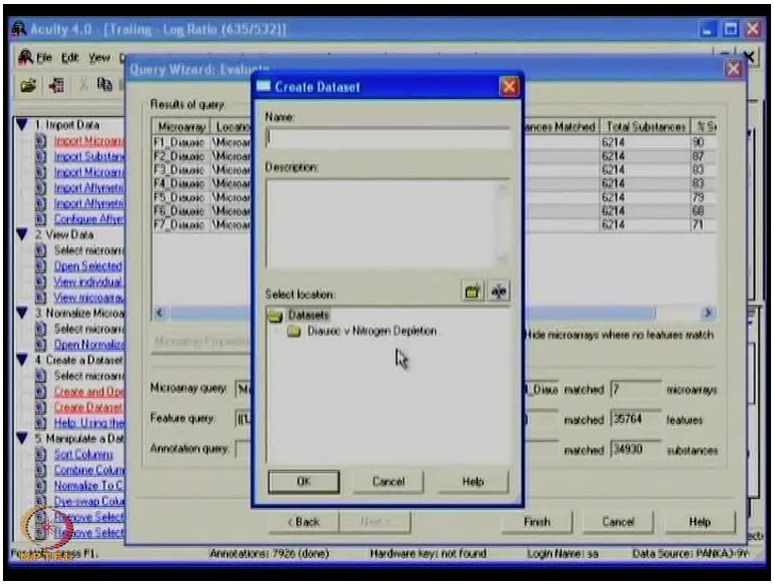
So you need to make sure your folder is right and then you have given only one copy.

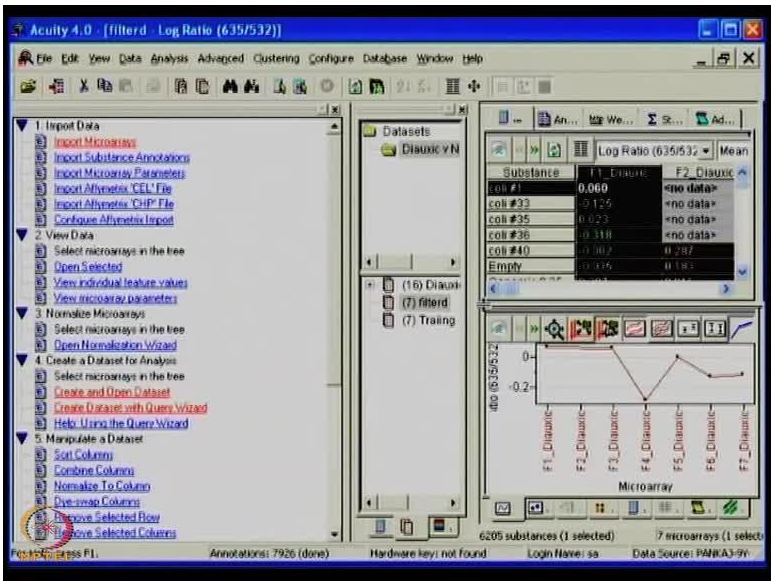
(Refer Slide Time: 22:53)

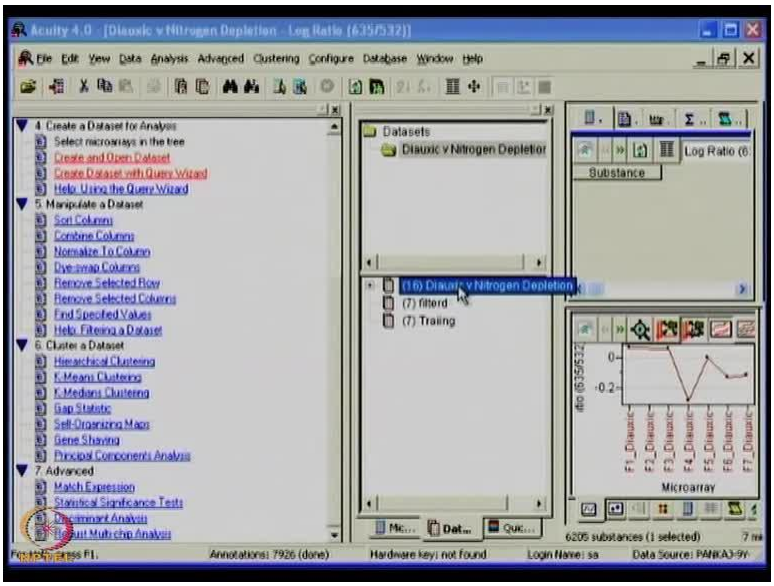

Acuity 4.0 - [Trailing - Log Ratio (635/532)]
Query Wizard: Filter By Microarray
Select From Folders
Select Microarrays:
Microarrays
Nitrogen Depletion
Pankaj
PK1
Name
D
Feat
Nor...
F1_Diauxic 1 6400 Ratio
F2_Diauxic 1 6400 Ratio
F3_Diauxic 1 6400 Ratio
F4_Diauxic 1 6400 Ratio
F5_Diauxic 1 6400 Ratio
F6_Diauxic 1 6400 Ratio
F7_Diauxic 1 6400 Ratio
Yeast 0 1000 None
OK
Cancel
Help
Step 1. Create New Search Condition:
Microarray Property
Step 2. Combine Conc
Step 3. Query (In Mic
Select From Folder
Load Condition
Clear Query
Next >
Evaluate
Finish
Cancel
Help
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa

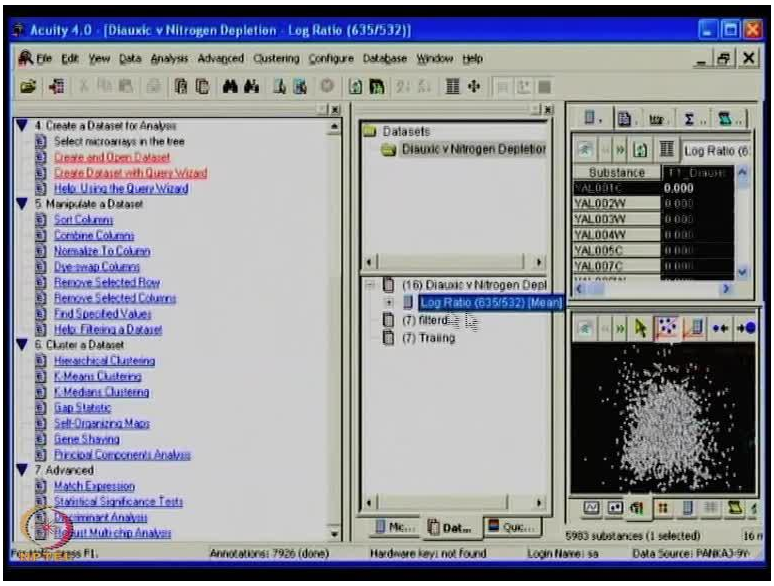

Acuity 4.0 - [Trailing - Log Ratio (635/532)]
Query Wizard: Filter By Microarray
Step 1. Create New Search Condition:
Type:
Parameter:
Operation:
Value:
Microarray Property
Add To List
Step 2. Combine Conditions:
'Microarray Name' in ('F1_Diauxic', 'F2_Diauxic', 'F3_Diauxic', 'F4_Diauxic', 'F5_Diauxic', 'F6_Diauxic', 'F7_Diauxic')
Remove
Select From Folder
Save Condition
Load Condition
Add To Query
Step 3. Query (In Microarrays Where) to be executed:
'Microarray Name' in ('F1_Diauxic', 'F2_Diauxic', 'F3_Diauxic', 'F4_Diauxic', 'F5_Diauxic', 'F6_Diauxic', 'F7_Diauxic')
Clear Query
Next >
Evaluate
Finish
Cancel
Help
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa

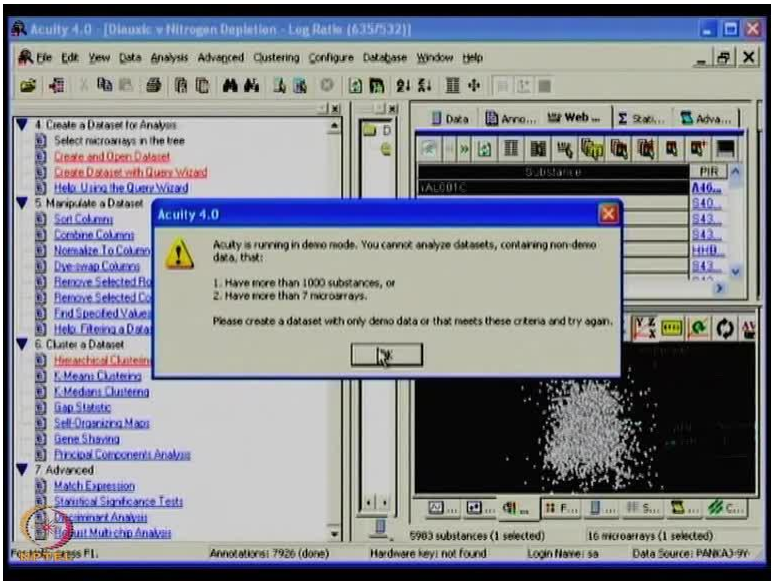

Acuity 4.0 - [Trailing - Log Ratio (635/532)]
File Edit View
Query Wizard: Filter By Data Value
Step 1. Create New Search Condition:
Type:
Data Type
Parameter:
Operation:
Value:
Step 2. Combine Conditions:
% > B532+1SD
% > B532+2SD
% > B635+1SD
% > B635+2SD
Autoflag
B Pixels
B532
B532 CV
B532 Mean
B532 Median
B532 SD
B635
B635 CV
B635 Mean
B635 Median
B635 SD
Apply And
Apply Or
Remove
Save Condition
Load Condition
Add To Query
Step 3. Query (Find Features Where) to be executed
Clear Query
< Back
Next >
Evaluate
Finish
Cancel
Help
1. Import Data
2. View Data
3. Normalize Microa
4. Create a Dataset
5. Manipulate a Dat
Sort Columns
Combine Colum
Normalize To C
Dye-swap Colu
Remove Select
Remove Select
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA):9n

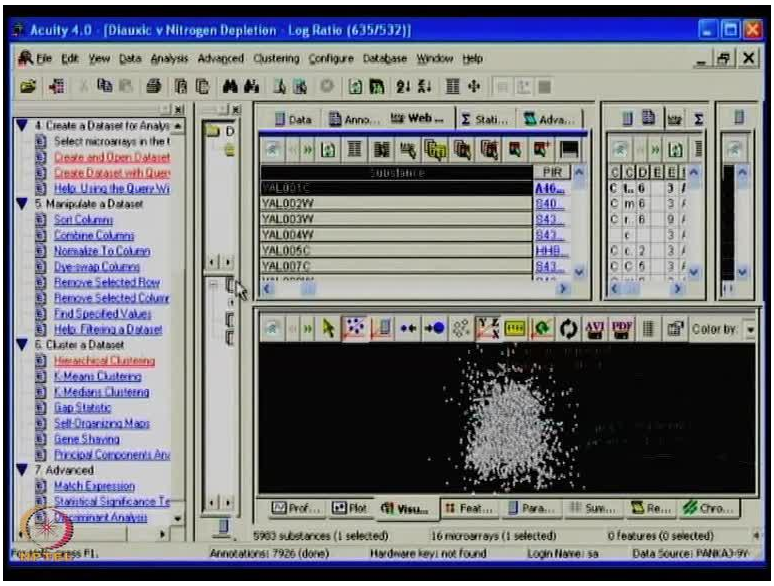

Acuity 4.0 - [Trailing - Log Ratio (635/532)]
File Edit View
Query Wizard: Evaluate
Create Dataset
Name:
Description:
Select location:
Datasets
Diauxic v Nitrogen Depletion
OK
Cancel
Help
Results of query:
Microarray
Total Substances
F1_Diauxic
F2_Diauxic
F3_Diauxic
F4_Diauxic
F5_Diauxic
F6_Diauxic
F7_Diauxic
6214
90
87
83
83
79
68
71
Hide microarrays where no features match
Microarray query:
matched 7
microarrays
Feature query:
matched 35764
features
Annotation query:
matched 34930
substances
< Back
Next >
Finish
Cancel
Help
1. Import Data
2. View Data
3. Normalize Microa
4. Create a Dataset
5. Manipulate a Dat
Sort Columns
Combine Colum
Normalize To C
Dye-swap Colu
Remove Select
Remove Select
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA):9n

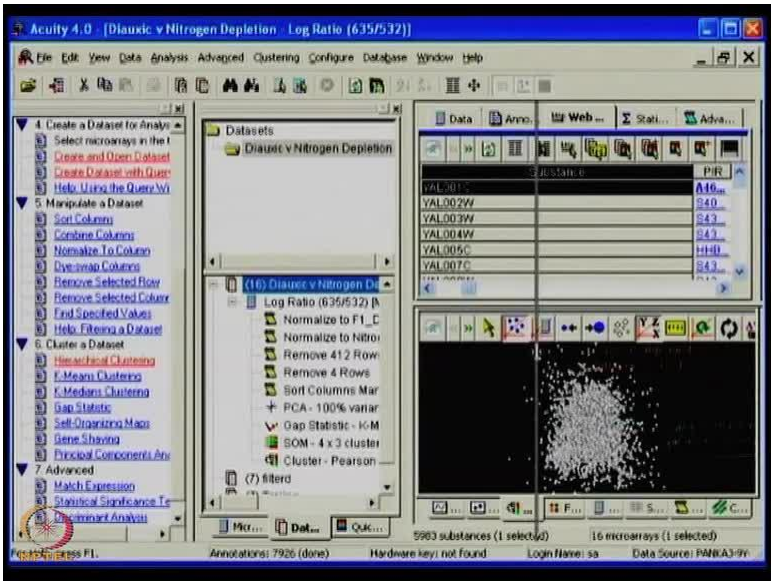

Acuity 4.0 - [filterd - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
1. Import Data
Import Microarrays
Import Substance Annotations
Import Microarray Parameters
Import Affymetrix 'CEL' File
Import Affymetrix 'CHP' File
Configure Affymetrix Import
2. View Data
Select microarrays in the tree
Open Selected
View individual feature values
View microarray parameters
3. Normalize Microarrays
Select microarrays in the tree
Open Normalization Wizard
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query Wizard
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Remove Selected Columns
Datasets
Diauxic v N
(16) Diauxi
(7) filterd
(7) Trailing
Log Ratio (635/532
Mean
Substance
F1_Diauxic
F2_Diauxic
coli #1
0.060
<no data>
coli #33
<no data>
coli #35
<no data>
coli #36
<no data>
coli #40
0.287
Empty
0
-0.2
F1_Diauxic
F2_Diauxic
F3_Diauxic
F4_Diauxic
F5_Diauxic
F6_Diauxic
F7_Diauxic
Microarray
6205 substances (1 selected)
7 microarrays (1 select
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-9Y

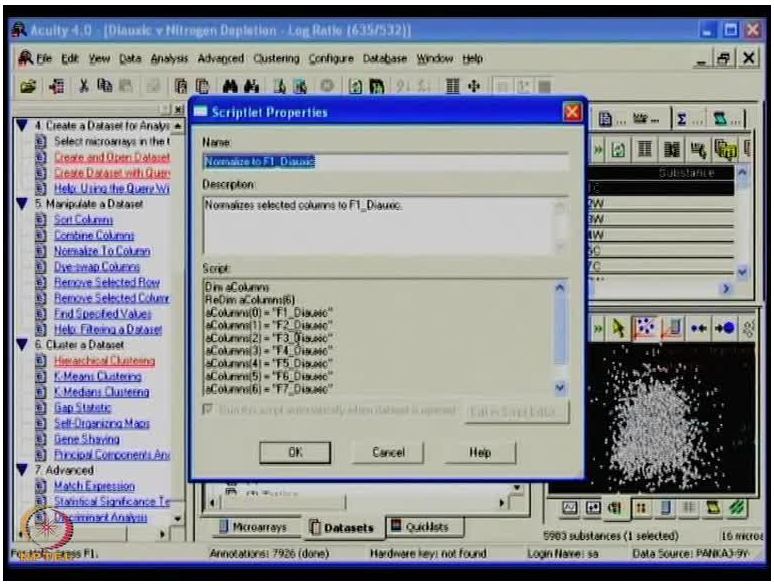

Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query Wizard
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Remove Selected Columns
Find Specified Values
Help: Filtering a Dataset
6. Cluster a Dataset
Hierarchical Clustering
K-Means Clustering
K-Medians Clustering
Gap Statistic
Self-Organizing Maps
Gene Shaving
Principal Components Analysis
7. Advanced
Match Expression
Statistical Significance Tests
Discriminant Analysis
Export Multichip Analysis
Datasets
Diauxic v Nitrogen Depletion
(16) Diauxic v Nitrogen Depletion
(7) filterd
(7) Trailing
Log Ratio (6
Substance
0
-0.2
F1_Diauxic
F2_Diauxic
F3_Diauxic
F4_Diauxic
F5_Diauxic
F6_Diauxic
F7_Diauxic
Microarray
6205 substances (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKAJ-9Y

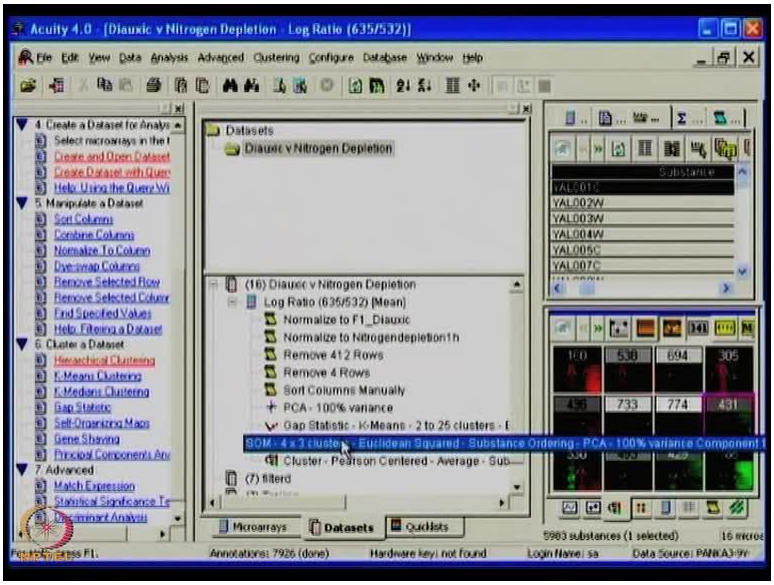

Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
4. Create a Dataset for Analysis
Select microarrays in the tree
Create and Open Dataset
Create Dataset with Query Wizard
Help: Using the Query Wizard
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Remove Selected Columns
Find Specified Values
Help: Filtering a Dataset
6. Cluster a Dataset
Hierarchical Clustering
K-Means Clustering
K-Medians Clustering
Gap Statistic
Self-Organizing Maps
Gene Shaving
Principal Components Analysis
7. Advanced
Match Expression
Statistical Significance Tests
Discriminant Analysis
Input Multi-chip Analysis
Datasets
Diauxic v Nitrogen Depletion
(16) Diauxic v Nitrogen Depl
Log Ratio (635/532) [Mean]
(7) filtered
(7) Trailing
Substance
YAL001C
YAL002W
YAL003W
YAL004W
YAL005C
YAL007C
5983 substances (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA3-9V

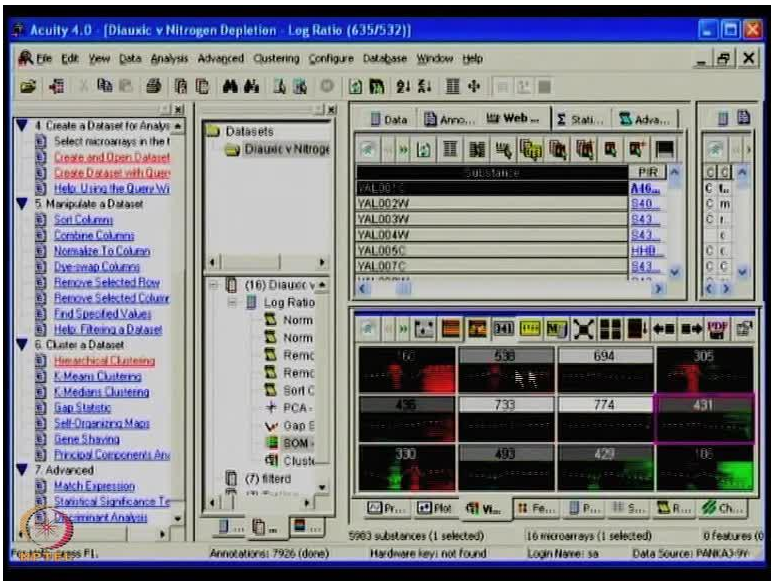

Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Acuity 4.0
Acuity is running in demo mode. You cannot analyze datasets, containing non-demo data, that:
1. Have more than 1000 substances, or
2. Have more than 7 microarrays.
Please create a dataset with only demo data or that meets these criteria and try again.
5983 substances (1 selected)
16 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa

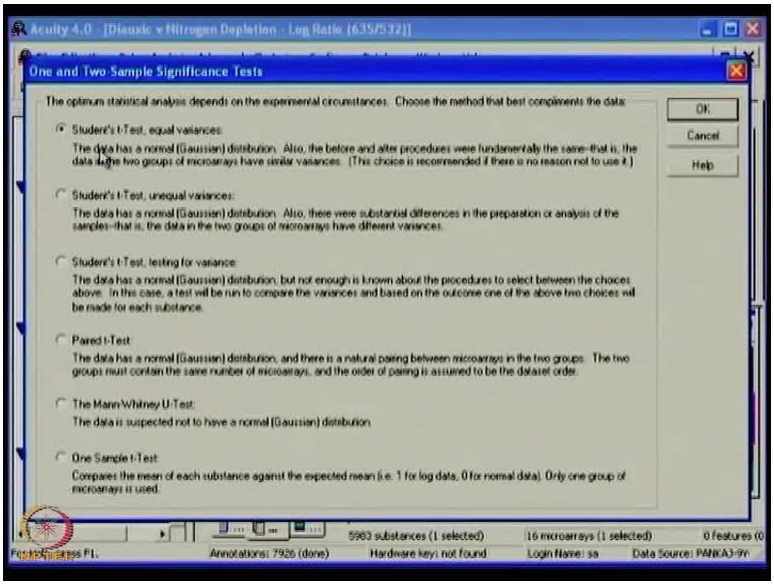

Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
4. Create a Dataset for Analysis
Select microarrays in the t
Create and Open Dataset
Create Dataset with Quer
Help: Using the Query Wi
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Remove Selected Column
Find Specified Values
Help: Filtering a Dataset
6. Cluster a Dataset
Hierarchical Clustering
K-Means Clustering
K-Medians Clustering
Gap Statistic
Self-Organizing Maps
Gene Shaving
Principal Components Ana
7. Advanced
Match Expression
Statistical Significance Te
Discriminant Analysis
Data
Anno...
Web ...
Stati...
Adva...
Substance
PIR
YAL001C
A46...
YAL002W
S40...
YAL003W
S43...
YAL004W
S43...
YAL005C
HH8...
YAL007C
S43...
Color by:
Prof...
Plot
Visu...
Feat...
Para...
Sum...
Re...
Chro...
5983 substances (1 selected)
16 microarrays (1 selected)
0 features (0 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA3-9V


Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
4. Create a Dataset for Analysis
Select microarrays in the t
Create and Open Dataset
Create Dataset with Quer
Help: Using the Query Wi
5. Manipulate a Dataset
Sort Columns
Combine Columns
Normalize To Column
Dye-swap Columns
Remove Selected Row
Remove Selected Column
Find Specified Values
Help: Filtering a Dataset
6. Cluster a Dataset
Hierarchical Clustering
K-Means Clustering
K-Medians Clustering
Gap Statistic
Self-Organizing Maps
Gene Shaving
Principal Components Ana
7. Advanced
Match Expression
Statistical Significance Te
Discriminant Analysis
Datasets
Diauxic v Nitrogen Depletion
(16) Diauxic v Nitrogen De
Log Ratio (635/532) [N
Normalize to F1_C
Normalize to Nitro
Remove 412 Row
Remove 4 Rows
Sort Columns Mar
PCA - 100% variar
Gap Statistic - K-M
SOM - 4 x 3 cluster
Cluster - Pearson
(7) filterd
Mic...
Dat...
Quic...
Data
Anno...
Web ...
Stati...
Adva...
Substance
PIR
YAL001C
A46...
YAL002W
S40...
YAL003W
S43...
YAL004W
S43...
YAL005C
HH8...
YAL007C
S43...
5983 substances (1 selected)
16 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa
Data Source: PANKA3-9V


Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Scriptlet Properties
Name:
Normalize to F1_Diauxic
Description:
Normalizes selected columns to F1_Diauxic.
Script:
Dim aColumns
ReDim aColumns(6)
aColumns(0) = "F1_Diauxic"
aColumns(1) = "F2_Diauxic"
aColumns(2) = "F3_Diauxic"
aColumns(3) = "F4_Diauxic"
aColumns(4) = "F5_Diauxic"
aColumns(5) = "F6_Diauxic"
aColumns(6) = "F7_Diauxic"
OK
Cancel
Help
Microarrays
Datasets
Quicklists
5983 substances (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa


Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Datasets
Diauxic v Nitrogen Depletion
(16) Diauxic v Nitrogen Depletion
Log Ratio (635/532) [Mean]
Normalize to F1_Diauxic
Normalize to Nitrogendepletion1h
Remove 412 Rows
Remove 4 Rows
Sort Columns Manually
PCA - 100% variance
Cluster - Pearson Centered - Average - Sub...
(7) filterd
Microarrays
Datasets
Quicklists
5983 substances (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa





To avoid that so I can select the folder and then I can say I am into this particular folder and I want to import all the data from that. So what happens is it selects again in the same shift fashion it creates query we can add and add query and it can be created in this fashion. So once you have done you want say which parameter you want to select.

Prof Sanjeeva Srivastava: right.

Mr Pankaj Khanna: I can just quickly do this for you that I want to take a ratio based one. I will just take log ratios and I can create a parameter of less than or equal to 0.3. 0.23 is log ratio changes something like 2 add to list and greater than 2 greater than say for example 2.

Prof Sanjeeva Srivastava: ya.

Mr Pankaj Khanna: Again I can add query so if I select both of them I can create an OR or AND, if I create OR we can apply that any feature which is this or that you select that for my import so I will able to just add as query on these to the final result and then you have the data available for you coming out. You can even select the database basis of annotation and filter only with are (())(24:14) specific, it depends on what questions you are ask is a quickly look at only that what you have say.

So it says how many of them are there in that so it reduce the numbers. So in this fashion I can import a limited number of database as well. Right? I can give a name to it like filter and it will be imported.

Prof Sanjeeva Srivastava: okay.

Mr Pankaj Khanna: So you cannot create files on the root directory you need to create your folder and do the job. So it can do job in this fashion. So once you are ready with your dataset what you want to do sometimes is based on your experiment it gives you an opportunity what all you can do.

Number one you can sort the column. Meaning F1 to F 7, I want to make F7 first. I can use a sorting by that and so also I can do a different ways of combining columns. One why you need to combine columns? Say for example I have given three technical replicates for each and then biological replicates of three. So I can just take an average of them, combine the columns and it will take an average of all of this and you are ready with data to go ahead for the analysis part.

So many a times you can also go for normalise to column the meaning I am having a zero to maximum. So I want all my data to get normalised to first column. This feature is used when you are essentially using point so all the standardize or normalised to one column which you have

define as a 0^{th} one. And very important one other would is dye swap usually we know that si3 and si5.

Si3 is little less in size, si5 is little bigger. So there is a variation in the incorporation of that. To take care usually people do at least one dye swap experiment to accommodate the variations happening because of the dye. So what you can do is you can apply a dye swap, I can quickly show how the dye the swap works. Dye swap just changes the way you look at. So you have taken ratios of one by another wavelength so it just reverse as it is.

so it is just - X. So - will become +, + will -. So it just changes the dye swap which will take care in the form of combining the data and avoiding the variability happening due to the dye swap affect.

Prof Sanjeeva Srivastava: you also talked about dye swap when I was talking to them about dye technology and how one need to use in fact labelling with different dyee and reverse dye swapping so that there is no dye bias in the analysis.

Mr Pankaj Khanna: True. So same can be accommodated here at the microarray to look at that is very nice (())(26:42) actually to go ahead with and then many times which you find at the few rows I want to remove I can remove few rows, I can select few of the columns, I can remove say example some QC has not passed. So I can just remove that which can allow me to do a different ways. And then once you have done that you can go and do clustering which visualisation method.

Prof Sanjeeva Srivastava: Correct.

Mr Pankaj Khanna: so technically the clustering is divided into two types, hierarchical and non-hierarchical. Hierarchical means that in starting that only one start point and then all other features are attached to them, other one is non-hierarchical type where each grouped behaves independently of each other. So they are K means, SOM, K medians and the particularly one the people use hierarchical when they do not know to where to start.

Prof Sanjeeva Srivastava: Right.

Mr Pankaj Khanna: So, they start with the hierarchical when they do not know how many groups could happen and how many results I am expecting and once you do have the results and idea you can cut down because hierarchical is little ram consuming it takes little bit of more to time.

Prof Sanjeeva Srivastava: Right.

Mr Pankaj Khanna: whereas beause you can imagine all has to be linked to one. And there are different ways of doing it central based the based like coefficient correlation based or distance matrix based and lastly binary based.

Prof Sanjeeva Srivastava: Ok.

Mr Pankaj Khanna: so binary based is usually used for only CGS kind of analysis where it is present absent type whereas the earlier two ones are extensively used in the microarray data.

The pearson's correlation with centric ones are being used for the median or the mean type variations. So, when you quickly do it we will able to see what we can do. Say for example I want to do all my filtered data which have already done one of the clusterings. Say I have got the filtering I have got the log ratios and another important thing is whichever is something like bluish in colour that particular one you have selected to work with.

Prof Sanjeeva Srivastava: Okay so refilter data.

Mr Pankaj Khanna: Ya so this is actually a filtered data and allowed to work with. And when I hit uhh kind of clustering possibility it says you need to create a quick data set.

Prof Sanjeeva Srivastava: okay

Mr Pankaj Khanna: The meaning of that one is you have half the data already in hand but now what you need to do is look at a third tab which allows one to look in the form which you want to create the data set for which you want to analyse going little lengthier right?

Prof Sanjeeva Srivastava: ya.

Mr Pankaj Khanna: So, I have a normalize the data and normalize the data to f1 so I can see the script what is happening. So I can just look at the one which I have performed at the level of

different processing so that I can look at the data. So, here when I look at SOM which I have some which I have done at level of 4 into 3 clusters, SOMs is something like nonhierarchical type so they are individually being blocked. So all different genes behave differently and based on the profile they are made into one group.

So because I have given 4 into 3 I will be able to see 4 into 3. So 4 number columns and 3 rows. So independent has been divided so, one can have any numbers the idea come from the hierarchical.

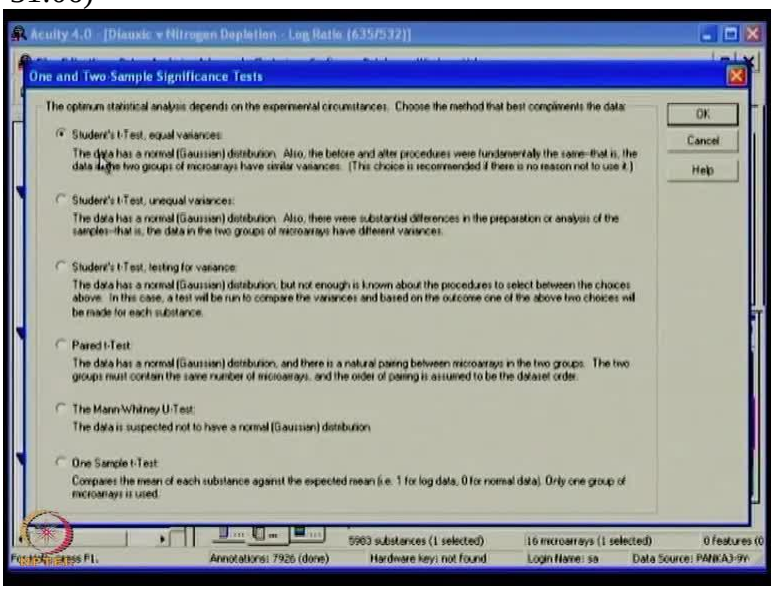
But as you see here immediately the clustering gives an immediate response that ok this was very low in these ones, again it went down and it you are able to see it again up. Prof Sanjeeva Srivastava: and different time (())(30:21) different expression

Mr Pankaj Khanna: correct so based on the behaviors you can see that okay this is being grouped up with. So there are different ways which people prefer but usually preferred ones are the (())(30:32)ones for the hierarchical type after the Pearson's correlation and further this one after SOMs people usually use (())(30:41) square using some of the nonhierarchical type. So k-means and k-medians is the better ones to use with.

So this actually gives you an opportunity how the data is being visualized. So once you have visualized the data so what you do is you can look at a statistics.

Prof Sanjeeva Srivastava: Right

(Refer Slide Time: 31:06)



Mr Pankaj Khanna: Statistics in the sense if you click on any of the ones which is something like statistical significant test it tells you I have different option go ahead with.

Prof Sanjeeva Srivastava: Different type of test.

Mr Pankaj Khanna: Different type of test one can perform like student t test and the reason is when it is equal variance with Gaussian normal well shape curved you prefer this test. When the variance is not equal you want to go with the second type where there is a small modification of the again student t test but without normal variance happening.

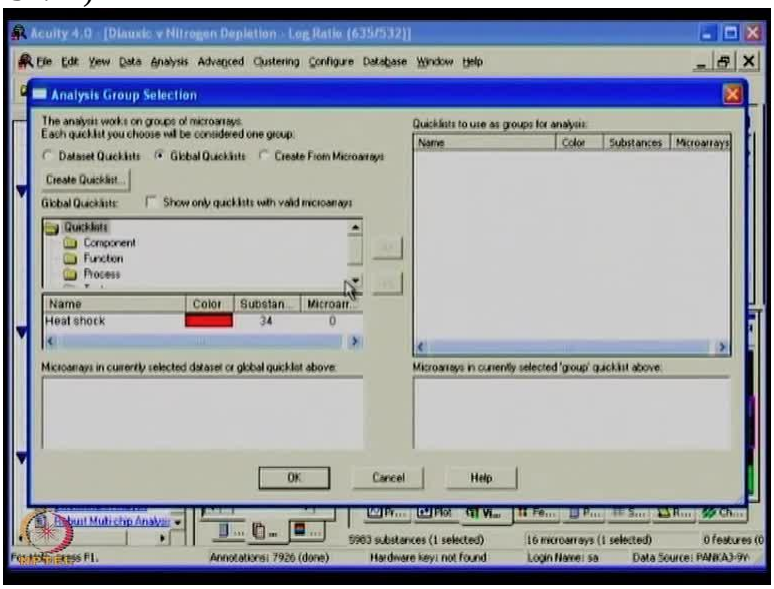
And there are student t tests where you are do not know what actually variance means to and you prefer paired test when the sample is coming from the same origin specially in the form of cancers where the when the cancer is being removed surgically. People remove some normal samples so that will be

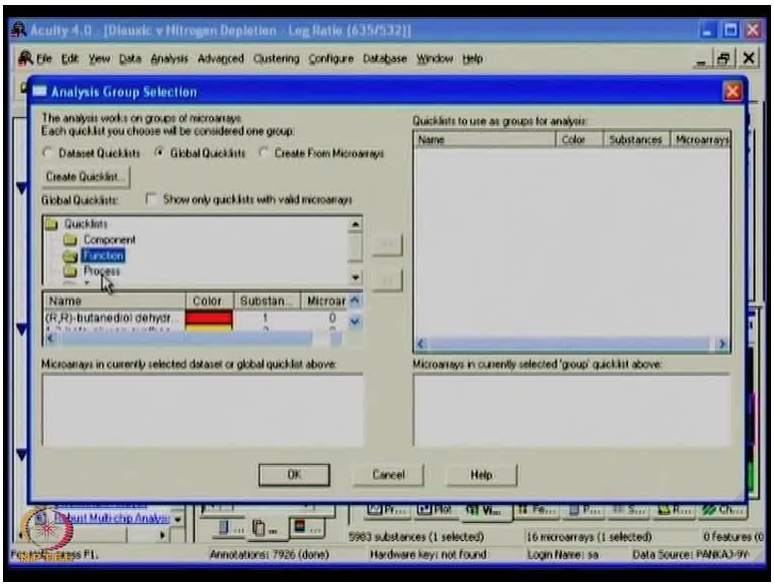
Prof Sanjeeva Srivastava: same…same

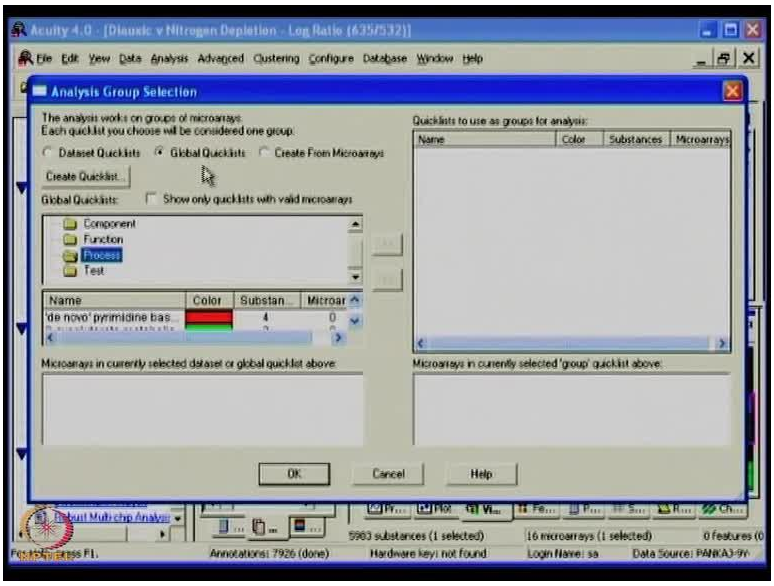
Mr Pankaj Khanna: ya same biological patients is the origin of that. So this helps this particular one will allow one to select based on what background you have and the other one is mann whitney test which actually people use where there is no normal Gaussian say for example few times it is only standalone ones so up and down. that is it.

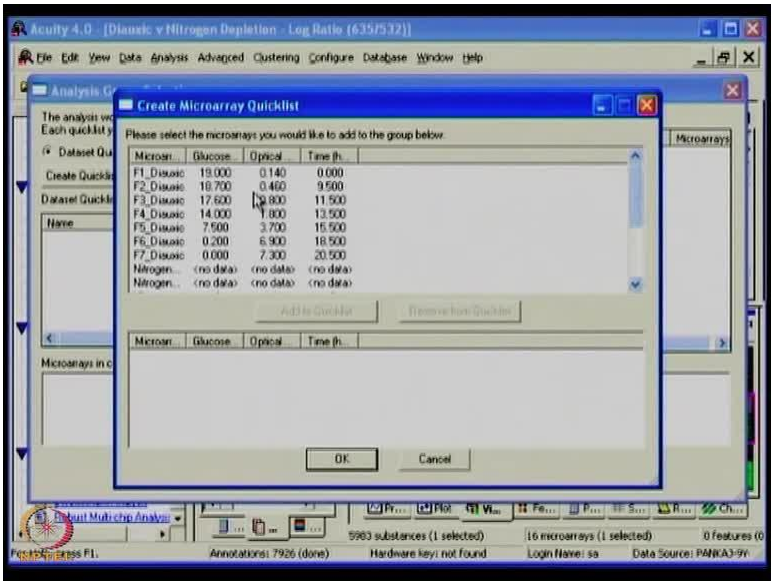
So, you want select this parameter for that. So based on right kind of design one will select the different kind of statistics available and you perform the analysis data.

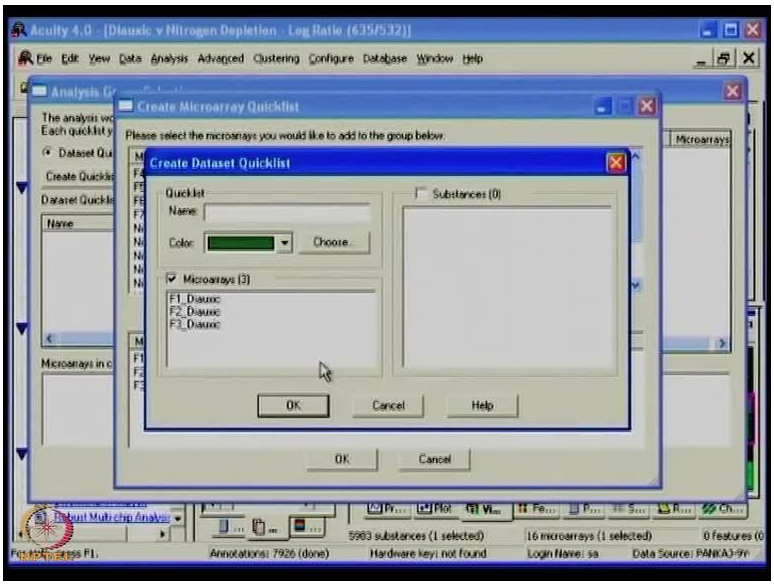
(Refer Slide Time: 32:21)

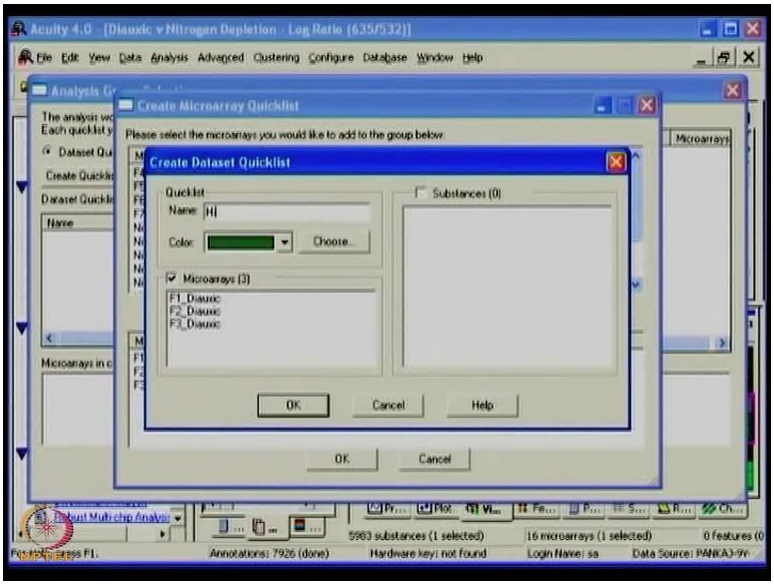


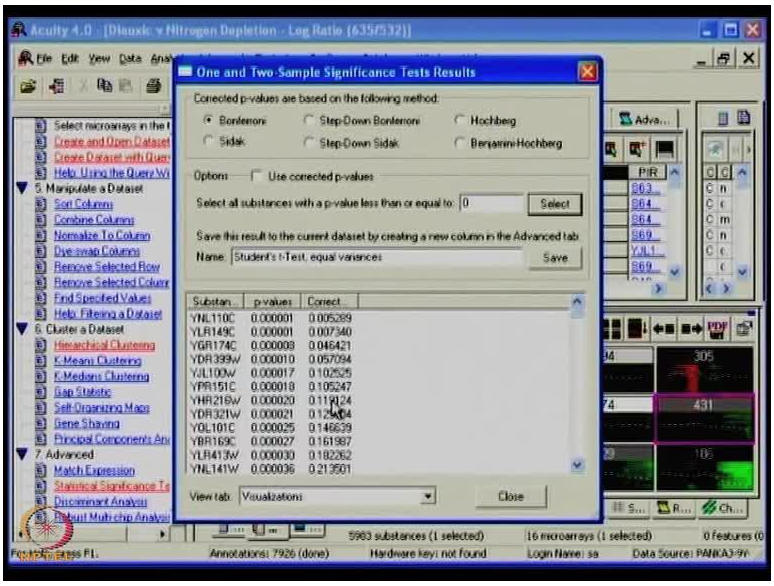


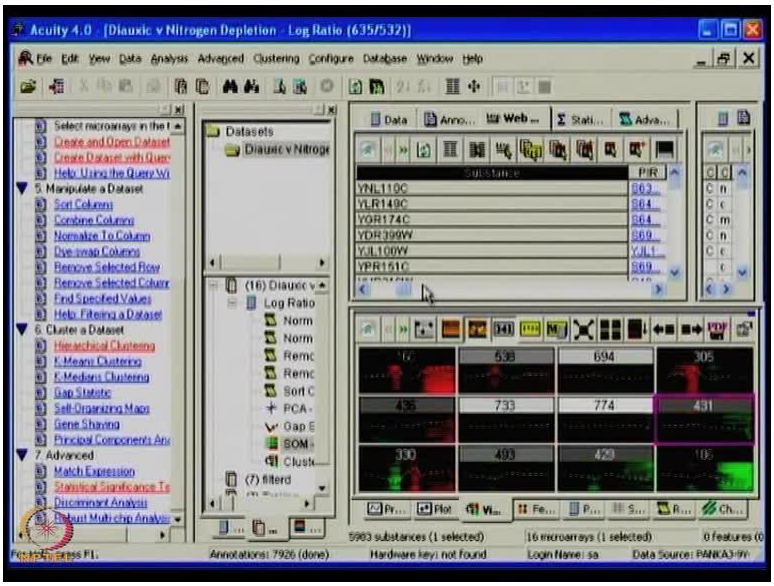

Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Analysis Group Selection
The analysis works on groups of microarrays.
Each quicklist you choose will be considered one group.
Dataset Quicklists
Global Quicklists
Create From Microarrays
Create Quicklist...
Global Quicklists:
Show only quicklists with valid microarrays
Component
Function
Process
Test
Name
Color
Substan...
Microar
'de novo' pyrimidine bas...
4
0
Microarrays in currently selected dataset or global quicklist above:
Quicklists to use as groups for analysis:
Name
Color
Substances
Microarrays
Microarrays in currently selected 'group' quicklist above:
OK
Cancel
Help
5983 substances (1 selected)
16 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa


Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Create Microarray Quicklist
Please select the microarrays you would like to add to the group below.
Microarr...
Glucose
Optical...
Time (h...
F1_Diauxic 19.000 0.140 0.000
F2_Diauxic 18.700 0.460 9.500
F3_Diauxic 17.600 2.800 11.500
F4_Diauxic 14.000 1.800 13.500
F5_Diauxic 7.500 3.700 16.500
F6_Diauxic 0.200 6.900 18.500
F7_Diauxic 0.000 7.300 20.500
Nitrogen... (no data) (no data) (no data)
Nitrogen... (no data) (no data) (no data)
Add to Quicklist
Remove from Quicklist
OK
Cancel
5983 substances (1 selected)
16 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa


Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Create Microarray Quicklist
Please select the microarrays you would like to add to the group below.
Create Dataset Quicklist
Quicklist
Name:
Color:
Choose...
Substances (0)
Microarrays (3)
F1_Diauxic
F2_Diauxic
F3_Diauxic
OK
Cancel
Help
OK
Cancel
Microarrays
5983 substances (1 selected)
16 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa


Acuity 4.0 - [Diauxic v Nitrogen Depletion - Log Ratio (635/532)]
File Edit View Data Analysis Advanced Clustering Configure Database Window Help
Create Microarray Quicklist
Please select the microarrays you would like to add to the group below.
Create Dataset Quicklist
Quicklist
Name: Hi
Color:
Choose...
Substances (0)
Microarrays (3)
F1_Diauxic
F2_Diauxic
F3_Diauxic
OK
Cancel
Help
OK
Cancel
Microarrays
5983 substances (1 selected)
16 microarrays (1 selected)
Annotations: 7926 (done)
Hardware key: not found
Login Name: sa





Prof Sanjeeva Srivastava: So one need to look at their experimental design and then select the right statistical p arameter for further analysis.

Mr Pankaj Khanna: Correct. So, in this fashion what it does is it tells you that okay I have got different kind of data sets which you want to select with, so I select them and I say group them according to like 0 and 1s, like all cases and all controls and then I perform the analysis.

So, let us quickly do that. This data is based on something like different kind of components, functions. So I want to understand the differences between different functions or I can quickly

say you can quickly create a data set. So I can create data set say which one you want to create with. So I said okay create a data set from all these.

Prof Sanjeeva Srivastava: Different time points we can (())(33:14)

Mr Pankaj Khanna: So, (base) basically I know that it is high glucose at the beginning. So I can group them all as very high.

Prof Sanjeeva Srivastava: Right

Mr Pankaj Khanna: Because this is a time point although it is decreasing more likely I will select only once with only nineteen and then compare with very low type and I want to see how they are getting different shapes

Prof Sanjeeva Srivastava: overall changes.

Mr Pankaj Khanna: Overall changes is happening in the low and high level change, I have to quick list data and it is available to me and the name I give is high sugar. So it is available for me now. And I can also create a one with low sugar and I am going to say I am going to compare these two groups.

I can differentially colour them and I can see how the things are happening. Hit ok and you are ready with the different things. Other important thing is for the multiple testing corrections. It also gives an opportunity of correcting for the different multiple correction type. Say for example bonferoni which is being used Hochberg, Benjamin Hochberg so they are like different ways correcting for the multiple test corrections. So you can apply different ones and look more preferably is bonferoni which is most stringent type, Hochberg and Benjamin is little lenient on the multiple correction type.

So, you select them and see still there is huge number of things which is getting significant pre value changes. So, these genes become really important for me. I can say I can create the image store them and look back what these genes are, what those functions are. So in this fashion one can visualize, look at the differential genes with different kind of statistics, with different kind multiple corrections and look at the data up. So in this fashion acuity help in understanding the data analysis.

Prof Sanjeeva Srivastava: okay so very nice to see that there are so many parameters and options we have here for QCingthe data and analyzing it further for obtaining some meaningful information I guess there is no end to doing all of this analysis till one really feels confident about that whole process has performed well.

So I think I will finish here uhh so thank you for giving a very useful demonstration on this software and at least giving a glimpse of the entire workflow, how different type of process are involved, I am sure there is lot more can be explained and lot more can be done here but just due to the overall time and this lecture I think we should finish here, at least the student have got the glimpse of the overall process involved in the analysis and how one can look at a stringent way, different type of statistical test one need to perform and then different type of filtering can be done to obtain a different type of four chain and different ratios one need to obtain and further one need to look at the trend for each of those and which can be colour coded and presented in different ways.

So thank you very much Pankaj for being here and giving a very useful demonstration on acuity software for microarray data analysis.

Mr Pankaj Khanna: Thank you Dr Srivastava.

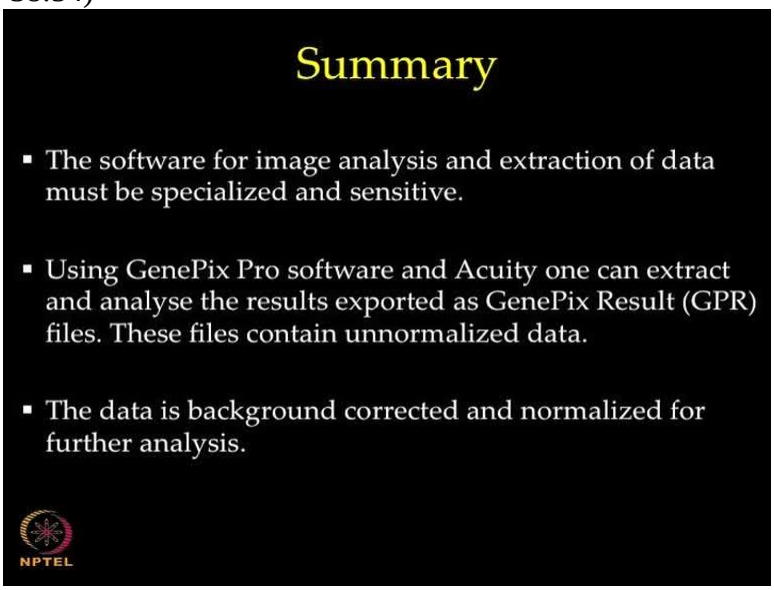
Prof Sanjeeva Srivastava: It was very informative to discuss various parameters and options which have been placed for QCing the data and analyzing it further for obtaining some meaningful information. There is lot more that can be performed here and there is possibly no end to doing the data analysis till one really feels confident about the whole process has been performed well.

But due to the time constraint it might be right to conclude this lecture here. I hope you have got the glimpse of the process involved in data analysis and how one can look into other data in much more stringent) manner, perform various types of filtering to generate appropriate threshold. One needs further look into the trend of each of these features which can be colour coded and presented in a different manner. Data analysis not only requires a good software platform but also requires good programming and statistical skills.

Experimental design is of paramount importantce in these microarray based experiments. !It is important to note thatth is software tool can help in analysis but it is also important to have a good understanding of both the biology as well as analytical techniques involved in performing these experiments. Rather than completely relying on one software we should also think about the biological contexts, look at the control and then after careful biological as well as analytical analysis you can obtain some meaningful information from these datasets. So, microarray experiment generate high throughput data in a short time but it becomes very challenging to analyse such datasets especially when you have to compare various slides from different experiments.

We have to normalize them equally so that you can compare all the slides on same platform. This careful image processing and data analysis becomes crucial in microarray based experiments which has been covered in detail now. Thank you.

(Refer Slide Time: 38:54)



Summary

- Clustering algorithms like hierarchical and non-hierarchial clustering can be employed to visualize data.
- Statistical tests like student's t-test followed by correction models like FDR, Benjamini Hochberg or Bonferroni can be applied to improve stringency of finding differentially regulated proteins.
- Adequate QC checking of data at every step is essential to maintain quality and accuracy of data

References

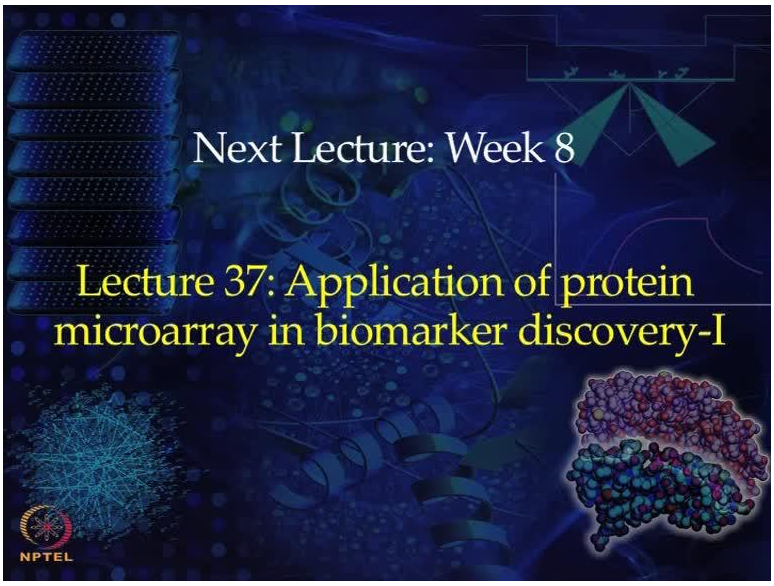
- Anna Pursiheimo, L.T.A. (2015). Statistical Analysis of Protein Microarray Data: A Case Study in Type 1 Diabetes Research. J. Proteomics Bioinform. 8.
- Bengtsson, A., and Bengtsson, H. (2006). Microarray image analysis: background estimation using quantile and morphological filters. BMC Bioinformatics 7, 96.
- Chandra, H., Reddy, P.J., and Srivastava, S. (2011). Protein microarrays and novel detection platforms. Expert Rev. Proteomics 8, 61–79.

References

- DeLuca, D.S., Marina, O., Ray, S., Zhang, G.L., Wu, C.J., and Brusic, V. (2011). Data processing and analysis for protein microarrays. Methods Mol. Biol. Clifton NJ 723, 337–347.
- Gupta, S., Manubhai, K.P., Kulkarni, V., and Srivastava, S. (2016). An overview of innovations and industrial solutions in Protein Microarray Technology. Proteomics 16, 1297–1308.
- Rydén, P., Andersson, H., Landfors, M., Näslund, L., Hartmanová, B., Noppa, L., and Sjöstedt, A. (2006). Evaluation of microarray data normalization procedures using spike-in experiments. BMC Bioinformatics 7, 300.

References

- Syed, P., Gupta, S., Choudhary, S., Pandala, N.G., Atak, A., Richharia, A., K P, M., Zhu, H., Epari, S., Noronha, S.B., et al. (2015). Autoantibody Profiling of Glioma Serum Samples to Identify Biomarkers Using Human Proteome Arrays. Sci. Rep. 5, 13895.
- Wellhausen, R., Seitz, H., Wellhausen, R., and Seitz, H. (2012). Facing Current Quantification Challenges in Protein Microarrays, Facing Current Quantification Challenges in Protein Microarrays. BioMed Res. Int. BioMed Res. Int. 2012, 2012, e831347.


Next Lecture: Week 8
Lecture 37: Application of protein microarray in biomarker discovery-I
NPTEL