LECTURE-15

Two-dimensional electrophoresis: Image processing and data analysis

TRANSCRIPT

Welcome to the proteomics course, in today's lecture we will talk about two-dimensional electrophoresis, image processing and data analysis.

So as in the last few lectures we have discussed about 2-DE and infact various types of gel-based proteomics techniques such as SDS-PAGE, Blue-Native PAGE. We have talked about different types of new advancements like OFFGEL fractionation as well as 2-DE. We also discussed in detail about 2-DE workflow, how to perform 2-DE based experiments step-by-step, various details of performing protein extraction, different types of challenges associated with the sample preparation followed by how to perform isoelectric focusing and then how to do SDS-PAGE analysis, in between these two what is required to make these strips compatible for doing SDS-PAGE which was discussed in the equilibration steps and then the second dimension separation using SDS-PAGE. Once you have performed all these steps and run the gel, after that you want to visualize the spots you need to use various types of staining techniques. So we discussed different types of staining methods, which are available and once you stain those gels you can visualize those spots, you can see how many proteins are visible on the gels and those spots can be further scanned for image processing. Now, good scanning image processing and followed by data analysis to obtain some meaningful information from these gels is very challenging.

So in today's lecture continuing our focus in theme of 2-DE. We will discuss about image processing and data analysis, various types of challenges associated with image processing and what are the steps required to perform data analysis.

Now when you have run a very good 2-D gel you will expect an image as showed here where you have well resolved spots through out the gel. This is the gel image shown from serum protein extracted from human and those proteins separated on the 2-D gel. Now if everything goes well you can see the well-resolved spots though out the gel and the whole protein can be well separated. But many times this is not the case. Many things go wrong, the salt has some interference, nucleic acids contamination will have some effect and you can see various types of artifacts. You can see various types of vertical as well horizontal types of streaking on the gels. These small details make difference when you are doing gel analysis for the image processing part. So how to analyze two gels by using software is not so tedious. But when you are talking about many gels where you have large number of control, large number of treatment then it becomes very challenging. Because all your gels should have run at the same time with the same dye front migration. But many times that is not the case. You cannot run 50 or 100 gels on a given day. So there will be some manual artifacts, day-to-day variations. Now on one hand it's a good idea to keep all your parameters very consistent so that the gel migration, all the protein resolution will be of the same parameter. At the same time even if all gels have not run the exact length still the software's can help you for image processing. You need to do the various type of image processing before you analyze your gels. We will talk about different types of image processing tools, how to prepare the images for data analysis and then how some of the commercially available software can be used for obtaining very meaningful information from these gels, whether you can see some sort of differential regulation of the proteins, up-regulation or down-regulation, or you may have some unique proteins emerging on your sample type, your treatments. So these types of the analyses can be performed on this software and then followed by statistical analysis to obtain confidence in the data, how meaningful this information is and whether I can proceed with these spots for further protein identification using mass spectrometry. So all of these details require very much practice, one needs to really spend hours on data analysis. In fact running gel is much easier as compared to doing very meaningful data analysis.

So to discuss about details for various challenges, which are associated with analysis of 2-D gels, the image processing tools, as well as the details required for the image processing by using various commercial software. Today I have invited a guest here to discuss the image processing and data analysis by using commercial software. So today we will have Dr. Srinivas from GE Healthcare who will discuss how to analyze the gel using Image platinum software.

So during the course of discussion we will talk about stepwise procedure how to take the scanned images, the raw tiff files and then align those and then followed by spot detection and do various types of manual refinement for image processing. And at the end we will discuss about obtaining meaningful information using various statistical analysis. So I will have a discussion with Dr. Srinivas now.

A. So today we have Dr. Srinivas, an application specialist from GE Healthcare who will be discussing with us the challenges of image processing and how to perform 2-DE gel image processing and data analysis. So welcome Dr. Srinivas and we would like to initiate discussion on image processing.
B. Hello Dr. Srivastava. Today I will explain the workflow of ImageMaster platinum. Now it is very important software to analyse 2-D gels. Especially when we have multiple numbers of gels it is impossible to analyse manually. This software plays a key role in the analysis role. As you can see there are major modules. Image Master is a flexible solution for the comprehensive visualization, exploration and analysis of 2-D gel data. Two modules are available in this software. The first one is ImageMaster 2D platinum and within this we have another module ImageMaster platinum, which is enabled for the DIGE application also.

A. So it means that the same software can be used for traditional or regular 2D gels as well as it can also used for the advanced techniques such as DIGE
B. Yes, it can be used. But upto some extend you can use. But if you want to full-fledged DIGE analysis one should have DeCyder. For conventional 2-DE, ImageMaster platinum (IMP) is a better tool for analysing the gels. You can see the overview of this particular image of this IMP software there are 3 major parts. One is the complete icons where you can see the file tools, as well as the edit view. After that you can see image view. These are all the images which you are visualising almost all 12 gels which you are visualising at a particular place and the thirdly which we can see the number of gels present here. The total number of gels which you can see here. These are all the major visualization part in the software.

A. So you are showing the interfase of the ImageMaster platinum where 3 parts are shown in the ppt about various types of tools one can use, how the images look like and what kind of parameters are there to display.
B. Exactly, those are the 3 major interfaces here and you can see even the workflow on the left hand once you go through the software analysis demo there you can see the particular detail of this feature also.
A. So, this software is made in such a way that can walk through the stepwise automatically.
B. If we can see the complete workflow this particular software, first we need to add images, then edit gel images. Editing is very important because during scanning one may have some extra portion of the gel as well as the streaking of the gels at the cathodic end or the anodic end. These all should be removed from the gel

A. Or even come from staining, staining artefacts or even some dust particles on the scanner. Some times water, air traps inside.

B. Air traps in between the gel. So it should be taken care of. So editing is very important, then visualization of the image. Construct matches hierarchically. There is a procedure to create a match set and how to create the class set. There are two differences. This we will discuss in detail in our software demonstration. Then processing gel images where we can detect those spots whether we can edit those spots like how are all the original spots or all of these the dust particles. Or else if there are couple of spots which are coming and showing as a single spot, in that case we can edit that particular spot or we can split that into two different spots and if somewhere if the boundary software itself gives a very nice, if you are not that satisfied with those boundaries you can easily draw yourself whatever boundary you would like to draw. This is what we can do. Then data analysis. Once you did all these things you will match all your gels after that you can get some statistical data like Annova and with the help of these all you can exactly see which are all up-regulated, which are down-regulated, and which are similarly present and these all can be seen in the output.

A. I guess preparing images for analysis that is very crucial because although there are lot of claims from all the software that its very automated procedure but regardless of that like I guess all the software pick-up lot of artefacts even small dust particles particularly when stains will detected as small spots on the gel. So one needs to have both automated as well as manual interference and balance of each to ensure whatever spots are picked up those are the real spots. I guess that's where you mentioned about the spot boundaries and tools which can enable us to remove a spot or add a spot and drawing a right boundary.

B. Here I will go through each and individual parameter.

A. Those will be very critical. I think those will be good to discuss more in detail.

B. In the image analysis workflow, first we will go through add image, how one can add the images to this software. There is an option “image pool” as you can see here. From there you can add any image stored in any area of your computer. Like once you add them you can see that it goes to that particular area. The software can analyse .gel, .mel, .tiff kind of images and very easily can be analysed.

A. So, various types of file extensions can be analysed.

B. Yes, exactly. Except simple .jpg or compressed tiff. These kinds of gels can’t be edited. It requires high-resolution images so that we can have much information with this image. At least image should contain 16 bit resolution. This is the main important thing in the software. During adding your images it will ask you what stain you used for your gel. So we need to mention the colour especially. Within the software we incorporated different stains like coomassie, silver or different fluorescent colours which you have to mention. Once you mention that then you can open your images.

A. Dr. Srinivas I want to ask a question here. So, what is use of defining the different types of stains? Is it going to change certain parameters based on what type of stain we are going to use?

B. By giving this particular staining one can give pseudo-colors for these images. So it helps viewing the images.

A. So, if you have stained with coomassie you can see in the blue background and with silver it is brown and different stains. But its pseudo, right? It has stored tiff images in greyscale and that’s what is going to be analysed and rest of things are just for visualization and getting a feel for the type of the stain.

B. Whatever you are doing like editing of the visualization that will be only for your visualization only. There will be no changes happening in the data. This is major important.

A. I think that's very important for the people to learn that even if they are playing with the image that is not going to change the raw data. That is important. All it is doing is helping in visualising the image and align it properly and helping you in the analysis.

B. During the analysis workflow there are some recommendations, which we can give like use viewing tools like contrast, overview, profile, 3D image to get an idea about the image quality. If once you can go through all these parameters definitely you can decide if this image can be analysed or it cannot be analysed. So once you get an idea of this you can go ahead with analysis.

A. So the user has to do this for each spot per say, right? For example if there's a doubt if it's a spilt spot or real spot one needs to look at 3D profile of it and look at the contrast and try to change different types of parameters to ensure if one is handling a small spot or artefact.

B. Yes, exactly. As we know if any dust particle gives you a spike in 3D view and one direct protein can give you perfect peak sort of thing. So this we can differentiate.

A. So we can take decision based on the 3D view of the spots.

B. Exactly.

How to adjust the contrast in our workflow? You can see there is an option from view there is an option to change complete contrast and brightness you want. In that case you can select any one of the box from the select tool. After selecting that particular box tool, you can change whatever you want from the x-axis or y-axis. The real changes you can see in this particular block. After changing whatever level of changes you want for visualization once you think this is perfect view

then you can press on “apply”. Then the same parameters will be given to all gels.

A. So, if you have a match-set of 20 gels, if you adjust for one it will be applied for all gels.
B. After adjusting contrast and everything, you will see the final image like this. Still there are, you can see, saturation. You can see like red colour. This can be removed.

A. So, these red colour spots are indicating the saturation are actually generated by software. The software is indicating that the contrast parameter is not suitable.
B. There is a zoom option in this. If you can see a particular area then you can see that particular area in a big screen so that you can see if there is any low abundant portion or if there are multiple spots binding together and showing as a single spot. In 3D view as well as you can see like this also. You can actually decide. Especially, manually when one can draw the outline when we are zooming that particular image. So this is very helpful tool again.
A. I guess when you have many gels and you want to look at one particular area, it’s not really possible to look at in a very small image. So you need to have a zoom tool for expanding the image.
B. There is another tool in this software like “profile view”. There you can select particular profile. Wherever you go for a particular spot it gives you profile in x- and y-axes. We can discuss this in detail during our software discussion live demo.
A. So, the x-and y-axes are showing the peaks, right? So that peak is actually for each spot when you click on it.
B. You don’t have to click. You move your mouse on the spot, wherever the cursor goes it shows the particular peak.

This is the 3D view of a particular area. Here earlier we were discussing about a particular spot which we can see as a 3D image but if you want to see a

particular area also you can see as a 3D view. With this you can easily identify the dust particles and the protein spots and this is how it helps you.

Now let us discuss about processing the images. Here first thing is to click on the spot detection. After the spot detection, we need to match them.

A. So, is this process (spot detection) automated?

B. About spot detection we will go in much detail in the next slide. I am just telling you the 3 majors steps here that is spot detection, matching and quantitation. These 3 are the major steps when we can see the spot detection, here there are 3 parameters majorly it will give you. Those are like smoothness. In smoothness it gives you when you set up the lower smoothness over-split the spots like you try to split the spots as many as possible. When you try to give high value it is under-splitting. So one can go through your image and how it is splitting you have to set that particular smooth value.

A. If you are just trying to detect the spot first of all you need to give certain commands and parameters to the software. So these three parameters smooth, saliency and minimum area were defined after playing with the numbers.

B. There are optimum numbers like if you set smoothness between 2-3, that's enough. There is not much variation in this. But whereas saliency is a very critical parameter. It takes the 3D image view and it sees how much slope is there for that particular protein. If you are giving a higher slope then it takes very higher probability. When you are giving the lower value it will take the lower probability.

A. So, higher the value, higher the probability and lower the value, lower the probability. What about minimum area?

B. Minimum area can easily remove the dust particles. Because any protein having minimum area above 5 at least. Whereas we are putting 5 as minimum area most of the times and removes all the dust particles and takes into consideration the proteins.

A. So, these numbers will be more clear when we look at the demo of the software where you can actually show that by changing the numbers what is the real output.

B. This is how it looks like during the spot detection whether it is over-splitting or optimal-splitting, everything can be seen on the screen and you can decide if the parameters are ok. So these red boundaries are software generated and these are all the parameters, which you can see which we were talking about like smoothness, saliency and minimum area. Here you can see they are detected that means they have given some wrong parameters. After that they gave some optimal parameters now you can see it is perfectly detected and perfectly boundaries.

A. So, first of all one would rely that software would detect as many spots as possible in very accurate way and then what ever still is left over, because I can see in the slide of your first image in the left-hand side on the top one there were some spots which were not able to detect by the software. So now those spots can again be detected by manual way by adding the spot. So that is the advantage of having this tool.

B. Even here there is another option that can remove the particular area. There is another option “exclude”. One can exclude particular area in analysis.

A. It means you are interested in certain region of the gel then one can include only that region of the gel.

B. In this experiment may be they are interested in this particular area only that’s why they have included this much data.

This is the 3D view; we can clearly see that the background is automatically removed here because whatever boundary we are putting which you see in the 3D image it is not taking the bottom of the 3D. That is the protein peak. It is taking some portion and leaving some portion and it is creating a particular boundary. So that the total volume it is going to take is above the boundary only.

It does not take below the boundary so that we are eliminating the whole background.

A. So, I think in the slide of the quantification of 2-DE the way it is shown that the green circle which is defining the spot boundary that is the settlement above the background. So we are discovering almost 75% of the maximum peak of the whole spot. Because you can expect 15-20% of the background in the bottom part.
B. So, even we can put this 75% for all spots so nowhere it is going to affect your analysis.

A. So, we can assume 20-25% common for the background for all these spots.
B. In analysis workflow you can see the quantification of 2-DE. Basically there are 4 parameters, which we see here. Background is automatically corrected lowest value in the neighbourhood. Then boundary; it is defined at 75% of the peak maximum. So one can easily remove the background. Then the volume is calculated above boundary level only, where you can see the green boundary above this only we are calculating as volume. Then normalization through relative volume that means the relative volume from all spots this 25% is reducing. So remaining whatever volume is relative volume. This is going to be analysed. In the analysis workflow quantification of 2-DE, one should look at the background of the gel as well as the boundary of the gel, volume and normalization.

A. These are the parameters one needs to look into.
B. Then background is automatically corrected lowest value in the neighbourhood. Boundary defined at 75% of the peak maximum. Then volume is calculated above boundary level so that you can remove the background very easily. Then normalization through relative volume. You can easily normalize your spots.

A. So in this slide what you have shown the green circle, which is defining the spot boundary that is showing 75% of the peak maximum. So it means you are

eliminating the background about 25%. But that should be kept consistent for the whole spots.

B. This is the normalization, which you are doing.

A. Right, that there is no artefact in the analysis and then one can remove the background artefacts only look at the peak volume and 75% of the peak area.

B. Most of the time if it is possible avoid spot-editing majorly. You can do the spot-editing but when you are doing number of gels more and more in that if you editing manually we don't know whether boundaries, splitting can not be done for all equally. Try to avoid this thing initially but if it is necessary then definitely you can do this. Because it introduces lot of subjectivity. This is not supported with DIGE gels. Then use alternatives like composite spots by defining multiple matches and propagate spots from one match set to another. Like once you are matching, these are all not matching perfectly and the same gel you can try to match another time and you can introduce few landmarks so that it is easily matched and detected.

A. So, your recommendation is that one needs to rely on the automated analysis and because the number of gels are more it is difficult to do the quality control checks with the manual way. By defining various landmarks, by repeating the process several times one can increase the precision of analysis.

B. Then if you would like to edit your spot how to edit the spot first. First of all this the way we can split the spot. There is a tool for spot splitting and then you just draw wherever you want to split then it can be created as two spots now. In this way you can edit your spot and as well as lets say you have to do many spots, by this way you can draw your line from both sets so that it can be much together. Another option in spot editing, if you are not satisfied with your boundary and you want to enlarge your boundary, you do not need to delete and re-draw your boundary. Instead you can just mention how much area you want to extend to this spot so that you can expand the volume of your spot.

In the same way if you want to decrease some portion of your spot like as you can see there is a tail sort of thing but you think this may not be your protein portion then one can easily cut that portion and hence can reduce the spot size.

A. So, defining the accurate spot boundary is crucial and different types of software views like 3D views and profile views can actually help one to take that decision of what is the best boundary one can define for the spot.
B. Then matching. Once you have finished your editing process, now this is the time to match whole gels so we can place one or several landmarks to support automatic matching. So if your matching, once you can match. But if you put landmarks, it can be very useful for matching. I will show you images how it effects the matching. Keep landmarks at minimum. If it is possible try to minimize it to 2, 3 or 4. Otherwise it will create some discrepancy in your spot searching.

    A. So I think it will be useful to have landmarks in various regions of the gel so that one can actually align all the different gels by using these landmarks.
    B. Place on well-defined smaller spots, which are clearly corresponding.

    Once you match these kinds of vectors you can see. What does it mean? This is deviating from the reference gel to remaining gels, how it is deviating for that particular protein whether it is one side deviating or it is differently deviating you can see here. It will show you overall trend of the matching. So there is no landmark at all. You can see once we put the landmark how drastically things changed. There is no tail. This is how it helps. So, this is another thing matched with one landmark.

A. So the vectors look well organised and have similar orientation as compared to the ones we have seen earlier where there was no landmark.
B. After matching everything, how to design your experiment here, how to create match set and all, this we will see more in our software analysis. This is how it looks like. From our image pool we have added images and after that we have to create match set, we have to create our class. We will talk about it in detail in our software demonstration. And these are all the analysis workflow, the complete data analysis, once you get the complete data from that what are the flow fold regulations, what are the statistical parameters, then you can count all the parameters and you can select your protein of interest. Shall we start with ImageMaster platinum demonstration?

A. Sure. I guess you gave an overview of the whole data analysis which one can perform by using this software and different types of challenges one may face during the analysis as well as what are the different parameters one can use for proper image processing. But I think it would be really useful to see the software layout and how one can take these images for analysis. So please go ahead and show the demonstration on the ImageMaster platinum software.
B. This is the software layout. Here there is an option like image loader. Go to the place where your gels are stored and add the gels. As I have multiple gels here, I will add few gels to the software to analyse. I am processing 6 gels at a time. These are all 6 gels basically replications of 2 gels. 3 replications for each gel. One is control and another is treated. As we discussed in our slides this is particular thing asking for staining. Here we can give coomassie, silver etc. So as it is stained with silver, I will opt silver and applying the same colour for all the gels.

Now you can see the 6 gels at a time in this overview. You can't see proper spots in all the images. They are very bright images. Now we need to edit the images for visualization. We need to change contrast, brightness. Now more spots are seen.

Otherwise there is another option. You can select a particular area and then change the parameters. Now you can see real time changes in the particular spots. After you are satisfied with the parameters you can click apply so that it can be applied for all gels. So now we can see more spots clearly on all the images. Now we should not disturb in our original images so we need to make copy of these particular images. For that we can crop and create new images. As we have some artefacts in our gels we can remove all those artefacts.

A. So there are some gel migration issues and some gels do not run uniformly you can crop the images and you can store a copy of the fresh image without changing your original image.
B. Otherwise once you are doing your analysis you have changed your original image you may lose your raw data. This should not happen and that is why we are going to crop them. When you are selecting in a particular gel it is going to be selected in a single gel. Instead of doing single selection just press on shift+control and select area which can be common in all gels. So that it can be common in all gels so that you can select at a time in all gels. Then you can adjust them in the same way. After pressing shift+control you can edit all boundaries at a time. Otherwise if you are editing a single boundary at a time there might be image-to-image variations.

A. So, manual adjustment is not required because we can do this automated way so that all the processing will be uniform.

B. This is what we prefer. Then we can crop them using the crop option. And click yes to create new images. It a very quick process. So now we can see all edited images here with extension of (1).

A. So the original images have been stored and created new images with the crop.

B. After that we close all the images and then the edited images are opened. One can create a grid lay on the gel from view option so that one can judge whether all the gels are showing reproducibility or not, in which grid I have what number of spots.

A. Also helps to decide which landmarks to use, also which area one need to look around.

B. This is very helpful tool again. If you want to remove the grid overlay click the same option. Another thing we can talk about is profile. One can click the profile and wherever the mouse goes it shows the particular spot profile, how it is looking. Initially we can see the gel quality. Then you can remove that particular tool.

Now we are going to plan the design of our experiment and how we can manage our gels. From the project we have to create new project. The project here we are doing is the IIT demo1. If you have any description about your own experiment, if you have specific details about the project then enter here and can be stored in this project. So now in this IIT Demo1 one can create the match set here by clicking match set. You have control and treated.

A. So, you have control and its 3 replicates and treatment and its 3 replicates.

B. Then create another match set. Two folders are created and shift the corresponding gels to these folders. Select the corresponding images and just drag them into the folders. Here you see one gel is half red and half green in one folder and it is the same in the other folder. This means this gel is set as reference gel. The remaining two gels are going to be matched against these two gels.

A. Can any gel be made as a reference gel?

B. Yes, exactly. Any gel can be used as reference gel. There is an option in view where you can change.

A. So there is no biasness in this and one can change the reference gel.

B. Yes.

Now display them and you have all of your gels. Once you have the gels we can directly detect the spots. Now the 3 parameters smoothness, saliency and minimum area. If you increase the smoothness it is going to under-splitting. If decreasing the smoothness it is going to over-split. So it is inversely proportional. Its average value is from 1 to 5 but 2-3 is sufficient to start. And then saliency; one can set upto 100-150 but the good value to start is almost all 15-30. Initially I am just giving 15 then minimum area 5 is good enough. If you think in your spot there are more artefacts you may increase this also.

A. So one can accept those values and visualize the spots.

B. Now all spots are detected. If you are satisfied with this particular detection then its fine.

Otherwise you can zoom this particular region and see how the spots are boundaries as well as how the detection is there.

A. Can you apply the same zooming parameter for all 3 gels?

B. Yes, definitely. First come back to original and select a region and the same region is zoomed in all the gels.

A. I think it's very useful to see this way. Now the parameters are well defined because it has properly defined boundaries and there are some spots, which on can look around.

B. Actually you want to edit these particular spots as you can see that there is no protein. But it is showing as a protein one can delete these kinds of spots also. One has to go edit the spot. There is an option "edit enable". You can go there and delete the spot you want to delete.

A. But before we take that decision can we have a look at the 3D profile to ensure that we are not deleting the real spot.
B. Definitely. Click on 3D view.

A. So the spot which we are interested in discussing is highlighted in the green boundary.
B. But it is only present in this particular spot. It is not there in the first image but we can see some portion of this spot in the 3rd image. That is why it is detected there also.

A. It is possible that your treatment may have the protein appearing due to the up-regulation of the protein.
B. But it is not exactly the case here because here these are three replicates of the same protein and hence they should be same. But somehow they are not and so we can delete these spots.
A. Or in this case it is doubtful, right?
B. Or let it be. When we finally see the statistical data where it is there in both gels as well as treatment also so that one can go with the statistical parameters. This will be helpful. So no need to delete also in this case.
A. Now let us look in detail about one of the real spots.
B. Otherwise, I can zoom more so that one can see it. Now this is the intense spot, which you are visualising.

A. And their boundaries are marked from 75% from the top.
B. Yes.
A. Actually we can zoom out and go back to the whole gel. So lets say in a region we have missed out some spots, like software missed out then there is an option to add the spot as well.

B. Let me select a particular area and let me zoom that. Now you can see that particular area in all gels.

A. I think it has defined boundaries over all. So that is why it is better to rely on automated detection.

B. If you want to include a spot wherever you are interested you can add. Like this end of extra portion can be deleted easily. Now we find all spot boundaries are ok and every thing is fine. Now we can match these gel and check for the reproducibility of the three gels. Zoom out to the original position and match them using the match option.

A. So, you are applying the same parameters for control and treatment gels and you are matching them.

B. No, we are matching control gels only to the reproducibility within one group. Suppose you have 5 or 6 replications instead of 3 replications, out of 5-6 replications one can remove that particular gel and they can take the remaining gels.

A. So sometime during various experimental run if there are 1 or 2 gels which are very bad which are going to effect the overall reproducibility by looking at these kind of parameters one can decide may be one out of 6 gels is very bad and going to affect the statistical parameter, then one should remove that. And select the ones, which have good number of matching.

B. Now in these two gels there are few vectors, which you can see. These you can remove by adding the landmarks. Now this is the landmark option. Initially you have added landmark in your reference gel only. After adding the landmark, you have to adjust the landmark to the same position in the remaining gels.

A. I think, now the vectors are improving based on the landmark position.

B. Again, match them.

A. From 146 it has increased to 746 by adding a landmark and our vectors are matching well.

B. You can see there are few vectors but they are still there. These are all moving to the same side. So that means there is no irregularity. One needs to go through each region carefully. So if you can add one more landmark here it can be removed very easily. This is what we can do. And now we can save this thing and apply the same parameters to our treatment gel also. If you can detect your spots automatically it gives same parameters earlier, which you used.

A. Is it a long process of analysing it or rather it is quiet quick.

B. It's quiet. Similarly you can go to each gel by zooming, 3D view and selecting the spots and non-spots and directly match them. 1274 matches are found. These gels are more reproducible as compared to the other set. There are no vectors even with out giving landmarks. So this is very fine gel. The replication with which we did was confirmed. These replications are ok. Now one can analyse in the class analysis between these two classes, like inter-class and intra-class.

A. So between the control and the treatment one can do the statistical analysis.

B. So now I am creating the classes here. Before that I would like to go to the report that is the gel table. These are all the gels, which are present here. By pressing shift+A you can select all the spots and go to the spot table.

A. So what parameters are described here in this spot table?

B. This is the file name, spot Id, pixels, match Id, x-axis, pixel at y-axis, pI, molecular weight, intensity of particular spot and area of that particular spot and volume of that particular spot, percentage volume of that particular spot, saliency of the spots. So these are all the parameters, which you can see in the software. If you are satisfied with these values you can go ahead for next level of analysis that is class analysis. Now I am going to create classes named 81 and 82, which are control and treated. And shift the gels to the corresponding folders. Another match set I am creating so where we can move easily before going to class. So I am moving all 6 gels into composite match set and now create again the same way, classes.

A. So by creating folders one can avoid any carry over mistakes and all the steps are stored.
B. Now I am shifting all the gels to the corresponding folders. This is for class analysis. We can see all the gels together, 3 gels from control and 3 gels from treated. No it is taking values as single gel of 81 and single gel of 82. It will average the first 3 gels and then average the remaining 3 gels. So a single value is obtained from the control and a single value is obtained from the treated. And quickly match them again. One can remove the vectors by adding landmarks, same process as we did earlier.

   Now go to reports where you can see the analysis classes and then go to table. Now we have the match Id. Match Id is a particular spot id as well as the maximum area.

A. For control and treatment, both the groups we had 6 gels, right? So what are the values in each six gels for each spot?
B. This is the averaged value of all the 3 gels from each spot. You can also see the ratio value between the up- and down-regulations.
A. So, this software gives you option for various statistical analysis.
B. You can also see the Annova value. Based on this we can easily select the proteins of our interest as well as you have the fold-change also. We have 1 fold up-regulated and 1.2 down-regulated. You don't have to go to each spot. You can filter them easily. We are going to take maximum Annova value of 0.05 as statistical significant value and so those min-spots can be highlighted. All the spots with this particular annova value are highlighted. This table can be exported to excel and from there you can have your data. This is the output of the analysis tool.

I think it was very useful to see what are the steps required for performing this analysis and also the software gives you lot of options for doing different fold-change and

statistical analysis of how significant the data is. And then one can still go back to the spots, which are significant and look at each spot manually to verify if it's a real spot. Thank you very much Dr. Srinivas. It was very useful discussion and demonstration on the software for image processing and two-dimensional gels. And we will continue our lecture flow for further difference in two-dimensional electrophoresis in the next class.

So the discussion with Dr. Srinivas on image processing by using Image platinum software was useful for you to understand the details required for image processing and how various types of parameters one need to pay attention during the analysis. So in today's lecture in summary we discussed about the challenges for the image processing associated with 2-DE, how one can process the images efficiently, how one can analyse these images stepwise using various software. Most of the software work with the similar principle and to give you a feel for data analysis we showed you a software demonstration during the discussion with Dr. Srinivas from the GE Healthcare by using ImageMaster platinum IMP7. At the end you must appreciate that one need to do different types of statistical analysis to ensure that the identified spots can be further persued for research and those are going to be further processed by mass spectrometry and if they can provide some meaningful information then one need to have full confidence and that's why we need to do statistical analysis.

So I would like to acknowledge Dr. Srinivas from GE Healthcare for discussion on image processing and data analysis as well as for giving us a demonstration of software for 2-D gel analysis.

In the next lecture, we will continue our theme on gel-based proteomics and I will introduce you with a new technology on 2-D difference in gel electrophoresis or 2D DIGE and then we will continue our lectures on 2D DIGE and its image analysis. Thank you!