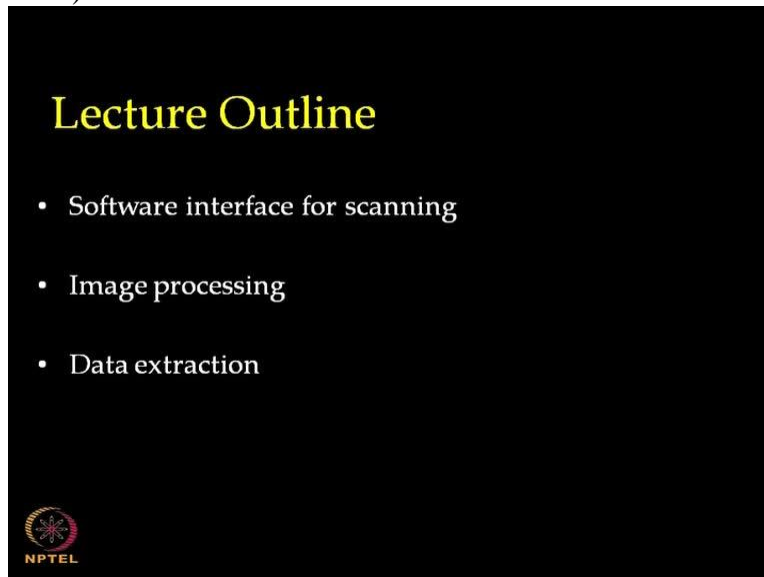


Interactomics: Protein Arrays & Label Free Biosensors
Professor Sanjeeva Srivastava
MOOC NPTEL Course
Indian Institute of Technology Bombay
Module 7
Lecture No 34
Software for Image scanning and data processing

In our previous lecture, we understood the vital parameters to be taken into consideration before data acquisition and scanning.

(Refer Slide Time: 0:28)

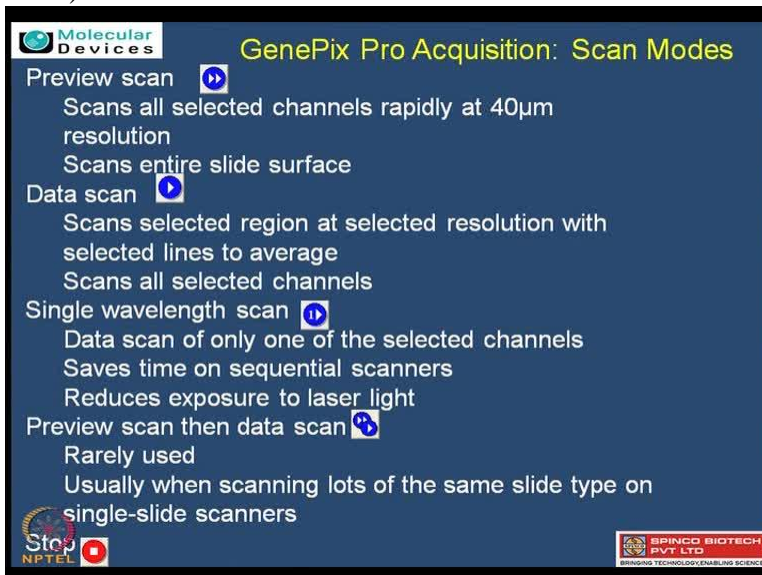


Today we will discuss about the software interface for scanning, image processing and data extraction. Mr. Pankaj Khanna would provide software demonstration for performing the scanning but before you move on to the live demo session on a scanning we can briefly discuss the software in general and how the user interface looks like.

Prof Sanjeeva Srivastava: So, welcome again to this lecture Pankaj. So, it will be useful if we can actually see the how the software works for performing the scanning but before we move on to the live demo for the scanning performance may be we can just briefly discuss about the software in general and how the user interface looks like. If you can just elaborate on some of the features while one need to look for during the scanning process I think that is useful for us to understand when we actually look at the demonstration for the software.

Mr. Pankaj Khanna: True. So the basic GUI which we are going to see in few minutes actually contains three different areas. The first one is image pleasure controls.

(Refer Slide Time: 1:46)



So, there you wanted to see which kind of laser I am going to use and second one is a different feature which is used for the controlling of the image and towards the right hand side we have a pane which is allowed or helpful in hardware controls. So, the basic one is in the hardware control is first as we said that we can look at auto PMT and other adjustment which is being done, for that we used preview scan.

So, preview scan and there are different tabs which allow you to go for the true scan and the preview scan and based on the laser which particular scan you are using. So, like for example if you start with the preview scan you decide on which best pixelation suits me, which different power of laser suits me. Once you are able to do these decisions being made you can go for your own data scan.

People prefer that because they are sometimes a bleaching effect on the () (2:37) so, they want to avoid the bleaching for a longer time. So, multiple scanning is avoided. So, once you see even a single wavelength scan so if your application requires only one wavelength to be scanned there you can select in the tab only a single wavelength scan and then so also a preview scan with the data scan forward by.

You can also do automation that once you do a preview scan you see that it is all good then you immediately it can go for that. But very rarely people use that because you see once and then only you want to go ahead.

Prof. Sanjeeva Srivastava : So I its like any general scanning even if you want to just scan sheets right? First of all you would like to preview it that how you you just want to get glimpse of the process like how the overall image looks like and then since you do your experiment you know your requirements like what wavelength need to be used, what type of (())(3:24) you have used, then one need to optimize and correct those things during the data scan which will actually scan performed and then after that one need review that whole thing and then how the slide looks like.

Mr. Pankaj Khanna: True.

Prof. Sanjeeva Srivastava: We will can talk little bit on the software which even I am using genepix pro software and how one can acquire image if one has system like 4000B.

(Refer Slide Time: 3:50)

GenePix Pro Acquisition: 4000B image capture

- Simultaneous acquisition of 2 wavelengths
- Adjust PMT voltage manually on-the-fly, or allow software to do it automatically
- Adjustable laser power (10%, 33%, 100%)
- 5 to 100 μm resolution
- Adjustable focus (-50 to +200 μm)
- Line averaging.

Hardware S...

<input checked="" type="checkbox"/>	635
PMT Gain:	600
Power (%):	100
<input checked="" type="checkbox"/>	532
PMT Gain:	600
Power (%):	100

Auto-PMT

Pixel size (μm): 10

Lines to average: 1

Focus position (μm): 0

Scan area: 1891 x 2089 pixels

Scan duration: 52 seconds

NPTEL

BRIDGING TECHNOLOGY ENABLING SCIENCE

Mr. Pankaj Khanna : Right. So, there is the button toward the side where you had having a control for the scanning time. So, as it is a dual channel that is the two lasers are present in that it is allowed to select whether you want to used one and two.

Prof. Sanjeeva Srivastava: okay.

Mr. Pankaj Khanna: And then based on based on one or the user application you select both the lasers and then at the level of live scan you control for the PMT and the pitch resolution you want to use for. So, these are within the same software towards the right hand side you see the pane where you have hardware control button there you can also look for all these different images which is being now suiting for your own biological applications.

Prof. Sanjeeva Srivastava: I think it will be more clear when we are actually looking at the software interface.

Mr. Pankaj Khanna : Sure

Prof. Sanjeeva Srivastava: okay, so one we are acquiring the image what the intensity histogram still is while scanning is in process and even after the scanning is done how one can really ensure that the scan is good and what type of balance one need to make in that.

(Refer Slide Time: 5:09)

Molecular Devices

GenePix Pro Acquisition: Intensity Histogram

- Real-time pixel intensity vs. normalized pixel count
 - Representation of image viewed in Image Tab
- Quantitative and qualitative measure of channel intensity balance
 - Helps user capture well-balanced images

NPTEL

SPINCO BIOTECH PVT LTD
EMPOWERING TECHNOLOGY ENABLING SCIENCE

Mr. Pankaj Khanna: yes so, basically in the preview scan when you are scanning your live data you can just switch on to the histogram graph like there what it gives is how much red and green channel are contributing towards the intensities. So, you really want that they are overlapping so they are really balancing.

There are could be small variability in the beginning going to the fact that they are just background and then the spots coming in. So there you want that they are really overlapping after little bit of lag that is few seconds of lag, that is it. So, once you are able to do then you can see and select whether yes this PMT is being good for me so, this is a way you check which PMT is more suitable.

So you select the PMT look at them see the overlap wherever there is best overlap without saturation you want to go with those settings.

Prof. Sanjeeva Srivastava: I guess now its time that we should really move on to the your software interface and just see like how really the scanning is performed by using this software.

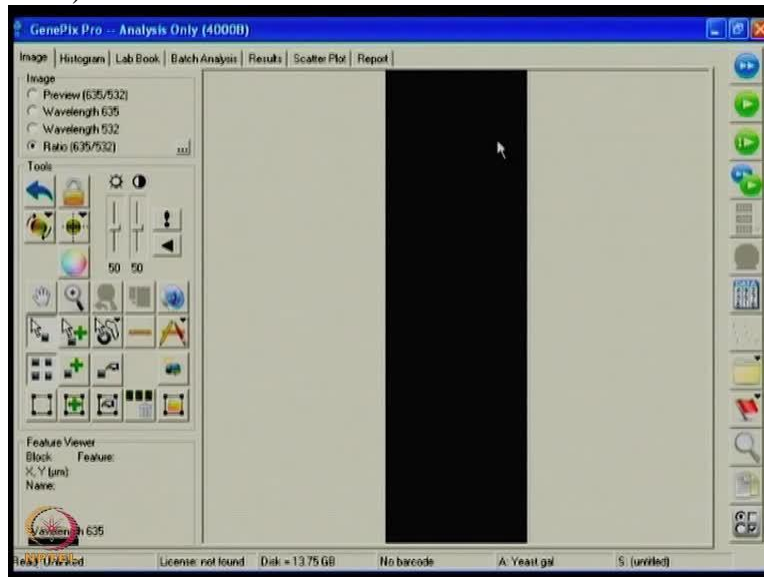
Mr. Pankaj Khanna: yes let us switch on to the GUI.

Prof. Sanjeeva Srivastava: Okay Pankaj so let us talk about the software how actually we can use this software for image scanning from the hardware so that now we can see it live that how one can actually control different buttons for acquisition of the good images.

Mr. Pankaj Khanna: let us go to the GUI of genepix pro.

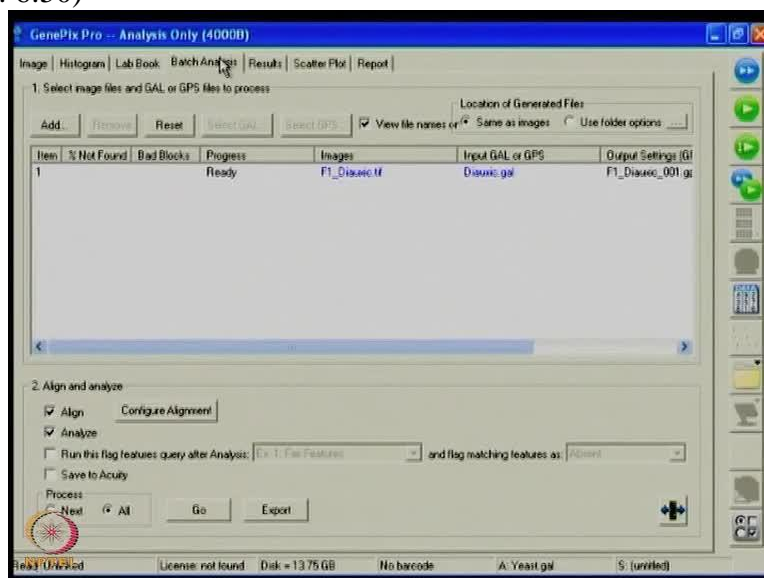
Prof. Sanjeeva Srivastava: Sure.

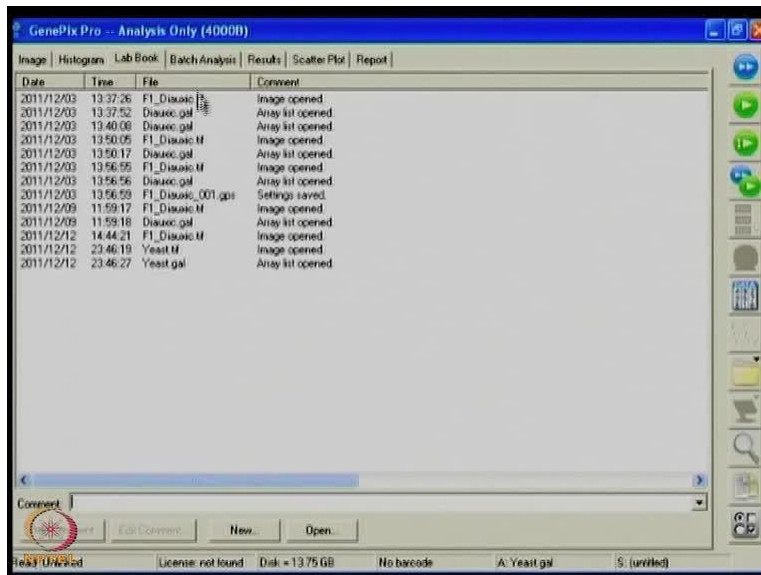
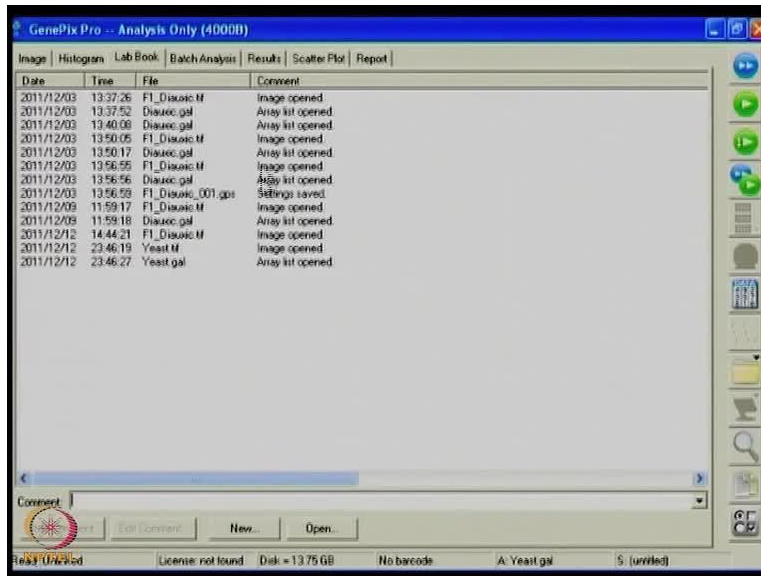
(Refer Slide Time: 6:21)



Mr. Pankaj Khanna: So, what you see now is a GUI or graphical user interface of genepix pro software. On top it is in the form of different tab buttons which allows you in different work group say for example image allows you in different ways and controlling of the image acquisition and histogram looks at how that image has performed. So this is what we were speaking about in the earlier slide where we can see a live kind of demo which is happening.

(Refer Slide Time: 6:50)



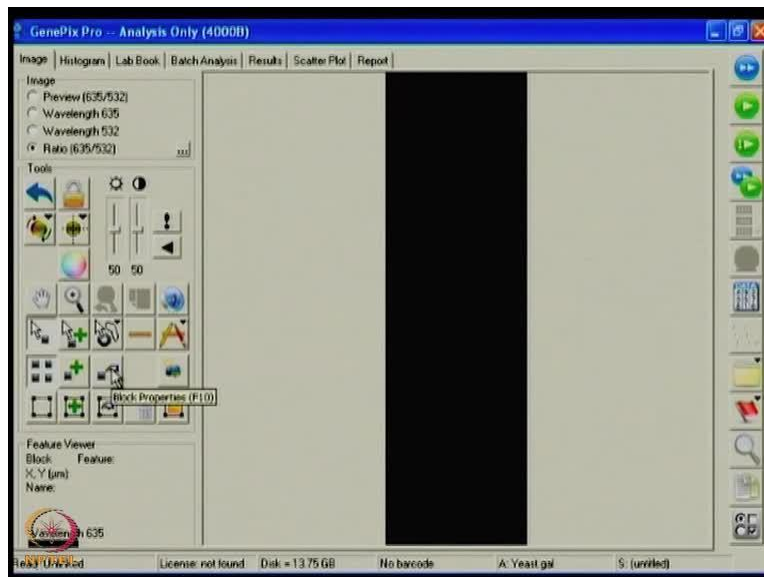
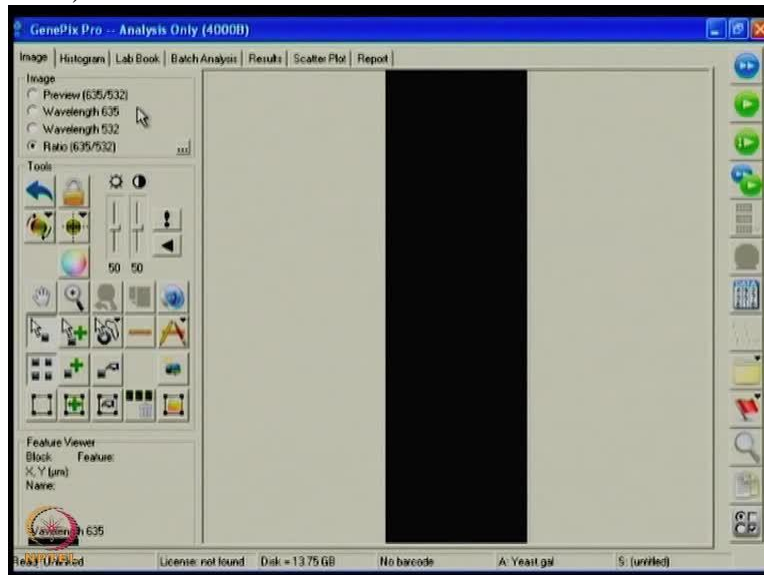


And then lab book actually gives what all you have done in the different stepwise. So what is being every movement of yours in this particular software is being logged in and analysis can be done in the form of batch form which allows multiple slides so that you can do alignment and the analysis which can be perform to the batch analysis and once the analysis is over, results can be seen and scatter plot can be now plotted at the level of this graphical user interface.

Once you are through you can look at the reports as well so, let us look at the major function of imaging so, what and how one control for the best image acquisition.

Prof. Sanjeeva Srivastava: It is important. Let us look at some of the buttons and how one can control those to acquire good image.

(Refer Slide Time: 7:37)



Mr. Pankaj Khanna: Sure. Let us quickly go through like different kind of buttons here. So, now the imaging can be done at different wavelengths and the like preview can be done at 635 and 532. In a single laser based the wavelength can be done at 635 or wavelength at 532. So, even the

ratio of the imaging how the both has performed together can be looked at looking at a button of ratio of imaging.

So, this one allows you to see how the image is being done after the scan. You can look at one channel, preview channel or different channels and now let us look at different tools which are available to you while or after the scanning. So, the major ones are here where you can move across the chips in the form of this hand tool, the + indicates the zoom tool and the other tools are actually this is who want to unzoom and you can also look at the whole image button.

So, once you have the image these two becomes activated. So, these two are actually for the blocks and looking and the controlling the blocks and these ones other features. So, many a times what happens usually you get the GAL file which is actually the feature information file.

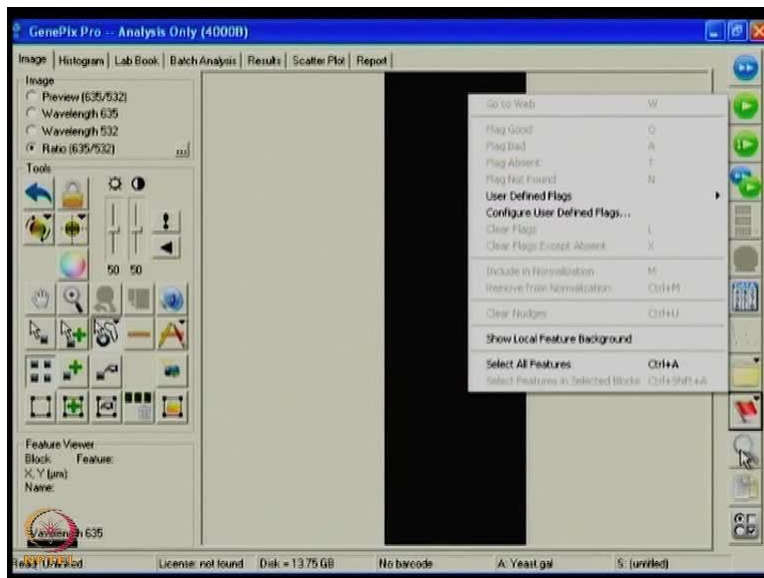
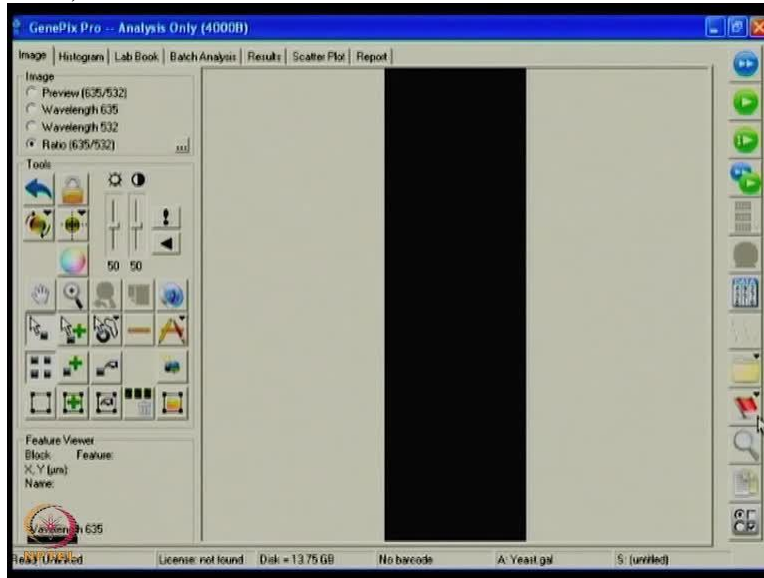
Prof. Sanjeeva Srivastava: Just actually you can just may be elaborate on GAL file because that is one of the very commonly used terms when people talk about microarrays right.

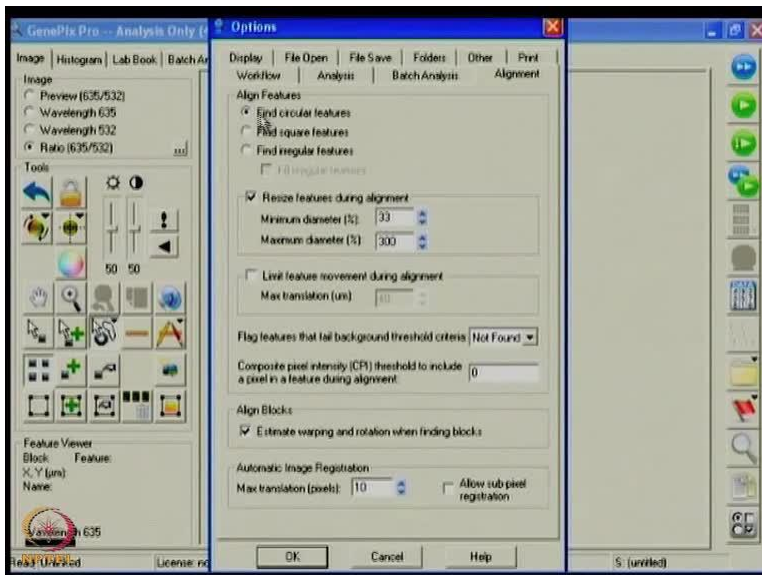
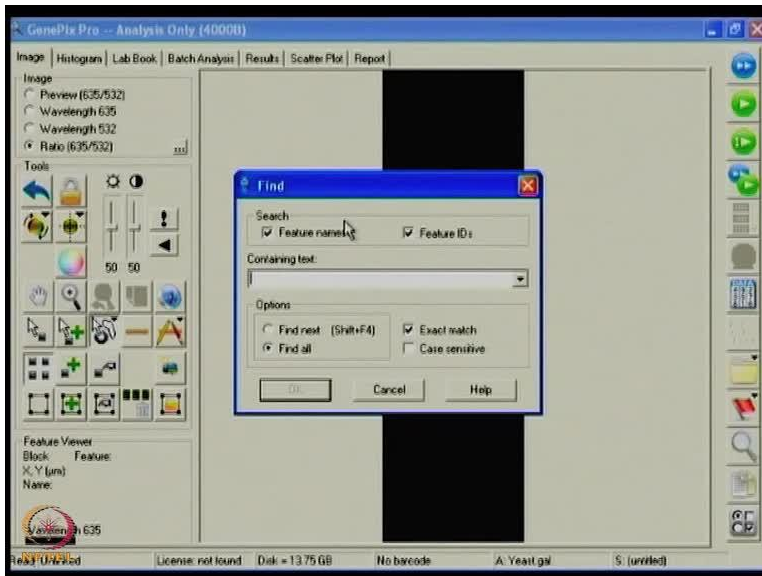
Mr. Pankaj Khanna: Sure. So GAL stands for gene array least file. So, actually it gives the x and y coordinates where each array usually is being presented in the form of blocks which blocks in turn are in the form of features. So these blocks and feature positions are being recorded in GAL file and then the information or annotation is given to each spot. So GAL GAL that is gene array least file essentially contains the x, y and the number of columns and so also the information of each spot how they are being annotated and placed on the chip.

So, if by chance if you do not know or you have prepared yourself these buttons here allowed to make here own blocks and we create here own GAL file with the help of the tool which is called as gene arrays least generator.

Prof. Sanjeeva Srivastava: okay.

(Refer Slide Time: 9:48)





Mr. Pankaj Khanna : So, now let us look at the control button which is towards your right side to the first one is a preview scan and then you can also have a data scan.

Prof. Sanjeeva Srivastava: okay.

Mr. Pankaj Khanna: one stance for one wavelength. So, its allows you to image from only one way and then you also have a multiple scans. So you do a preview scan, then you do a scan with this button. You also have other buttons which will light up as and when you acquire the image and this is for the analysis. So the once analysis is being done if you click this button the analysis will be performed after the alignment.

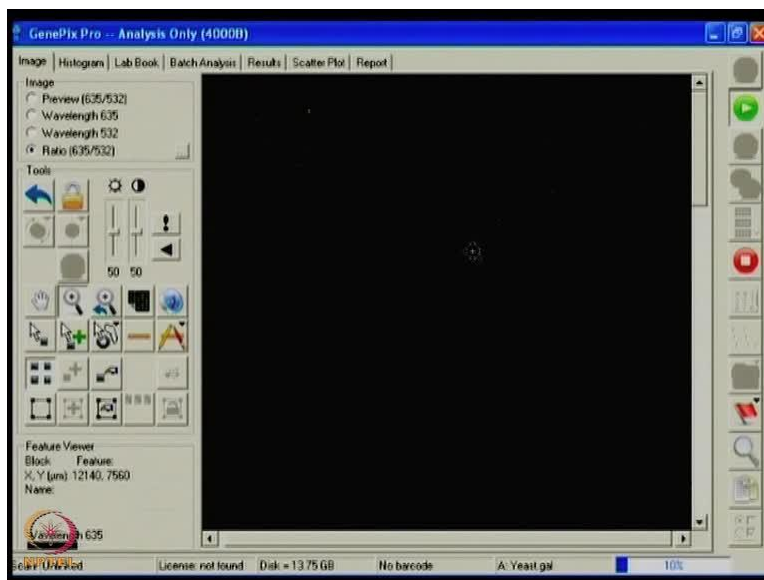
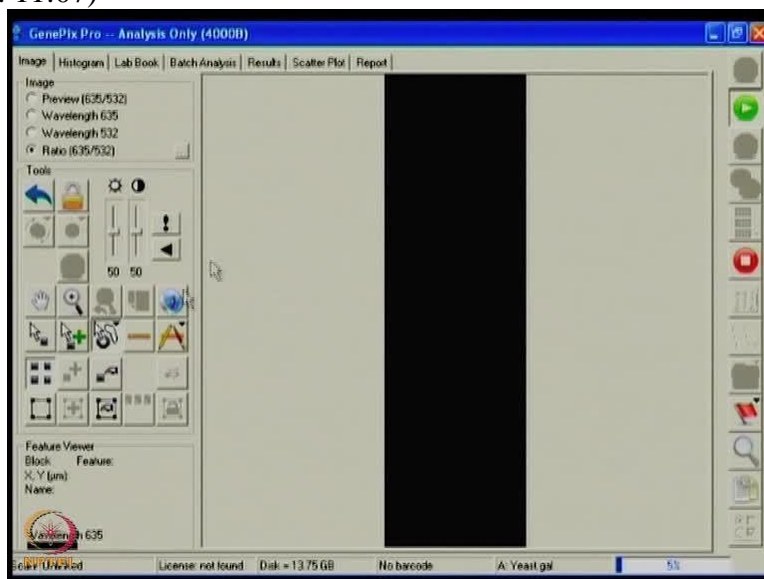
So, this is actually open button so this is like normally your file where you want to open or save your images and this one is actually a flag. As we discussed the different features can be flag we can look at once the images available to you you can look at good, bad or absent and you can give them different ratings.

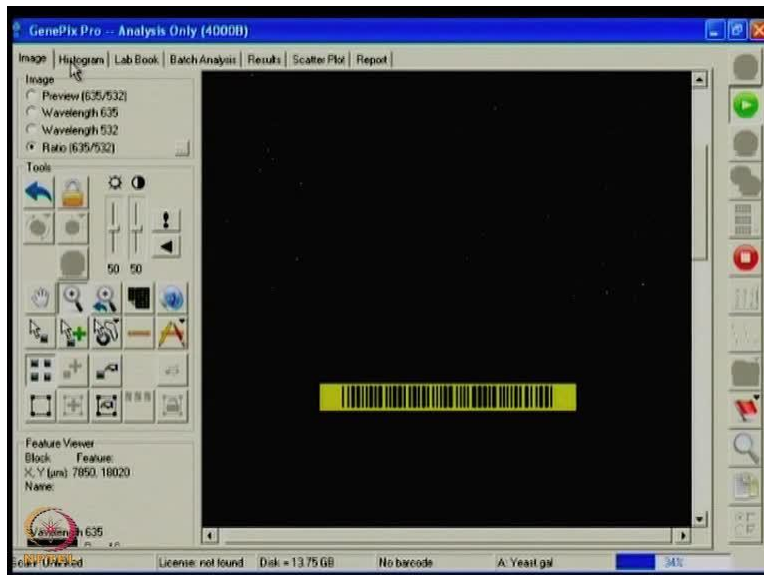
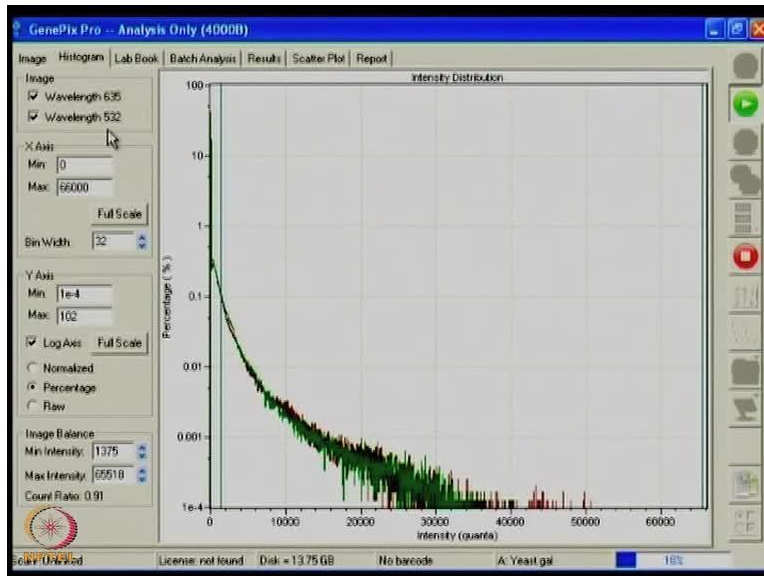
Here again is looking at different zoom buttons so which allows you that which you want to focus are, feature names so the feature ids where you want go for what particular (())(10:52). The major one here is particularly this which allows a different workflow controls right? So, this allows in form a different ways can be happened.

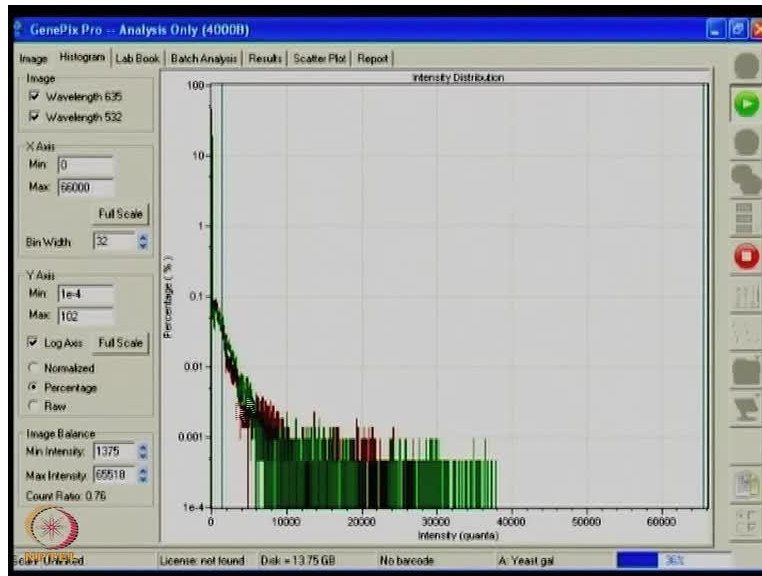
Now let us quickly go through a one particular scan which is a simultaneous simultaneous scan. So both lasers would be acquire to the same time.

Prof. Sanjeeva Srivastava: okay.

(Refer Slide Time: 11:07)







Mr. Pankaj Khanna: So if I pressed on a data scan button the image the after putting inverted slides in the hardware. It is scanning so you just see on the top which is very less visible so let us try to zoom inside. So if I just put this button and allow you to zoom you can see particular how the scanning is happening. So, you are looking at different image type. So, if I click on only one wavelength this because its live after the scanning you can see it is going for the ratio image scan.

So now quickly and a histogram you see it is start coming up. Because it is scanning is going on live so its start reducing. So basically as we discussed it should be overlapping so, my settings are usually looking very nice in this particular work.

Prof. Sanjeeva Srivastava: So I guess like as the scanning is progressing one need to also keep looking simultaneously the histogram.

Mr. Pankaj Khanna: Correct.

Prof. Sanjeeva Srivastava: To determine like how si3 and si5 are well aligned.

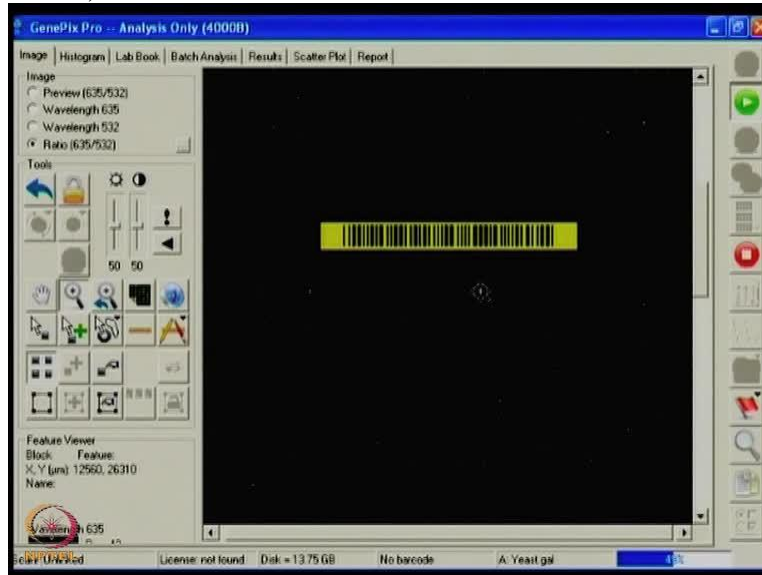
Mr. Pankaj Khanna: Correct. So how well allied with the help of auto PMT so si3 si5 we can adjust auto PMT you can adjust laser power so that you can see this one.

Prof. Sanjeeva Srivastava: So if we see some variation then we need to come back here and adjust this parameter so that there are (())(12:17).

Mr. Pankaj Khanna: Yes. So in this fashion the image acquisition is being happening.

Prof. Sanjeeva Srivastava: Right.

(Refer Slide Time: 12:30)



Mr. Pankaj Khanna: And you also have you we give you a power that in between the slides usually people keep barcodes.

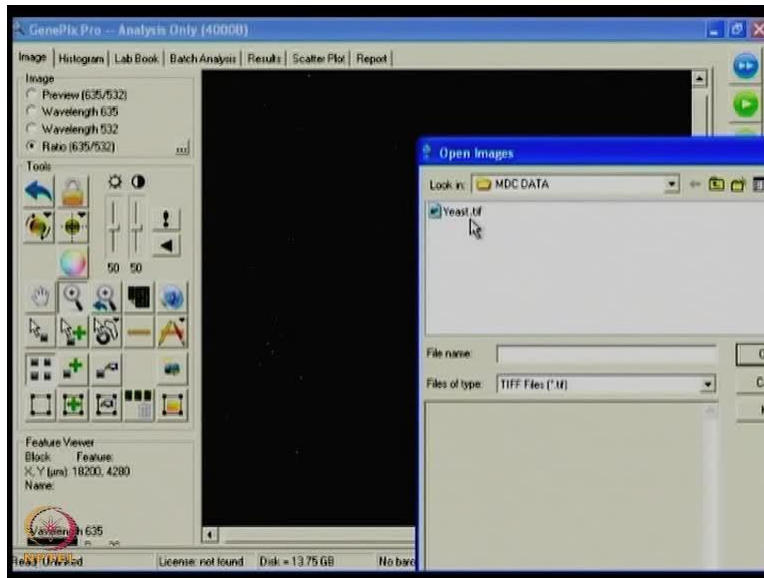
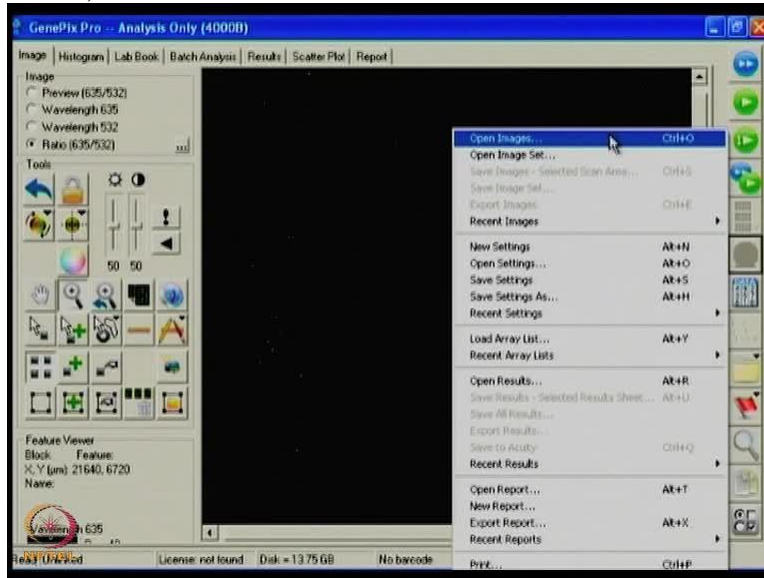
Prof. Sanjeeva Srivastava: Right.

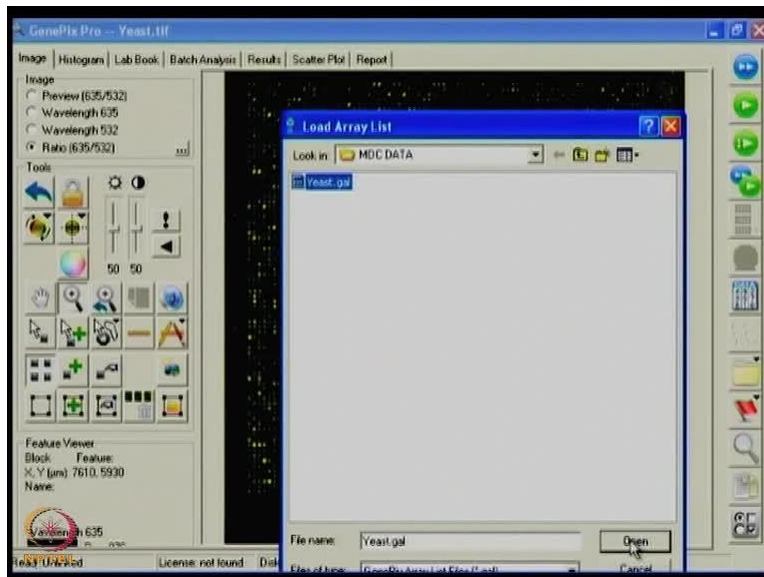
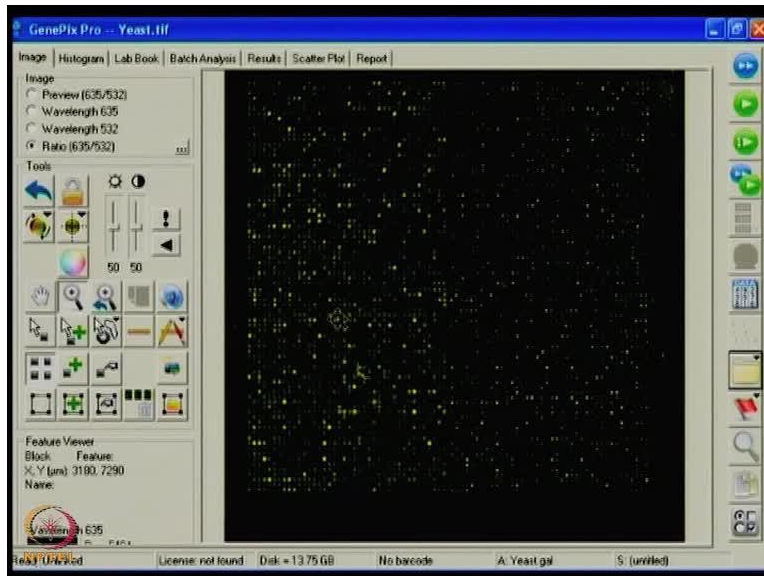
Mr. Pankaj Khanna: and our system or GenePix pro is compatible with reading the internal barcode which is being done. So so that you can have multiple scans also being possible. So nowadays each slides is coming with multiple arrays. Because of the variable densities people are focusing on the custom type. So, this can also be done with the new software upgrade developers.

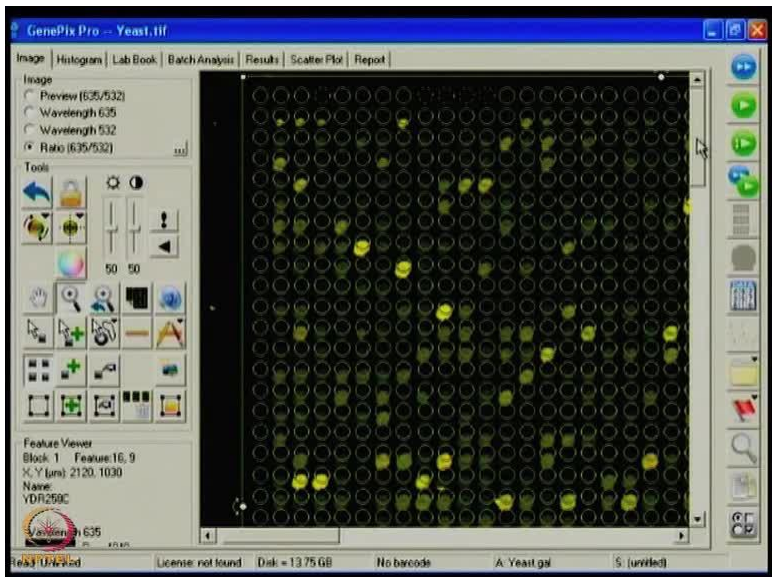
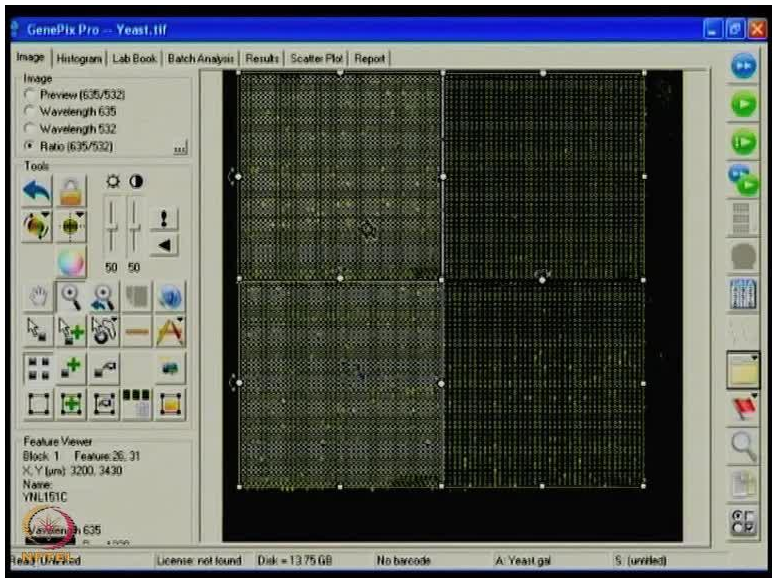
So, now as the scanning is being performed let us look like I will save the image and once I save the image I would like to see how the different processes is being done.

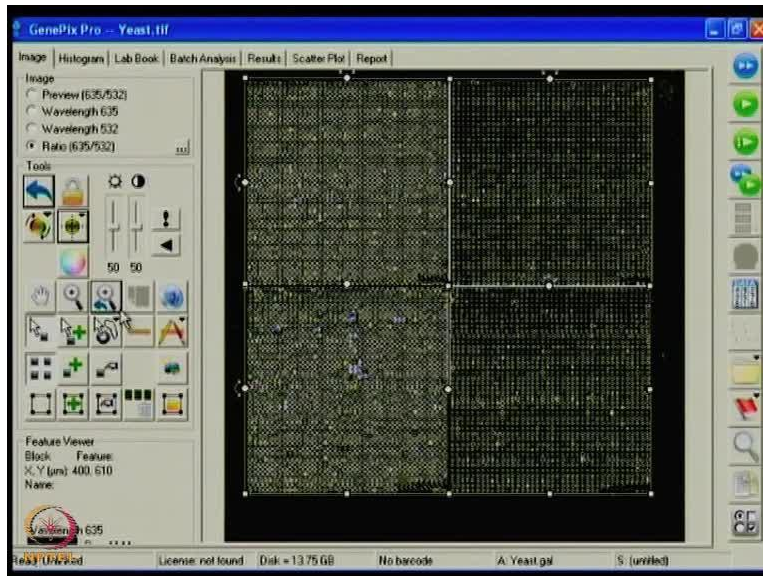
Prof. Sanjeeva Srivastava: ya.

(Refer Slide Time: 13:12)









Mr. Pankaj Khanna: So say for example I have saved this image in the form of say this is a yeast. So I just want to open an image which I have just saved. So, basically as we discussed each particular array can be divided into the blocks. So this particular array of yeast contains four blocks and each block is having features. So number of features information is given in GAL so basically terminology is array block and features.

So I need to align that GAL information of positions on top of this so I have to put a GAL file and do my further analysis. So, what I am going to do now is open a GAL or GAL file which allows me for alignments.

Prof. Sanjeeva Srivastava: Right.

Mr. Pankaj Khanna: So, best feature of GenePix pro is its capability of identifying feature by itself which totally (())(14:13) so let us just see how the zoom button looks like.

Prof. Sanjeeva Srivastava: And one need to fine tune that alignment for overall proper image extraction.

Mr. Pankaj Khanna: Correct. So, only thing you have to do here is just take your block and allow it to move to the first alignment and then what you can do is click the button over here which is for the align. Align can be done in different ways. I recommend to use the first which finds all features all blocks and do automated fashion. So if you click once you see software automatically

finds all its features wherever by chance the features are absent or they are some physical deformity is say it is not present or it flag it as back.

Prof. Sanjeeva Srivastava: So one thing which is good here like its automatically adjusting according to the spot size.

Mr. Pankaj Khanna: Correct.

Prof. Sanjeeva Srivastava: The overall width it is adjusting according like with their some spots are not so uniform right.

Mr. Pankaj Khanna: correct.

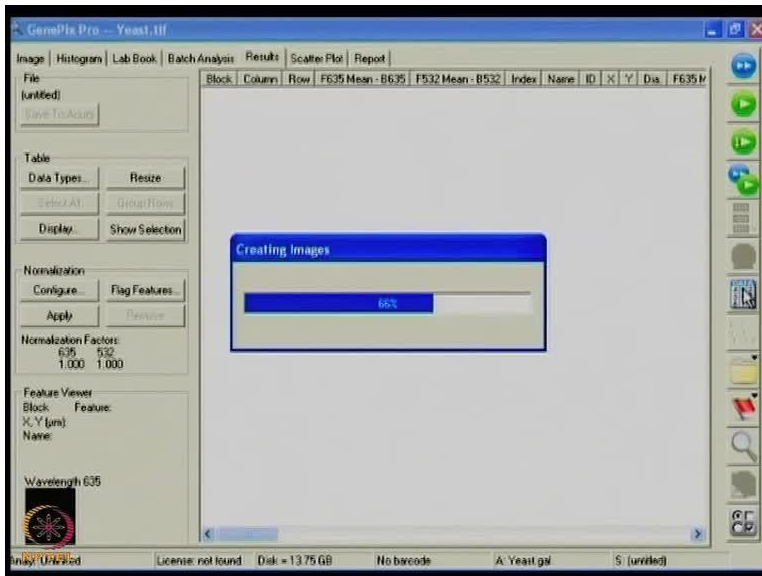
Prof. Sanjeeva Srivastava: So it is making that correction here.

Mr. Pankaj Khanna: Correct. So as you see you can actually move it but it does not affect live as long as you have just kept once and the data being stored but usually people ask me is it good idea if I am by chance move to do once again. So it is not bad idea because it would takes few seconds to do it. So, once you have done this particular alignment let us look this two slides which I said it can zoom out.

Prof. Sanjeeva Srivastava: Ya

Mr. Pankaj Khanna: So you can see whole particular slide now is being scanned and aligned as well. So it is a very quick process which software performs very easily for doing the job and once you have done this you can always hit a button of results.

(Refer Slide Time: 15:51)



GeneDix Pro -- Yeast.tif

Block	Column	Row	F635 Mean - B635	F532 Mean - B532	Index	Name	ID	X	Y	Dia	F635	
1	1	1	2255	20503	1	G	5	2	50	115		
1	2	1	22680	20590	2	G	6	2	50	255		
1	3	1	19991	19811	3	G	7	2	50	236		
1	4	1	27	-24	4	E	8	2	80	132		
1	5	1	783	1288	5	p	9	2	70	174		
1	6	1	1434	2104	6	L	1	2	50	154		
1	7	1	29918	25222	7	p	1	2	60	315		
1	8	1	28378	3059	8	s	1	2	50	301		
1	9	1	-19	-51	9	E	1	2	80	121		
1	10	1	-3	-94	10	E	1	2	80	124		
1	11	1	-51	-2	11	E	1	2	80	116		
1	12	1	-27	-61	12	E	1	2	80	120		
1	13	1	17066	15138	13	cyl...	Y	1	2	60	226	
1	14	1	9070	9037	14	PIP...	Y	1	2	60	110	
1	15	1	964	1658	15	ace...	Y	2	2	60	226	
1	16	1	6018	5974	16	ho...	Y	2	2	60	682	
1	17	1	5489	6113	17	p	2	2	50	494		
1	18	1	4388	4040	18	K	2	2	60	687		
1	19	1	5261	5497	19	c	2	2	60	588		
1	20	1	2214	2823	20	c	2	2	60	322		
1	21	1	4352	4202	21	c	2	2	60	534		
1	22	1	1783	2461	22	c	2	2	70	276		
1	23	1	4517	4630	23	c	2	2	60	440		
1	24	1	9	-16	24	E	3	2	80	121		
1	25	1	2465	3321	25	FIN...	Y	3	2	70	344	
1	26	1	28330	23705	26	tran...	Y	3	2	80	400	

License: not found Disk = 13.75 GB No barcodes A: Yeast.gal S: (unfitted)

Column	Row	F635 Mean - B635	F532 Mean - B532	Index	Name	ID	X	Y	Dis	F635 Medial	
1	1	27255	20503	1	G.	5	2	50	11240		
2	1	25800	20590	2	G.	6	2	50	25304		
3	1	19991	19811	3	G.	7	2	50	23612		
4	1	27	-24	4	E.	8	2	80	1528		
5	1	753	1288	5	p.	9	2	70	1748		
6	1	1434	2104	6	L.	1	2	50	1540		
7	1	29618	25222	7	p.	1	2	60	31508		
8	1	2831	3059	8	s.	1	2	50	3012		
9	1	-19	-51	9	E.	1	2	80	1216		
10	1	-3	-94	10	E.	1	2	80	1240		
11	1	51	-2	11	E.	1	2	80	1188		
12	1	27	61	12	E.	1	2	80	1204		
13	1	17066	15138	13	cyt.	Y.	1	2	60	22844	
14	1	9070	9037	14	PIP	Y.	1	2	60	11040	
15	1	864	1658	15	aca.	Y.	2	2	60	2256	
16	1	6018	5974	16	ho.	Y.	2	2	60	6824	
17	1	5489	6113	17	p.	2	2	50	4940		
18	1	4388	4040	18	K.	2	2	60	6872		
19	1	5261	5497	19	ca.	2	2	60	5880		
20	1	2214	2823	20	ca.	2	2	60	3224		
21	1	4352	4202	21	ca.	2	2	60	5340		
22	1	1783	2461	22	ca.	2	2	70	2784		
23	1	4517	4630	23	ca.	2	2	60	4404		
24	1	9	-16	24	E.	3	2	80	1216		
25	1	2465	3321	25	RN.	Y.	3	2	70	3440	
26	1	29330	23785	26	tran.	Y.	3	2	80	40007	

Now if I go to results actually this is an empty. So, if I click on results results are being calculated and there are some 40 different columns which will be output in the form which GenePix pro understands different ways.

Prof. Sanjeeva Srivastava: Sure.

Mr. Pankaj Khanna: So just quickly looking it to the major ones. The major ones here are looking at this F means the intensity from different channels 635 or 532 and this background calculation is being done accordingly in the same laser range. So, once you do a corrections what happens is you want to correct your intensity mean values with the values of the background.

Prof. Sanjeeva Srivastava: Okay

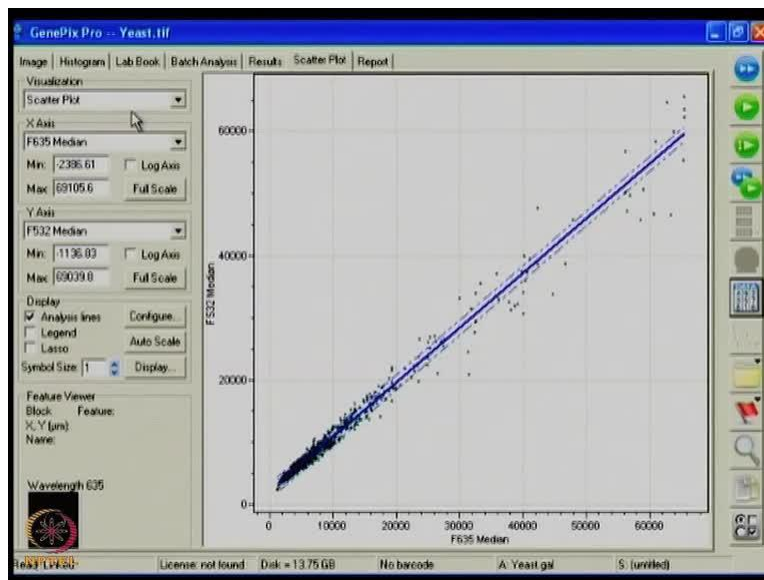
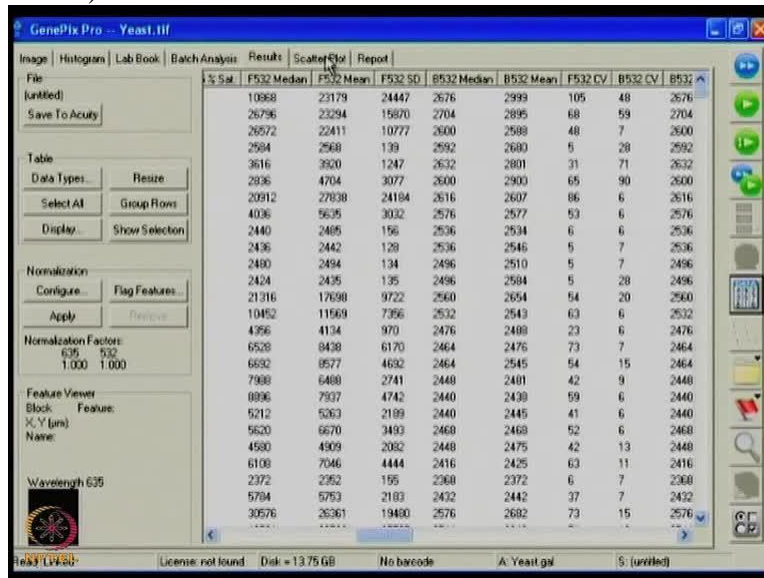
Mr. Pankaj Khanna: So this is what is the most important which usually people use for the further calculation apart from a ratio of means or ratio of medians which can be calculated again and being presented to you in different column formats. So each column signifies different ones. Like for example SD standard deviation, CV coefficient of variations and then different channels coming up.

So, in this fashion the result would be outputted if you are image acquisition first controlling the part then allowing you to align and then do the analysis. So this is the basic steps which anybody

or everybody want to do in microarrays steps. So once you see the images the the people end up in the form of results and you have different columns available to you.

Prof. Sanjeeva Srivastava: Right. After that what is a next into look for like how would set plots are or...

(Refer Slide Time: 17:25)



Mr. Pankaj Khanna: So there are different ways people want to visualize how my column because number make very less sense. So the best way to look is a scatter plot. Scatter plot allows you in different ways what you are plotting at X and Y axes and here if you see I am

just plotting actually towards F635 median over the F635 median. So you are comparing two different channel how they have behave.

So essential rule is they should mostly the microarray assumes all the sick chips are having the spots and which are genes which are not bearing too much. So you expect most of them to stand nearby to the origin of the center. So this is what you want look at. How there plot are aligned. Yes so they should be not too far away from each other. So that they are not need actually to each other because you expect there are few differences but not very very significant which can be seen at very large scales.

Prof. Sanjeeva Srivastava: So I guess we talked so far about how to use the hardware to scan a slide and by using this software interface which is GenePix pro here to acquire all the data. Now next challenge is how to really obtain some meaning full information from this whole data which we have already acquired.

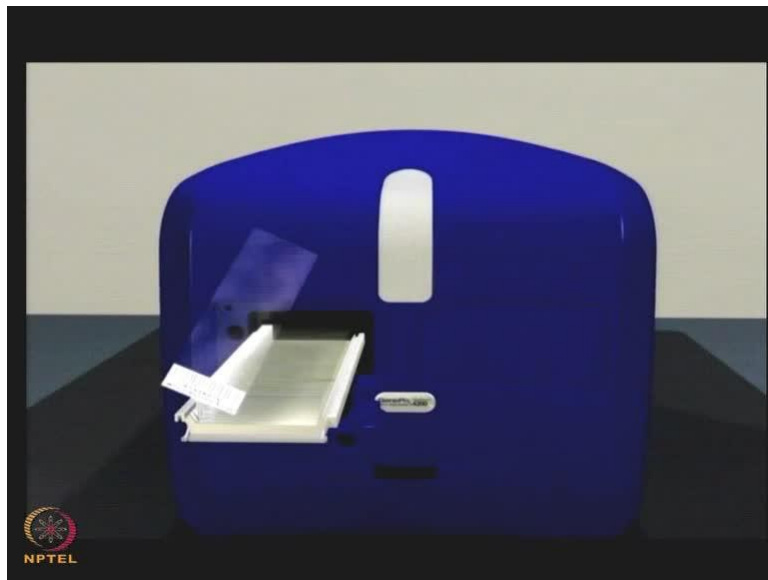
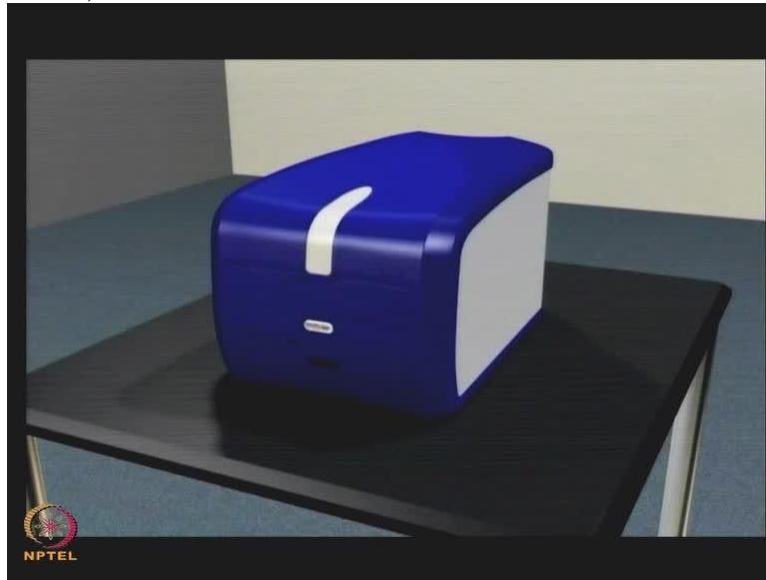
Mr. Pankaj Khanna: True. So GenePix pro as we discuss it acquisition software and the molecular devices recommend acuity software for further data analysis which can be at the level of secondary or tertiary based on that. So you do statistics as well as visualizations on a single or multiple data to handle that.

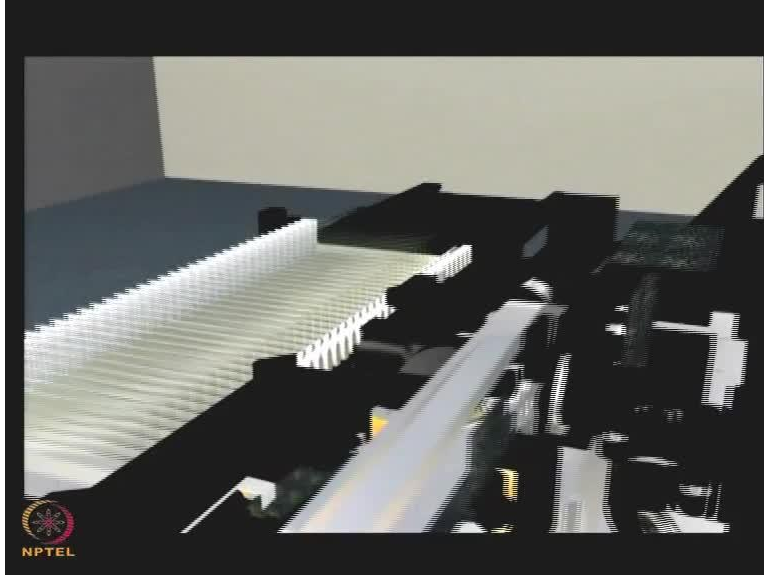
Prof. Sanjeeva Srivastava: So Pankaj it was good to talk to you about a how one can use the DNA microarray (())(18:59) protein microarrays, different type of slides and use the hardware the scanner to scan these slides by using this software which has various features and depending upon experiment one can actually take a decision that what type of features one need to use for that scanning.

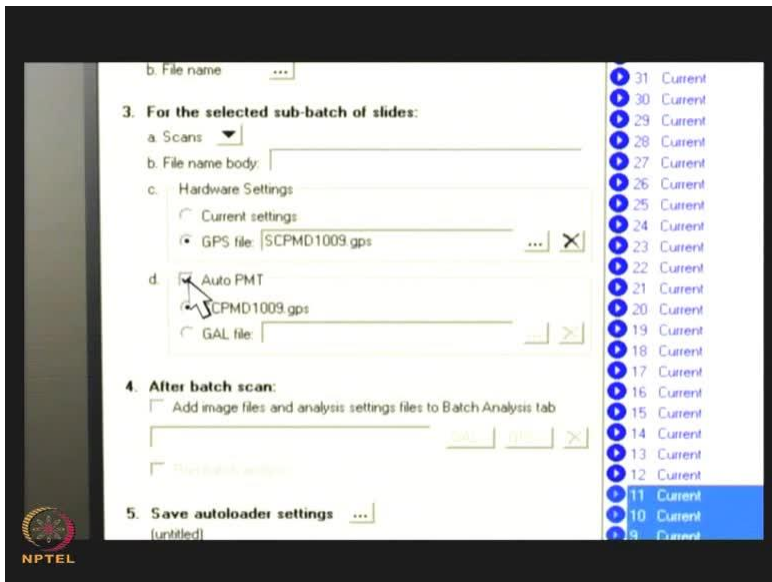
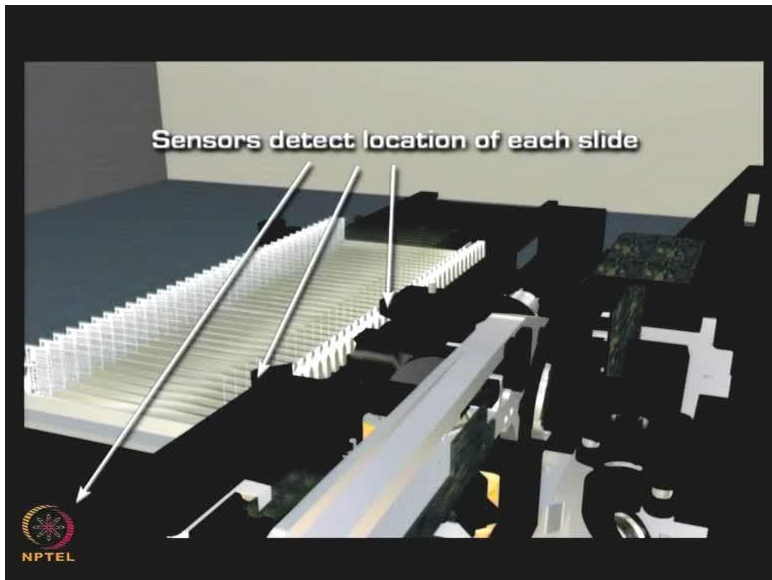
I think it would be very good if you can just share with us the video which describes overview of the whole process is starting from the sample preparation followed by a image acquisition as well as image analysis. So if you can show us that video that will be great.

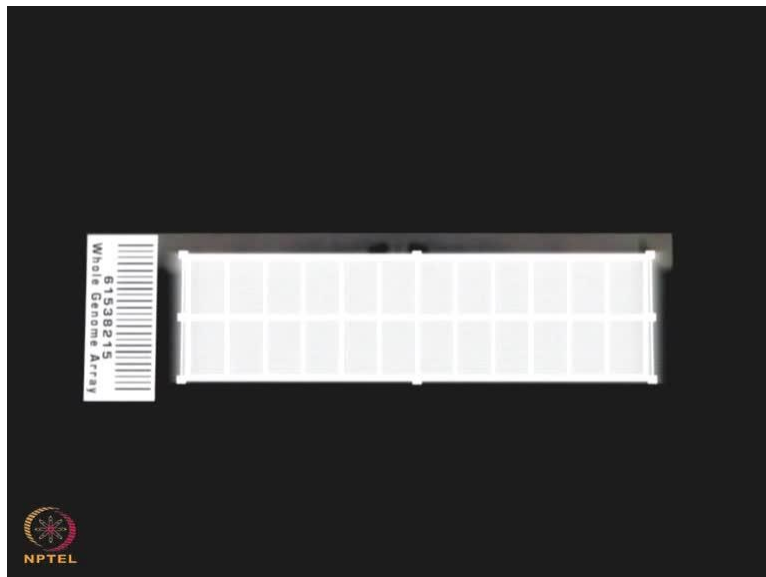
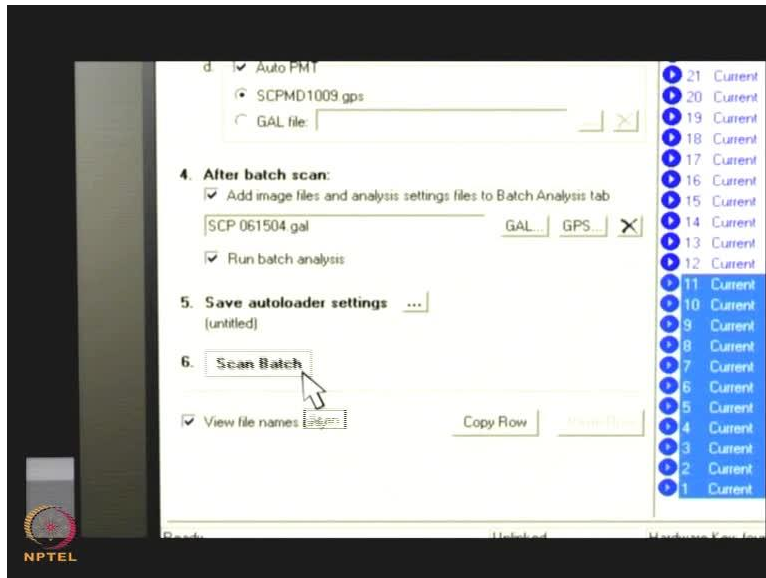
Mr. Pankaj Khanna: Sure. Thanks to molecular devices they have provided me with this video which actually takes you from the basic process of biology in very very brief to importance software parts and so also the hardware design which is being emphasis to the label of result. So let us watch that video.

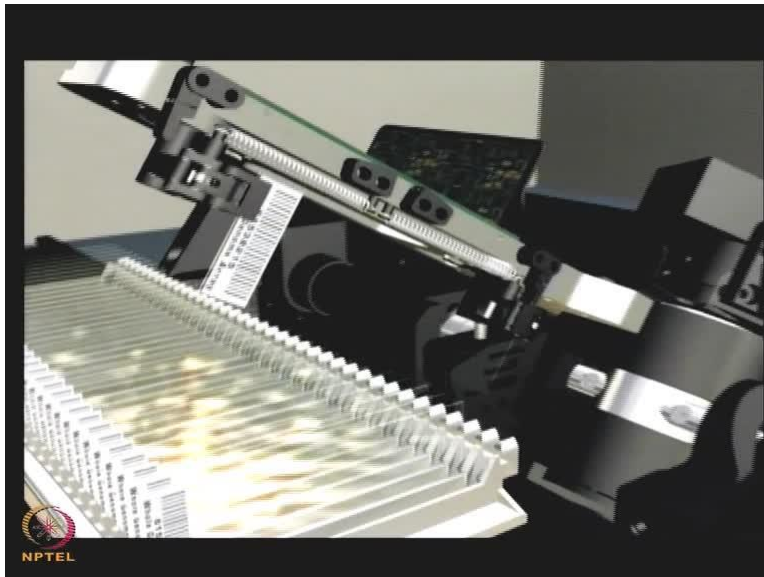
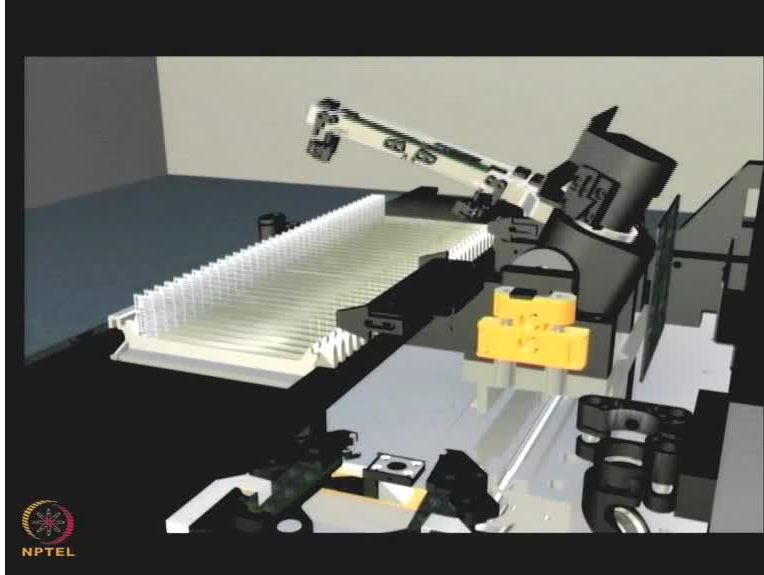
(Refer Slide Time: 19:52)

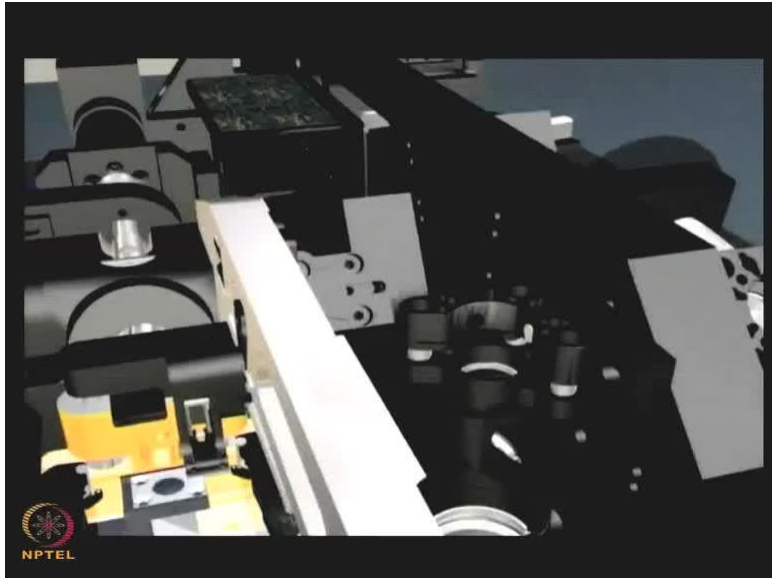
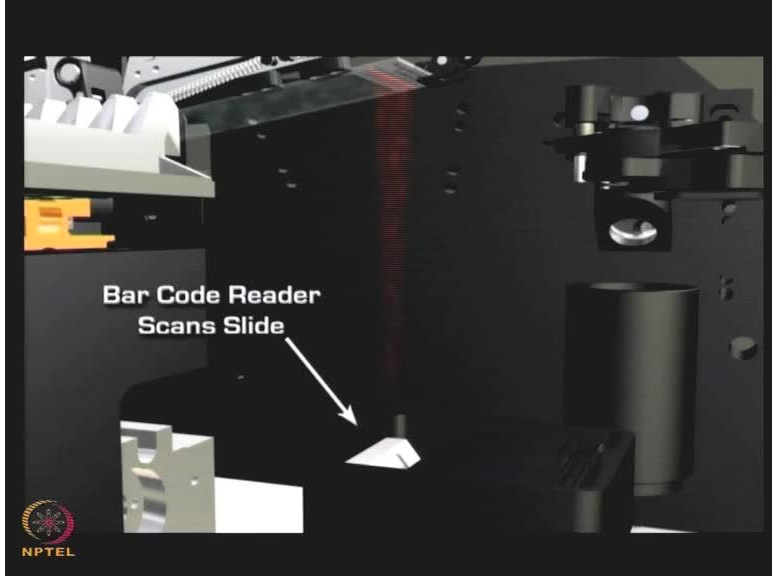




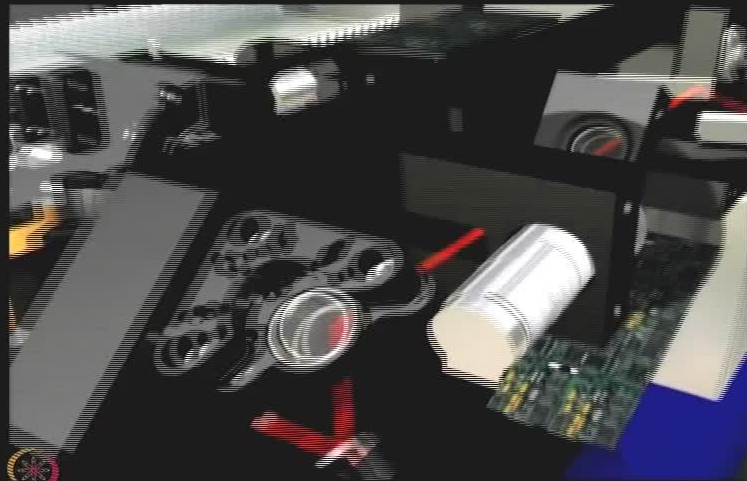
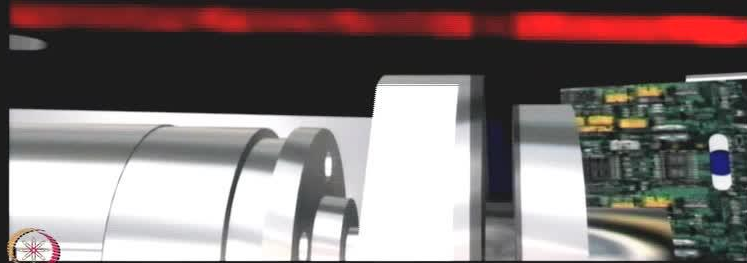


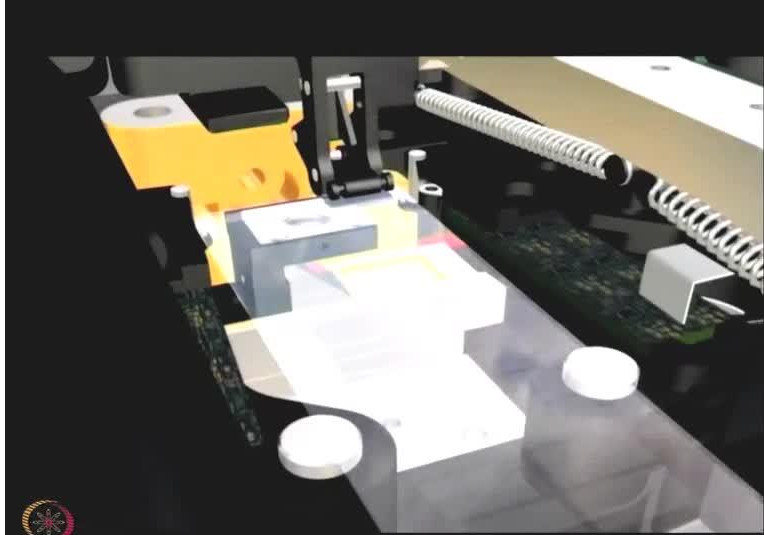


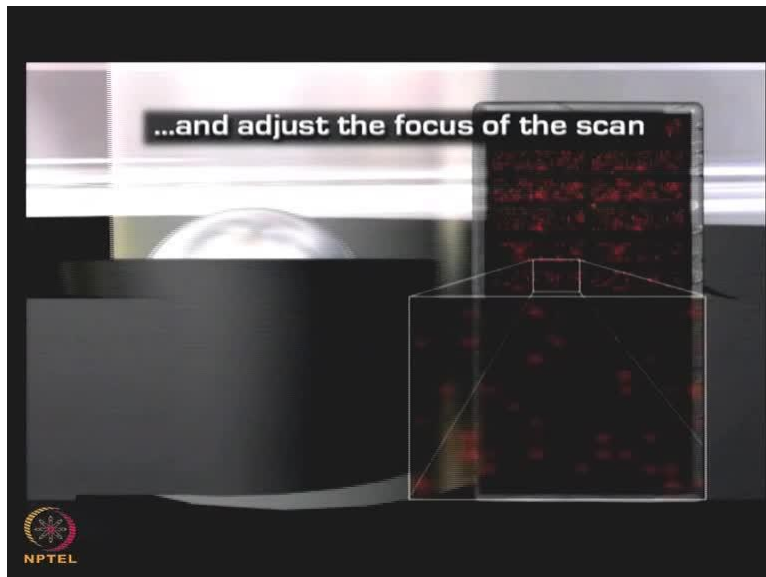
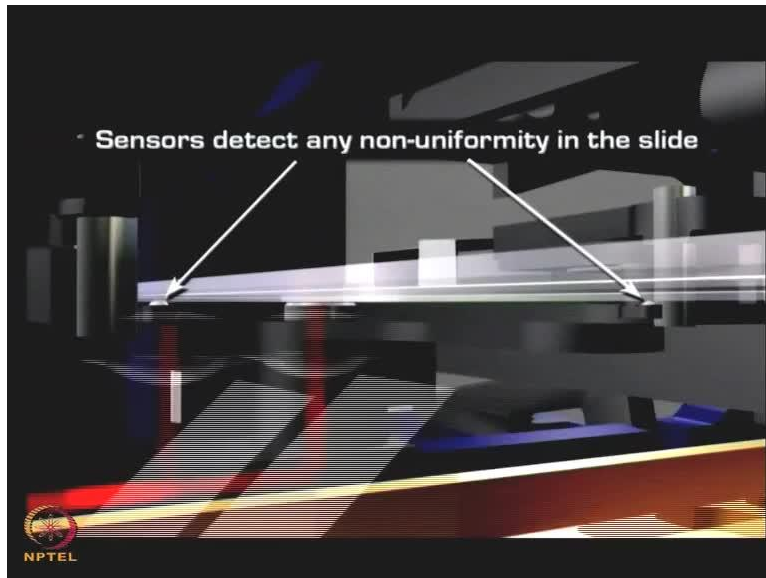


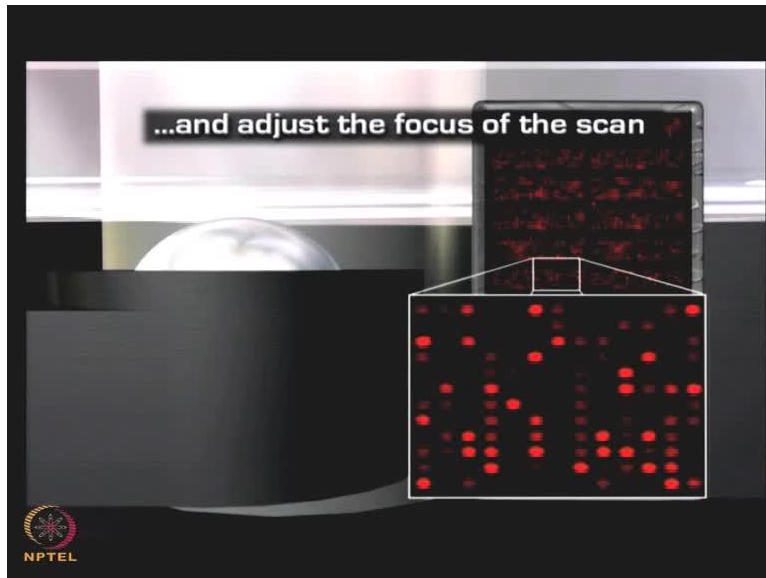


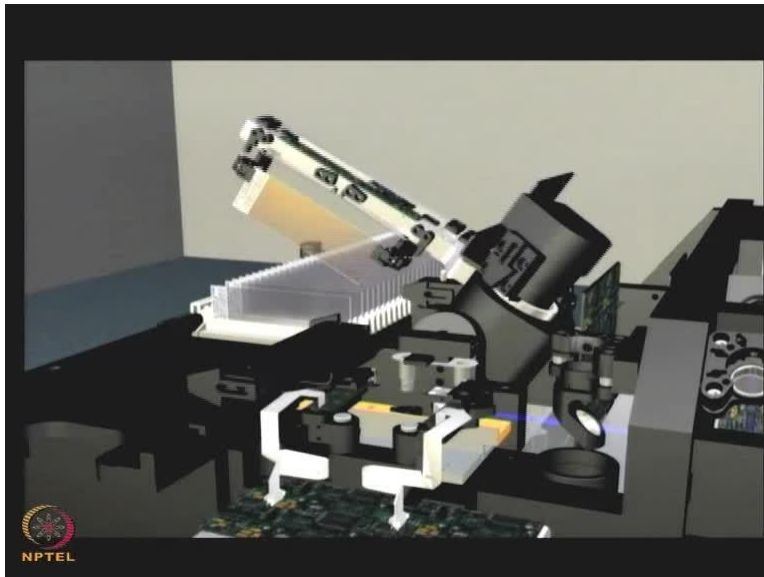
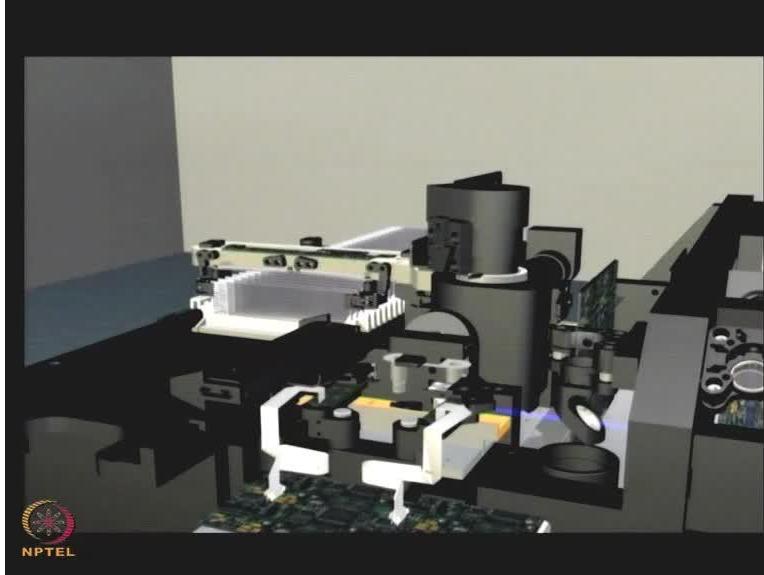
Laser Power Wheel Adjusts Laser Strength

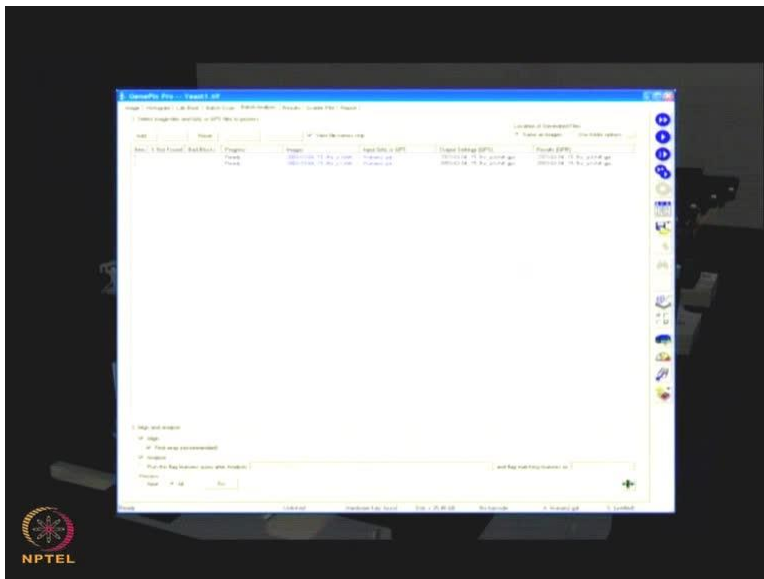
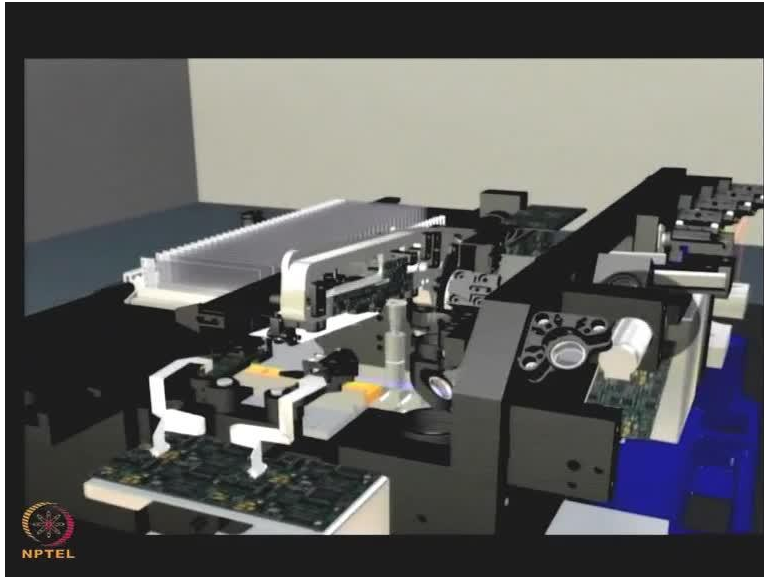












GenePix Pro Microarray Image Analysis

Gene Name	Ratio of Medians	Ratio of Means
bcl-2	0.400	0.358
interleukin-1	1.000	0.850
caspase-6	5.000	4.500
p-53	2.000	1.780
c-myc	0.600	0.516
CD34	1.000	0.880



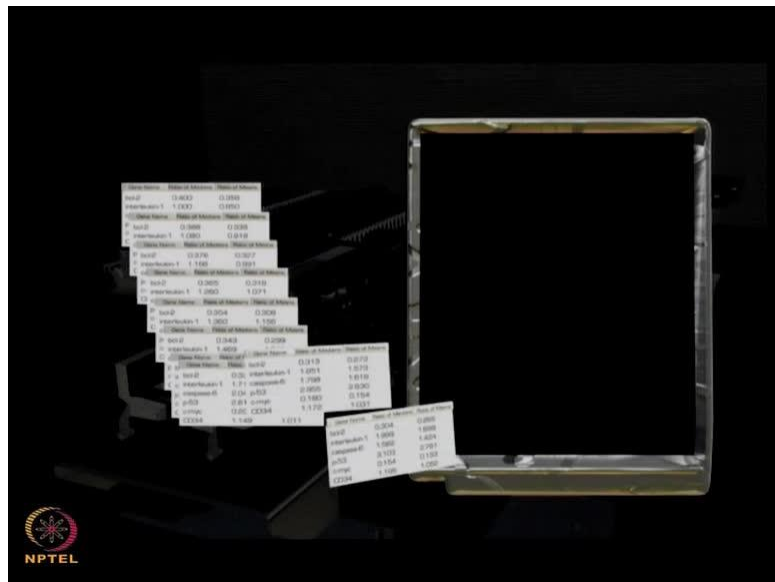
GenePix Pro Microarray Image Analysis

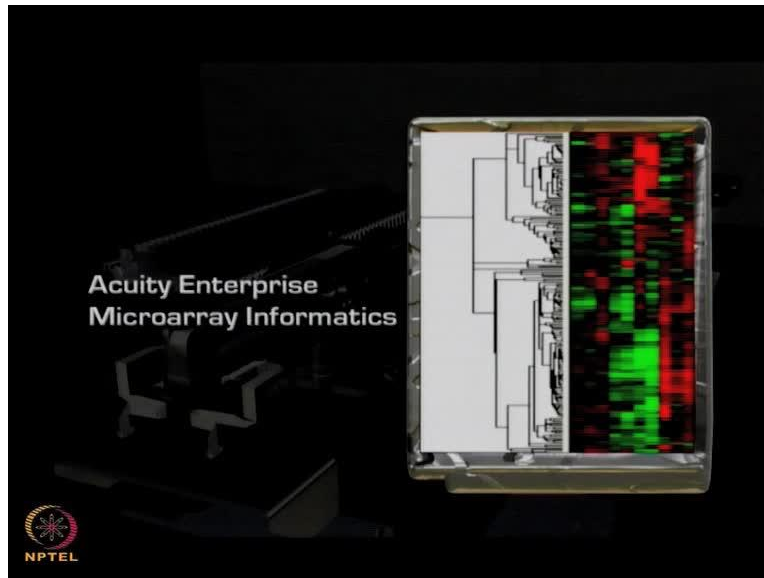
Gene Name	Ratio of Medians	Ratio of Means
bcl-2	0.365	0.318
interleukin-1	1.260	1.071
caspase-6	3.407	3.067
p-53	2.315	2.061
c-myc	0.382	0.328
CD34	1.061	0.934



GenePix Pro Microarray Image Analysis

Gene Name	Ratio of Medians	Ratio of Means
bcl-2	0.343	0.299
interleukin-1	1.459	1.249
caspase-6	2.539	2.375
p-53	2.553	2.272
c-myc	0.282	0.243
CD34	1.104	0.972





Molecular devices introduces the world's simplest most reliable automatic microarray slides scanner. Now you can walk away from scanning while the GenePix Autoloader 4200AL automatically loads, scans and alive and saves results for up to 36 slides.

The Autoloader accommodates microarrays on standard glass microscope slides labelled with up to 4 fluorescent dyes. These microarrays can contain just a few 100 spots or 10s of 1000 of spots representing an entire genome. As many as 36 slides can be loaded in to the convenience slide carrier. As the carrier is insulated into the scanner sensors detect the location of each slide indicated by a blue power on the slide carrier map.

On the batch scan tab in GenePix pro you have complete flexibility to define the most appropriate settings and analysis parameters for each slide or for group sub slides. You can also choose to automate scanning analysis and file saving steps Enter an email address and GenePix pro notify you monthly when your batch is complete.

Using the defined scanning primers the precision robot arm leaps into action and moves to the first slide. Our unique never let go grippers securely clamp the slide and carry it to the scanning area. A barcode reader records the barcode and then the slide is positioned for scanning.

The GenePix Autoloader 4200AL can be configured with up to 4 lasers. A neutral density filter wheel can be used to attenuate the laser power if necessary for especially bright samples. The

laser excitation beam is delivered to the surface of the microarray slide. The beam scans rapidly across the short axis of the slide as the robot arm moves the slide more slowly than the long axis.

Fluorescent signal emitted from the sample is collected by a photomultiplier tube. As the scan proceeds sensors detect any non-uniformity in the slide surface and the robotic arm adjust slide position accordingly to ensure the array surface is always in perfect focus. Each channel is scanned sequentially and the developing images are displayed on the monitor. The multichannel thick images are saved automatically according to file naming conventions specified by the user.

After the slide has been scanned the precision robot arm replaces it safely in the slide carrier before picking up the next slide. As each slide is scanned a list of each saved image with its associated settings and analysis files accumulates in the batch analysis tab until the batch is complete. GenePix pro automatically finds the spots and calculates up to 108 different measurements for each spot. The results are saved as a GenePix results or GPR file.

GPR files can be saved automatically to the acuity database for statistical analysis, clustering and other advanced investigation.

Prof. Sanjeeva Srivastava: okay it was a very useful discussion Pankaj with you for knowing more details about various type of features one need to look for to obtain very good images because that is the most important part for doing the microarrays. Once you have acquired the good images then only one can do good data analysis from that. So thank you for coming here and sharing your experience for overall microarray scanning and the data acquisition. So, thank you once again for coming here and discussing and sharing about your experience for GenePix pro. Thank you.

Mr. Pankaj Khanna: Thank you.

Prof. Sanjeeva Srivastava: In protein microarray experiment intensity of a spot is a representative of the interaction between the sample and analyte. To achieve this target proficiently in image processing and data acquisition is required. As discussed the artifacts due to the contaminants such as dust particle or even very high background issues make automation in image processing

and data acquisition very challenging. Researchers have devised several segmentation algorithms to reduce the manual interventions.

However you have seen there a degree manual flagging is necessary to the mark the low quality spots. This data would now be preprocessed for background correction and normalization. You also seen the demonstration of (25:13) software GenePix pro microarray image analysis software for data acquisition and you will further see how it will be used to analyze data in subsequent lecture.

There are several commercial software like proto array prospective software which are available along with their compatible scanning devices for image processing, data acquisition and preprocessing. In our next lecture we will continue our discussion on microarray experiment workflow and how to analyze the microarray data obtained from the images generated from today's lecture. Thank you.

(Refer Slide Time: 25:58)

Summary

- The GenePix Pro software for data acquisition has an image processing control which allows user to select the laser and feature under focus.
- There are hardware controls which allow user for controlling PMT, preview or a true scan etc.
- Intensity histograms allow users to check if the scan has run smoothly and access balance of signal intensity.
- The GUI has several tabs which allow image acquisition at various wavelengths.



Summary

- It allows overlay followed by alignment of .GAL file which is a crucial aspect of data extraction
- Data is further visualized and analysed using statistical approaches using Acuity software



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

- 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 7

Lecture 35: Microarray Data Analysis: Part I

