LECTURE-17

Difference in gel electrophoresis (continued)

Discussion and data analysis

TRANSCRIPT

Welcome to the proteomics course. In the last few lectures we have discussed about gel based proteomic techniques. You have how 2-DE can be performed. There are certain limitations of this technology. So then there are new creative ways, more advance methods have emerged to overcome those challenges. In the last class we talked about difference in gel electrophoresis or DIGE and you have seen how to perform the DIGE experiment in detail.

In the previous lecture we discussed about different challenges, which are associated with 2-DE. I give you few examples, how to prepare sample, different types of reagents or even isoelectric focusing settings. All of those factors can contribute towards a very bad gel. So if you have performed everything properly and all your reagents and the protocols are good at the end you can expect a good gel but often there are various gel-to-gel artifacts and variations in traditional 2-DE method. So we discussed about how to overcome those problems. A new method such as DIGE has emerged.

In today's lecture we continue our discussion on gel based proteomics, certain advantages and challenges of performing DIGE experiments, how to analyze DIGE data by using few very specialized software's, how to interpret that data, how to obtain some meaningful biological information from those analyses. So today I have invited a guest Dr. Srinivas from GE Healthcare who is going to talk to us about DIGE technology and give us a demonstration on software to perform DIGE gel analysis.

An interview with Dr. C. Srinivas Rao, application specialist-Life Science Wipro GE Healthcare.

Discussion on "advancement in gel-based proteomics; 2D-DIGE and data analysis":

A. This is my great pleasure to introduce Dr. Srinivas Rao from GE Healthcare. He is an application specialist in the research product division of GE Healthcare. Dr. Srinivas, thank you for coming for discussion about 2-DE and DIGE technologies.

A. So, what are the major advantages and disadvantages you see by using 2-DE approach?

B. This is very user friendly as well as very low instrumentation cost. The complete proteomic the differentially expressed proteins we can see exactly during this complete 2-DE technique. As well as the novel proteins we are able to isolate from this technique. This we can identify easily, the software's are very user friendly. This is what we can do. There are some disadvantages; the reproducibility will not be there in most of the times. This is the major disadvantage of 2-DE and again software; they are very user friendly. Still there's lots need to be developed so that we can use very effectively.

A. So as you rightly mentioned so 2-DE is still very powerful to resolve thousands of proteins on the gel. One can also analyse isoforms and post-translational modifications including phosphorylation and overall the process is very user friendly. So still it is one of the very powerful technologies currently available for proteome analysis. Now, I also agree with your comment that there are some reproducibility issues because users have to run the gels from the control and treatment separately. And then there will be lot of manual artefacts regardless of how meticulous they are doing their experiments. So in that regard what is your suggestion like how a user can overcome the limitations of traditional 2-DE technology?

B. One should have to prepare very good sample preparation so that the reproducibility thing will overcome and the second point is these days the new generation 2-DE electrophoresis system that is the DIGE technology is available where you can use up to three samples in a single strip and finally there will be

an internal control as well as control and treated. These three can run in a single strip so that these kind of reproducibility issues can be overcome as well as there is a very powerful scanner available in this technology that is typhoon trio which is based on the laser based technology so that you will have very efficient gels in your hand.

A. As you mentioned that DIGE technology is one of the very powerful solutions to overcome several challenges which people face in running traditional 2-DE gels. So in that regard I think if you can elaborate little bit about DIGE technology so that will be useful to the students.

A. I have a small video regarding the DIGE technology. I'll show you that video so that everyone can be able to easily understand what exactly is DIGE technology.

Video - As we were talking about the 3 samples we can labeled with Cy2, Cy3 and Cy5. Now we have two different samples from different sources. Now you are taking these 3 samples into 3 different eppendorf tubes and you are labeling with Cy3 dye as well as Cy5. Then you are taking those two samples, pooling them and labeling it with Cy2. Then your mixing those 3 into a single tube then you are running in a single strip. There you will be analysing the sample in the first dimension; the IEF. Where as in the first dimension after analysing IEF you will be placing the gel into second dimension where you have completely analysed in the second dimension. Now this gel can be scanned in the Typhoon Trio plus scanner where you have 3 images from a single gel. Then this can be analysed with the help of DeCyder software. Then where the complete analysis can be done with the help of DeCyder. In DIA there will be co-detection. The co-detection is nothing but there is a standard gel which is labeled with Cy2 from the same area where expanding remain two images also. So the artifacts can be minimized. Again this is fully automated analysis. There will be very less manual interference in this analysis. After finishing the DIA analysis we can go to BVA where we can compare all DIAs together as you can see in this video. Firstly it matches between the standard gel and from standard gel again to the corresponding DIA. It matches each and individual spot and it will give you the final data. Then these BVAs can be analysed in the further

in EDAs that is extended data analysis. Here you will have lots of stringent statistical data that's differential expression analysis, pattern analysis, discriminant data analysis and PCA analysis.

A. So Dr. Srinivas it was very useful to get the glimpse of the overall process and DIGE technology. Would you like to elaborate or demonstrate some more details about steps involved in doing the labeling of this technology?
B. Yes, there is another video where you can see now. This is the second video which tells you more elaborative. This is particularly a protocol, which is developed for membrane protein analysis. Now you can see this video.

Video - In the cell surface labeling protocol seen on top you label the cells while there are still intact. During the labeling process the dye will only have access to the cell surface proteins. After the labeling step the cells are lysed. To verify cell surface specific labeling the label sample was fractionated into membrane and cytosolic proteins. A known fractionated sample was prepared in parallel for comparison. This fractionation analysis is not necessary but was done just to show that a cell surface protocol is specific for cell surface proteins. You also performed the Standard Ettan DIGE labeling protocol seen below. The cells are lysed before labeling and in this way all cellular proteins are accessible for labeling. After the labeling step the samples are subjected to 2-DE.

Sample preparation - Adherent cells are detached using a non-enzymatic procedure to avoid digestion of the cell surface proteins targeted in this protocol. We use enzymes free cell dissociation media is also an option. Count and divide the cell suspension into aliquots of 5-10 million cells per tube. The cells are then pelleted and washed in HBSS pH 7.4 to remove traces of cell culture media. Contamination from serum proteins and other components can interfere with labeling and detection. Cells growing in suspension are directly pelleted and washed before the labeling step.

Labeling with Cy dyes - After the wash the cell pellet is re-suspended in 200 µL ice cold labeling buffer containing HBSS pH 8.5 in 1 M urea for optimal labeling conditions of cell

surface proteins. Always check pH before labeling. We used 600 picomoles of Cy dye for 10 million CHO cells. The optimal ratio of Cy dye to cell number will vary depending on the cell type. Since we do not the exact protein concentration on the cell surface. How to determine the optimal conditions for Cy dye labeling of proteins is described in the 2-DE principles and methods handbook.

Incubation and quenching

The cell are incubated with Cy dye DIGE flour minimal dyes for 20 min on ice in the dark. After the labeling reaction the un-reacted dye is quenched by adding 20 µL of 10 mM lysine. The labeled cells are now washed twice in cold HBSS pH 7.4 buffer to remove the excess Cy dye. There will, therefore, be no free dye left for unwanted intercellular labeling of proteins in the next step, which is cell lysis.

Cell lysis

The proteins on the cells surface are now labeled and the cells are washed and are ready to be lysed. The pellet from the last washing step is re-suspended in 150 µL cold lysis buffer containing 7 M urea, 2 M thiourea, 4% Chaps, 30 mM Tris, 5 mM Magnesium acetate, pH 8.5 and left on ice for at least 1 hour with occasional vortexing.

Rehydration of IPG strips

The samples are now ready for 2-D gel separation. The first step in 2-DE is to prepare IPG strips for rehydration. Prepare this rehydration solution by adding IPG buffer corresponding to the pH in one of the strips used and add the solution in the lanes of the rehydration tray. Remove the protective film of the IPG strip and place the strip carefully with dried gel facing down in the rehydration tray containing the rehydration solution. Close the lid of the rehydration box and rehydrate the strips overnight.

Isoelectric focusing (IEF)

In the first dimension IEF the proteins are separated according to their pI. The rehydrated strip is placed in the manifold and the electrode is mounted on top. 50 µg from each sample is applied using sample application caps. We have either directly applied non-fractionated samples without prior fractionation or fractionated samples into membrane and cytosolic samples before they were applied. The lid is closed to protect

the fluorescent samples from the light. The instrument was programmed according to the recommendation and run overnight.

Gel assembly
Large 12% lamelli gels were casted using a dual 12-gel caster. The displacing solution was added to avoid polymerized acrylamide gel in the tubing. The gels were allowed to polymerize overnight at room temperature prior to use.

Equilibration of IPG strip
After the IEF the strips are removed and equilibrated in SDS containing buffer in two steps using DTT to reduce the disulfide bonds of cysteine residues followed by alkylation IAA to avoid modification by acrylamide. The IPG strips are dipped in running buffer and carefully placed on the top of the large 2-D gels. Avoid trapping air between the strip and the gel. Seal by adding melted 2% agarose solution with bromophenol blue on top. The gels are now ready for second dimension SDS-PAGE separation.

SDS-PAGE: In the second dimension SDS-PAGE the proteins will be separated according to the molecular weight and this is performed with Ettan Dal6 system. Fill the electrophoresis chamber with anodic running buffer, insert the gels and fill the top compartment with cathodic running buffer. Program the power supply according to the recommendations and run the second dimension protected from light for 4-5 hours or until the dye front reaches the bottom of the gel.

Gel scanning - After the second dimension electrophoresis, the gel cassettes are placed by using the grippers in the typhoon imager. 2 gels and 3 channels can be scanned simultaneously.

Results: DIGE gels
The result from this 2-D gel shows high resolution of membrane proteins in the sample. Even if there are some known restrictions for hydrophobic proteins to be detected in a 2-D gel the results show many new cell surface label spots shown here in red that are not detected using the standard labeling protocol shown here in green.

Analysis
These results show that the cell surface labeling protocol is highly specific for labeling cell surface proteins. Since cells of these proteins exclusively labeled they are easily visualized and attenuation by abundance cytosolic proteins is avoided. The fluorescent gels with known fractionated membrane fraction or cytosolic fraction are shown on top. Below are the images of the same gels co-stained with silver the result show no fluorescent labeling of cytosolic proteins but the silver staining shows that there are proteins the gels. The result also shows similar map patterns known fractionated and membrane fractions demonstrating that there no need for fractionation prior to 2-DE which makes this protocol both simple and convenient. Cy2, Cy3 and Cy5 show similar labeling patterns and all compatible with the cell surface protocol.

Analysis using DeCyder
A DIGE experiment was performed using all 3 Cy-dye DIGE flour minimal dyes for studying different expression of cell surface proteins in CHO cells of serum starvation of different length of time. Cy2 cell surface samples from samples in the experiment are pooled and used as an internal standard. The differential changes of several cell surface proteins could be followed during the starvation period. Over 18 novel cell surface proteins were discovered during the cell surface protocol that were not detected with the standard Ettan DIGE protocol. For finding their identity in the protein preparative gel it was necessary to spike with the cell surface protein to facilitate matching back with the analytical data set.

A. So, Dr. Srinivas it was very useful to see all the steps, which are involved in doing this DIGE technology. There is a discussion going on in the proteomics field that due to advancement in the MS and microarrays. How do you foresee the future of 2-DE and DIGE technology? Do you think it is think it is going to be simultaneously working along with MS and microarrays or do you think it is going to be behind because of more advanced approaches have emerged. So what is your suggestion on that?

B. Yes, the MS and microarrays developed differently. They have their own advantages as well as disadvantages. Like disadvantages like very high instrumentation cost, high protein requirement as well as technical variance should be there that is why still 2-D DIGE will be the core technology to analyse proteomics. One can easily analyse.

A. So, you are still very convinced that by using 2-DE or DIGE approach that will still remain the core proteomic technology for proteomic analysis and it will be used simultaneously along with MS and microarrays. So at the end I would like to request you to give a brief overview of DeCyder software, which is being for the analysis of DIGE gel.

B. DeCyder is the software to analyse DIGE gels where we can analyse DIGE gels in different stages. In first stage we upload our gels in to our software. This basically works on the oracle database where we are incorporating our gels into database. There you can add these gels into software this way and you can select whatever gels you want to upload and once you have uploaded then you can be able to edit these gels here itself like cropping etc.

Now you see this is the overlay of 3 images and now you can crop like whatever portion you want to take. You have to remove the area, which is not the area of your interest then you can directly crop those images. Once you are done with editing you can directly save it. Then you have to create a new project. Your project name would be according to your project. In this case it is Demo. Now, you can import these 2 gel images into the new project, which created. Just click on “import” and these gels will be imported into this particular project. Now here you can see all 6 images. Though they are basically 2 gels, within 2 gel images there are Cy2, Cy3 and Cy5 kinds. So in total there are 6 images. This is the image editor and image loader.

The next one is difference in analysis or DIA. In this you can see we can create a new project using the option “create workspace”. Now it will take you to the place where we already saved our gel in our database.

From the database we can select any particular project and from there I am selecting gel number 1 as we saved this one. Now you can see the gels we have uploaded. So now, after uploading here you can process the gel images. During the process you have to give some number. This is the threshold which you are giving actually this may be 2000 so that it takes care of background issues also. Then basically in DeCyder co-detection happens.

I would like to explain you what is co-detection? Co-detection uses information from all 3 channels and creates a geometrically identical spot boundary for a spot across all the channels. That means there are 3 channels, Cy2, Cy3 and Cy5. Out of these 3 channels inside image it creates particular volume and same area can be applicable for the remaining two gels also. This is the way it works. In this way quantitative and qualitative results much more accurate than in a single detection. In DIA each image is co-detected with its internal control producing two images pairs. The ratio of standard sample is calculated for each protein in each image. So as we see here these are the number of spots, which has been detected in which there are some red color spots which are down-regulated compared with control or treated. And the blue color spots where you can see these are all up regulated spots when comparing with control and treated. In between there are some blue colored spots. These all are similarly regulated. So this is what we can see in DIA. Now you can go through each individual spot and you can view the 3D view of that particular spot. You can go one by one and see if it is a real spot or some background. If it is background you have to remove that. Suppose it is background and there is no spot but it is detected. So you can exclude it from here by clicking and confirming it. So this protein has been removed from the gels.

The same way you can do for each and individual spot and exclude it if it is background. By this way you can check all spots and have more accurate data with you. You see how accurate would be the spot detection. This is what you get in DIA.

Now, we are done with DIA. Now, this is the BVA. BVA stands for biological variation analysis. One of the internal standard images is selected as a master image and all

internal standard images are matched into these samples, standard, spot ratio for each protein, each sample then compared giving T-test value, fold-changes, Anova value for each protein. How to create BVA workspace? You can open the BVA and create a BVA workspace and go to DIA workspace where we have DIAs. From there select DIAs. Minimum of two DIAs are required for BVA. After selecting DIAs click on “create BVA” and a new BVA is created. From here all Cy2 gels automatically go to a standard folder and the rest of the gels remain in the unassigned folder where we need to assign these gels according to the gel types or sample types. Then click “add” and then create a group like control or treated. The first one is control and apply some color draft and then confirm it. Then create another group for treated and give a color draft and confirm it. Now we have two folders control and treated. So as we have in the assigned folder both control and treated, the control gels can be transferred into control folder by dragging those images into control folder. Similarly treated gels can be transferred to treated folder. Now we have our images here. After shifting the control to control and treated to treated. We have to match all gels. Just click on match and the matching process has been finished. As we discussed out of all gels, one gel has been selected as a master image and you will compare the remaining gels with this master gel. Now we have the comparison data.

After that we need to calculate statistical parameters. To do so click on statistical parameter button. Now we have some prior statistical parameters like independent T-tests, average ratio, student-test, one-way Annova in between different groups. We doing between controls and treated and calculate that. So the statistical parameters have been calculated. Now we can see exact results of statistical parameters. You can go to the table view and you can see t-test, one-way annova. We can select from here the statistically significant and which are not significant

After analyzing the statistical data, now we can see the complete results here. Here we can see four views like this is the image view and it is the histogram view. In histrogram view we can see clearly how protein is behaving through out control and treated. You can see this is the standard gel that means this is mixture of control and treated.

And we can consider the standard somewhere at zero and then you see the control is showing up-regulation and after the treatment it is showing down-regulation. And in the table view you can see complete protein data where the t-test value, one-way annova value can be seen.

So you can see four views at a time here. After this we can filter this according to our interest. So select few parameters which are available which spots maps which are present when a particular protein is there and student t-test as well as average ratio then one-way annova value then filter it. So there 2299 spots available in all of the gels but 107 spots passed these parameters. So with these parameters we can select proteins of interest and assign peak-list so that these proteins can be saved in a file. Further this file can be given to the spot picker.

These are all the things in BVA we can identify. This is very user-friendly and there not much of manual interference. This helps you to analyse your DIGE gels.

A. Can you elaborate what is EDA or extended data analysis and what it can do which we are unable to do in BVA. So there is layers here. right? One is DIA followed by BVA and then ultimately EDA.
B. Basically, what we can do here is we can compare two BVAs directly. Here we are talking about a particular disease or particular set of data only. There we can analyse different BVAs together in EDA. There you will get majorly differential expression, PCA and discriminant analysis. These kind of statistical data you get in EDA. I will show you briefly.

A. If I understood you correctly, probably the statistical parameter is more stringent over there in the EDA where you can have better biological significant information from the data set. Because in lot of clinical data and different types of treatments people like to do several gels and lot of treatment. So your number of sample to be analysed is very large. And

obtaining the meaningful data is one of the major challenges in proteomics field. So I would like to know about EDA.

B. In EDA the first thing we can see here is differential expression analysis. In which you can see differences in between controls as well as two experimental data, different treatments, which are given. Here you can see how particular proteins are expressing through out these controls and pf and pv. You can see this kind of data for each and individual protein so that from here you can see which one is of your interest?

A. So you are actually analysing the data spot-wise now?

B. Exactly, we are looking spot-by-spot. Even the index number shows that there is a master gel. From that master gel you can exactly see this number. This is what we are seeing for each individual spot. Here you can see result as well as you can see principle component analysis for this data. Here you can see 89 proteins. Out of these 89 proteins as you can see many proteins are present inside the circle and some proteins, which are outliers, are present out of the circle. In the circle there are about 95% statistically significant proteins and the outliers are non-reproducible spots or they might be highly up-regulated or highly down-regulated. So these can be used as marker also. Then we have to go back to our BVA data and check the data for the regulation of the protein. Then we can identify the protein and use for further analysis.

A. So this is a powerful statistical parameter and by using it you can identify some outliers which can be the potential discriminator between the control and treatment. And once you identify those proteins then you can go back to your original data in BVA and get all the analysis done. This is very interesting.

B. Next is pattern analysis. You can see the whole proteome and how they are all different from each other. This is the heat map made from the total of 82 proteins, which we took into consideration. And you can see in which area they are up-regulated. The green area is completely down-regulated area and red area is up-

regulated area and the black area represents similarly regulated proteins. This kind of data helps you to represent your whole analysis.

A. So Dr. Srinivas it was very helpful to have you here and to get an overview of DIGE technology, how people can use DeCyder software and analyse the data using DIA, BVA and EDA. Although there was no enough time you have given very good demonstration in a very short time to give the glimpse of the processes involved in doing this analyse and how different types of statistical parameters can be applied to get some powerful statistical information from biological data. Thank you very much for coming here and giving a very good introduction of DIGE technology.

So I hope our discussion with Dr. Srinivas was useful and now you can perform these analyses by using specialized software and obtain some very useful biological information from your dataset. Probably you must appreciate there are lot of meticulous steps are involved in performing these experiments but in the end this provide very useful quantitative multiplexing approach for separating proteins and to analyze different types of variations. So in summary, in this module we have discussed about gel-based techniques, we have talked about basic separation by using SDS-PAGE, BN-PAGE and then how to perform IEF and by combing those two methods how to perform 2-DE. I give you workflow for performing 2-DE then we have discussed about different types of challenges for this method and I briefly introduced few creative ways of performing gel-based proteomics method, including OFFGEL fractionation methods. Then we discussed in much more detail about difference in gel electrophoresis. I hope at the end of this module and this lecture you will be able to perform gel-based proteomics experiment. But please keep in mind these protocols and methods are only for giving you feel for performing these experiment. Each experiment, each sample, each biological question, brings it's own unique challenges and depending upon those conditions and the sample types you need to optimize the methods. There is no one technology, which can answer all of your questions but it's a good idea to know what are different methods, which are available for you to use. I hope by taking this module on gel-based proteomics, now you are familiar with different types of gel base

techniques. These are very few, there are many other methods available but these are the most commonly used methods which people are applying in the field of proteomics. Thank you!