LECTURE-18

Applications of two dimensional electrophoresis

TRANSCRIPT

Hello. Welcome to the proteomics course. In today's lecture we talk about applications of 2-DE. In the last lectures in this module we discussed about 2-DE technique. We discussed in much detail the workflow, how various steps are involved in performing 2-DE and DIGE. Today we will talk about how these techniques can be applied for understanding some of the biological questions. This is time for us to assimilate all the knowledge, which we have acquired during the last several lectures. For example what are the things we need to take care of during protein extraction then isoelectric focusing, second dimension protein separation followed by gel analysis. All of this knowledge actually integrates when want to apply these techniques in any biological applications

So today I will just first give an overview of 2-DE technique just to refresh you about various things, which we discussed in this module. Then I will give you few example applications. We will take 2 case studies. This 2-DE technique is now widely applicable in almost all the areas of biological sciences. You name any application you can there is some published reports on using 2-DE or DIGE to apply for such applications whether it is looking at various human dieseases, looking at drug response or investigation about plant stresses or any drug treatment in bacterial species or any other microbes. So as I said you name any application and you will see there are several published reports where people have used these techniques and try to look for different types of proteome changes. Now what you can look at the differential protein expression, which is most commonly used if you want to compare control with a treatment. You have certain level of exposure to a disease or a chemical or a drug and you what to know what are all the proteins that have been modulated due to that particular stimulus or treatment and this is known as differential expression. Now if you are interested to investigate all the proteins in a given sample that will be considered in the global proteome profile. So

many times if the proteome of a given species is not known then proteomics techniques have been used for global proteome profiling. But most commonly they have been used for differential proteomics. So today we will discuss 2 papers where authors have applied differential proteomic analysis by using 2-DE.

So before I discuss the application let's just refresh whatever we have discussed in the previous lectures on various gel-based proteomics techniques. This slide reminds you that we have talked about different approached using the gel-based proteomics including SDS-PAGE, 2-DE and various advanced levels of protein electrophoresis including difference in gel electrophoresis. Then data analysis that becomes common for all of these gel-based approaches and that becomes actually very challenging aspect of it. So performing the proteomics experiment is not so tedious as compared to analysing the large amount of data, which is generated from these techniques and most importantly how to get some meaningful information from these experiments. Therefore data analysis and statistical considerations are very important parameters for obtaining very meaningful information.

So, first of all you need to extract the protein from the given sample and that actually becomes very determining factor for various applications. Regardless of questions you have in you have in your mind which you want to apply using 2-DE, first of all you need to make a very good protein extract. When I say a very good extract it should have minimal interference with any interfering components. Ideally you want to separate as many proteins as possible. It means you want to get a very high coverage of the whole proteome. So considering the diversity of the samples which one can apply on this techniques it comes very important that people should first standardize very good protein extraction. Once protein extract has been obtained and you have determined the quantity of the protein then you can apply on the first dimension, which is isoelectric focusing or IEF. So in this you can separate the proteins based on their pI values as shown in this slide here. You can use different lengths of IPG strip or immobilized pH gradient strip and separate the proteins in the first dimension based on their isoelectric point.

After dimension separation if you remember you need to do some treatment, which is equilibration you want to prepare your strips for the second dimension separation. In equilibration there were two components, which were very essential and if you recall we talked about DTT and IAA. After doing the equilibration you can separate the proteins in the second dimension based on molecular weight of the proteins. As you can see in the slide you can separate the protein in two dimensions based on their properties like their isoelectric points and molecular weights.

This is a representative of 2-DE gel image which shows that large number of proteins can be separated on these gels and immediately by looking at the given protein spot one can obtain two information, what is its molecular weight and what is its isoelectric point. This information can be obtained immediately and then after comparison of the gels one can know what are the proteins, which are differentially expressed.

As I discussed in much detail in fact in a complete lecture earlier the data analysis is a very crucial aspect of 2-DE experiment. You can run large number of gels but when it comes to compare these gels it takes lot of time. You have to be careful to obtain statistical significant spots, which could be meaningful for that biological application.

Having discussed the broad view of the 2-DE now let's move on to the case study. In first study we will talk about serum proteome analysis of vivax malaria and insight into disease pathogenesis and host immune response studied by Ray et al. So this study was performed by my group and we tried to investigate the host response against the malarial pathogen vivax.

So if you look at this map of world population overall Asia and Africa are mostly affected due to malaria. Asia is suffering from challenges of drug resistant strains of malaria and Africa is still the single largest cause of death, which is due to malaria. But it is not only in Asia and Africa it is an epidemic in 103 countries around the world. More than 300 million cases of incidence are recorded per year and leads to the death of 1.1-2.7 million people a year. These stats give us an indication on the seriousness of this

problem. Among the various species of malaria *Plasmodium vivax* and *P. falciparum* account for 95% malarial cases worldwide.

So the estimated clinical cases of vivax malaria ranges from 70-390 million, excluding Africa. *P. vivax* accounts for more than 50% of all the malarial cases worldwide. But the morbidity associated with this disease and the spectrum of infection still remains neglected. Nearly 2.6 billion people around the globe are at the risk of having *P. vivax* infection, even slightly greater than the risk of *P. falciparum* infection. Although *P. falciparum* infection is historically regarded as benign but recently the severe and the fatal incidence of vivax malaria has been reported from different regions of the world which rendered this clinical paradigm clinical deceptive. From the complex life cycle of malaria now investigating how *P. falciparum* and *P. vivax* behave differently in human and what are different types of proteins which are elucidated due to the response against these pathogens. These are some of the very challenging questions in the malarial field.

This slides gives you various species of malarial parasites, which cause different types of malaria. *P. vivax* causes benign tertian malaria, *P. falciparum* causes malignant tertian malaria, *P. ovale* causes ovale tertian malaria and *P. malariae* causes quartan malaria. In addition to these *P. knowlesi* can also cause acute and severe illness but mortality rates are quiet low for this species. So there are different species of *Plasmodium* but *P. vivax* and *P. falciparum* are the prime targets for investigation because as we discussed earlier more than 95% cases worldwide are due to these two species.

This slide gives you the overview of the workflow, which we followed to investigate the host response against *P. vivax* in human serum. Serum samples were obtained, proteins were extracted and the protein samples were compared from the healthy and the vivax patients by applying gel-based proteomics including 2-DE and DIGE. We separated large number of proteins on the gels and looked for those proteins, which are significantly differentially expressed in vivax as compared to the healthy controls. And

then those proteins were subjected to mass spectrometry analysis. Further we validated few targets and analysed the data for various types of pathways which could be involved in such disease.

So before we move to various experiments performed in this study let me give you an overview of the workflow involved in this experiments.

Protein extraction from human serum - Carefully withdraw intravenous blood into a vaccutainer tube and store the tube on ice to allow the blood to coagulate. Centrifuge the contents to separate the coagulated blood cells and clotting factors from the serum which forms a clear supernatant. Collect the obtained serum in a fresh tube and then sonicate the serum proteins to break down any large protein complexes. Serum contains several proteins in a wide range of concentrations of which albumin and immunoglobulin G are the most abundant. Presence of such high abundance proteins can interfere with experimental analysis while trying to detect other proteins at lower concentrations. It is therefore preferred to remove these high abundance proteins before electrophoresis. This is done with the help of commercially available depletion columns, which bind only the high abundance proteins on to their matrix through affinity interactions. Once the serum has been processed using a depletion column, precipitate out the remaining proteins using trichloroacetic acid and acetone.

Quantification of proteins - Prepare the Bradford dye reagent and label the tubes suitably for standard and test samples. Add the standard and sample solutions to their respective tubes. Then add the Bradford colour reagent to the tubes and mix thoroughly. Adjust the zero reading on the UV spectrophotometer using a blank solution and then measure the absorbance of all standard and unknown protein samples at 595 nm. Protein concentration of the unknown can be determined from the standard plot based on the absorbance value obtained.

IPG strip rehydration

Clean all the apparatus thoroughly before use to avoid any contamination. Remove the IPG strip from its cover and place it carefully in a well of the rehydration tray. Then add the reconstituted protein sample on to the strip. After around ten minutes, pour some mineral oil over the strip to prevent the strip from drying up. Cover the tray and leave it overnight to allow rehydration to occur.

Isoelectric focusing

Cut the paper wicks required for focusing to a suitable length and wet them with a small amount of water before use. Initialize the instrument and clean the surface well with a dust free cloth. Place the focusing tray on the instrument and ensure that it is properly balanced. Carefully remove the IPG strip from the rehydration tray and drain out any excess oil by blotting it on a tissue paper. Place the strip in the focusing tray and immerse it with mineral oil. Then place the two wicks at either end of the IPG strip followed by an electrode at each end. Fill all the adjacent wells with mineral oil to ensure uniform current flow. Then input the desired protocol on the instrument software along with details of strip length, pH range and number of strips and start the focusing process. The voltage-time curve will appear based on the protocol that has been set. Protein separation occurs on the basis of net charge on the protein. Proteins will migrate along the strip and come to rest at a point when their net charge becomes zero, known as their isoelectric point.

SDS-PAGE, which constitutes the second dimension of two dimensional gel electrophoresis, involves assembly of the gel apparatus, gel casting, equilibration of the IPG strip followed by placement of the IPG strip on the gel and protein separation.

Clean all the components of the gel assembly thoroughly with water and dry them with paper towels. Carefully arrange the glass plates in the gel casting assembly interspersed with separator sheets, depending on the number of gels to be run. Prepare the gel casting solution consisting of Acrylamide-bis-acrylamide, Tris-chloride, SDS, APS and TEMED and pour it with the help of a funnel into the central channel of the casting assembly. Ensure that the solution spreads evenly such that the top edge of the

gel is uniform. Allow the solution to stand for polymerization to occur and spray a solution of SDS across the top to ensure that the gel does not become dry.

After second dimension run, process involves removal of the gel from the electrophoresis assembly followed by treatment with a fixing solution, staining solution and finally the destaining solution.

After watching this animation now you must be clear about various steps required to perform this experiment. In clinical studies, first of all, obtaining the clinical samples and getting the right type of clinical samples is a very challenging question. To obtain the sample one needs to get the ethics committee approval, scientific community has to look at the proposal and review and approve it and then the patients which fall under the category which you want to investigate have to be enrolled after taking their consent and ethically it should be approved. So various types of information from such patients are tabulated. And then these patient samples can be compared with the proteomic analysis. In addition to the age, gender, physiological status, whether patients have any alcoholic background, previous history of diseases or they have taken any treatment. In addition to all these various types of clinical tests, which are performed they are all tabulated which eventually helps them analysing the data when you want to compare the proteomic response for various patients. Selection of healthy control becomes critical because you want to compare healthy with the disease samples. Now if your healthy sample is suffering from some disease that might give you some artifacts. So that's why you need large number of clinical samples. Then there are issues if somebody wants to analyze only individual samples of the clinical patients or they want to pool and then analyze. Consider you have 200 cases and 200 controls. Now running individual gels for all of patients and controls can be very challenging. And then finally running the gels in triplicates and analyzing the data becomes very tedious. So there's discussion going whether for such large number of clinical cohorts whether group of patients can be pooled and analyzed. For example, when analyzing 200 samples then you make pool of 10 patients then you have 20 samples. But it all depends on the investigators and the kind of experiment design which people have on their mind for

performing such studies. Then sample collection process becomes very important, if you recall the previous lectures how different types of storage conditions can also affect your proteome. So all these parameters one needs to take into consideration.

So, similarly we also tabulated various information for different vivax patients and healthy controls. This slide gives you an overview that before even starting the proteomics experiment, lot of thought has to go into your experiment like experimental design and what type of samples one has to collect, and then you need to obtain various clinical information for those healthy controls and your patients.

As you have seen in the animation, first of all, you need to collect the blood sample then serum has to be separated from that and after serum is collected you need to store that serum in small aliquots at -80°C. So this slide gives you various steps involved in this process. So ideally 5 mL of blood is sufficient for getting serum to perform such proteomic investigations. 5 mL blood if you collect and keep it on ice until the serum isolation. Allow it to clot for an hour and then centrifuge at 20°C for 10 min followed by separation of the supernatant, the serum, immediately. Collected serum can be stored at -20°C or -80°C. -20°C if you want use that immediately for your further use or if you want to use that after long time you can store it at -80°C.

Serum poses challenge of abundant proteins. So first of all we have to remove the abundant proteins from the serum and there are various standard columns available from various commercial manufacturers from which one can remove selectively very high abundant proteins such as serum albumin, IgG and various other abundant proteins. In fact there are columns, which can deplete almost 14 other abundant proteins in the serum.

What sample should be ideal for such proteomic sample? To address this question we first of all tried to separate various proteins from different types of samples. We took crude serum. Because serum will have mostly the protein content, so we thought can we directly apply the serum on IPG strip, separate the proteins and see how many proteins we can separate on 2-DE gels. But serum contains lot of salt. So IEF itself

becomes challenging if you have crude serum. So whether we remove the salt content of it and after the desalting if we do the IEF. Since serum contains high abundant proteins, these commercial columns which can deplete these abundant proteins selectively, those quiet costly and considering that large number of patient sample we need to process, we thought can we disrupt this abundant proteins and remove these low abundant proteins from these abundant proteins by using sonication. So effect of sonication and as well as desalting was also investigated and then also used commercial columns to remove the abundant proteins followed by the desalting step.

Now sample was collected and processed with all these four variables, which I taught you and after that standard procedure of 2-DE was followed. When you are doing this serum protein separation the IEF becomes very critical because you have lot of salt in it. And if your salt removal is not effective your IEF run may not be very smooth. So from the software you can monitor how your IEF is progressing.

So we tried various types of effects, crude serum alone, sonicated or desalted, desalted alone, depleted and desalted and then tried to compare the effect of these types of processing on overall proteome coverage.

We also tried to see how many proteins we can separate on the small gels and then as expected small gel of 7 cm couldn't provide us very large number of proteins. But these gels can be used for standardization process. As you can see starting from the crude serum to desalted and then finally sonicated and desalted gives us better coverage of the proteome. We tried the same with the larger strip, 11 cm and again it was quiet convincing and then we moved one to very large strips, 24 cm for the clinical studies.

In general we took 40 µL of the serum which was precipitated with 4 volume of ice cold acetone containing 10% of TCA, incubated at -20°C, centrifuged and then 1 mL of ice cold acetone to wash the precipitate. Then incubated it on ice for 15 min and again centrifugation was performed. Acetone containing supernatant was removed and then pellet was dried in the lysis buffer. The recipe for lysis buffer was told earlier when we

discussed about sample preparation, which is also mentioned in the slide here. You can use urea, CHAPS, IPG buffer, DTT and BPB.

So from the previous study we looked at replicate gels and as three here and we found desalted, sonicated and depleted samples gave us largest number of spots on the gel as compared to crude alone or desalted alone or desalted and sonicated. So these studies were performed on small strips to standardize the procedure but variable to draw conclusions based on the reproducible pattern obtained on these gels and then we applied the desalted, sonicated and depleted conditions for processing all the samples.

After processing of these samples, now how to extract the proteins? So protein extraction protocol was also optimized and to optimize this protocol first of all we looked at the available methods in the literature. So people have applied crude serum, TCA-acetone precipitation, sonicated desalted serum, trizol extraction method as well as abundant protein removal and TCA-Acetone precipitation.

We used this modified procedure of TCA-Acetone precipitation where we used depleted, desalted and sonicated serum. We also compared the effect of acetone or TCA-Acetone precipitation. As you can see in the slide the left gel is showing the protein overall present in the acetone alone treatment and right gel showing the TCA-Acetone precipitation. So overall more number spots were resolved on the TCA-Acetone gel.

Since we standardized our procedure on small gels earlier we tried to apply those conditions on the large gel of 24 cm because that's where we have to perform all our analysis. So we again compared crude versus depleted response on the large gel and as expected the depleted serum sample showed more number of spots as compared to the crude serum samples.

We also tried to look at the effects of various types of staining whether it is coomassie blue staining or silver staining. We definitely identified more number of spots from the silver staining. But just because of we want to compare various large number of clinical samples with the different healthy controls it was difficult to standardize conditions to

keep it uniform staining for the silver. So to overcome this limitation we used the colloidal coomassie and biosafe coomassie staining and we standardized the same conditions for all the gels.

This slide gives you an overview of various parameters, which we investigated from different types of serum, crude serum, sonicated and desalted, TCA-Acetone precipitated and trizol extraction methods. We looked at the effect of sonication, desalting, rehydration, amount of protein loading, type of strip, staining solution, how many number of spots we can resolve on these gels and then since software analysis is automated gives artifacts how many real spots we can obtain after the manual refinement. I am trying to give all the details what one need to actually perform before reaching to the stage where you can compare the healthy controls with the clinical samples and this lecture is trying to cover various types of concepts which we have talked in the sample processing and different lectures of 2-DE. After doing lot of comparative analysis and standardizing the protocol for sample preparation and protein extraction, now we were ready to perform the comparative study on vivax serum and the healthy controls. This is one of the gel images showing that what is expected on 24 cm large 2-DE gel. And then the representative gel showing the healthy control and *P.vivax* treated samples. So these samples as you can see were showing large number of protein spots resolved on these gels.

And then after separating the protein samples from these 2-DE gels from various patients and controls then most challenging task was the data analysis. Automated matching and manual steps were performed to obtain the how many spots are significantly modulated due to vivax infection. As you can see in the slide we have looked at various controls and various vivax patients. These are only few patients samples here but we analyzed large number of samples. But those spots which we reproducible in all the patients, for e.g. you can see in the top panel two spots are showing significant alteration in the vivax as compared to the healthy controls. These types of spots were considered further for the mass spectrometry analysis.

Now those spots which we found were showing good differential expression we looked at the 3D views of those spots to ensure that those are not artifacts and those are the real spots showing up- or down-regulation of these proteins in human serum.

After completing the analysis we were able to obtain large number of proteins, which were differentially expressed and as you can see in the graph there are proteins in the graph which vary in different range of the fold change. Many proteins show very less fold change between 1 to 1.5 fold and there are few proteins, which show more than 10 fold up- or down-regulation. These protein spots were conceded interesting because those were statistically significant and then these spots were further analyzed using mass spectrometry techniques.

This slide gives you an overview of tandem mass spectrometry followed by how LC-MS/MS data can be analyzed by using Mascot search engine. So different types of mass spectrometry techniques are available including MALDI-TOF/TOF and LC-MS based approaches and we will talk about these techniques in much more detail in the next module when we talk about mass spectrometry.

Just to give you an overview here that one needs to analyze the spectra and one needs to have standard proteins to ensure that correct spectra is obtained. So you have to do lot of quality control checks and mass spectrometry to ensure that your data is of good quality.

I will cover the details of mass spectrometry experiment when we talk about mass spectrometry in more detail in the next module. And again may be I will take this study at that time that what are the different detailed procedure required to perform such analysis.

But at the end from these experiments of mass spectrometry one can obtain the identity of these proteins and as you can see in the slide that there are several proteins in the host human were differentially expressed due to the *P. vivax* infection. These are some of the proteins listed here such as haptoglobin, apolipoprotein A-1, serum albumin

precursor and clusterin precursor these were down-regulated and serum amyloid A, ceruloplasmin precursor, leucine rich alpha-2-glycoprotein precursor, alpha-1-antitrypsin precursor these proteins were up-regulated among many other proteins.

So conclusions from this case study, that few differentially regulated proteins were identified in this study including ceruloplasmin, haemopexin, alpha-antichymotrypsin, alpha-1-antitrypsin which have not been previously reported in the vivax malaria pathogenesis. A very important role of few proteins such as serum amyloid A and P, haploglobin, apolipoprotein A-1 and E was established in the vivax malaria. This study finally enhanced our understanding for the basis of the disease pathogenesis and provided few potential targets, which could be used for further investigations.

Today we talked about how 2-DE technique can be applied for various biological applications. Although I thought I could cover two case studies on malaria one on the host side, like how response can be monitored due to the infection of *P. vivax* and then the second study which I wanted to take was based on the effect of the drug on *P. falciparum*. So we could not discuss the other study on the drug treatment of the malarial parasite, which we will discuss in the next class. But by giving you a case study and details of various experiments one need to perform probably you got an idea that before performing the final experiment on your cases and controls lot of optimization has to be performed and starting from sample processing and how you can expand the coverage of your proteome these kinds of quality control checks are very essential. If you can increase the overall proteome coverage you can separate over 1500-2000 proteins on the gel reproducibly then you have good chance of identifying various potential targets which could be novel drug targets or potential biomarkers. But if you have not done the quality control experiment you have not optimized the protocols properly and your proteome coverage is poor on the gel then you are comparing only partial proteome and there's a good likelihood that you will miss out many important changes. So not only we discussed today about one application how host serum proteome changes due to one of the *Plasmodium* parasites but also we discussed various experiments one need to perform to achieve such kind of comparison. We will

continue our discussion on some more applications of 2-DE and also two-dimensional difference in gel electrophoresis because of the variations in the 2-DE people also try 2D-DIGE experiments and also that is very sensitive so it will also expand the overall proteome coverage. So we will continue the discussion in the next class. Thank you!