

**NATIONAL PROGRAMME ON TECHNOLOGY
ENHANCED LEARNING
(NPTEL)**

**Applications of Interactomics using
Genomics and Proteomics technologies**

**Course Introduction by
Prof. Sanjeeva Srivastava**

MOOC-NPTEL

**Applications of Interactomics using
Genomics and Proteomics Technologies**

**Lecture-20
Basics and Applications of Reverse Phase Protein
Arrays-III**

**Dr. Sanjeeva Srivastava
Professor
Biosciences and Bioengineering
IIT Bombay**

(Refer Slide Time: 00:26)

MOOC-NPTEL

**Applications of Interactomics
using Genomics and Proteomics Technologies**

LECTURE-20

**Basics and Applications of Reverse Phase Protein
Arrays-III**

Dr. Sanjeeva Srivastava
Professor
Biosciences and Bioengineering
IIT Bombay

MOOC-NPTEL Applications of Interactomics using Genomics and Proteomics Technologies IIT Bombay

Welcome to MOOC course on applications of interactomics using genomics and proteomics technologies. In continuation of Dr. Ummanni experience lectures where he talked to you about basics of reverse phase protein arrays.

We are continuing to discuss about this novel technology platform RPPA. In today's lecture where Dr. Omani is going to mainly focus on RPP analyzer which is a tool for RPP data analysis and visualization.

So let me welcome again Dr. Ramesh Ummanni to continue his lecture and discussion on reverse phase protein arrays.

(Refer Slide Time: 01:10)

Dr. Ramesh Ummanni
**Scientist, CSIR-Indian Institute of
Chemical Technology, Hyderabad**



Lecture outline

- RPPAnalyzer - tool for RPPA data analysis and visualization
- Absolute quantification by QconCAT-RPPA
- Applications of RPPA



Dr. Ramesh Ummanni: I will show you what kind of data we have generated, so here let's say for example these are the different samples, these are the different mutations in the cell lines, so you can see that this is the data we could generate, so in this corner on the drug resistance can be incorporated here, drug resistance can be incorporated so this kind of numbers you will get, so these are the target names, (Refer Slide Time: 01:49)

The screenshot shows a Microsoft Excel spreadsheet titled "RPPA data sample file - Microsoft Excel". The spreadsheet contains a large matrix of data. The columns are labeled with various numerical values, and the rows are labeled with text descriptions of samples and mutations. The data is organized into a grid format, with a header row and multiple rows of data. The spreadsheet is displayed on a Windows operating system, with the taskbar visible at the bottom showing the time as 12:03 PM on 24-02-2018.

so your complex matrix will be like this.

So from this if you load this data into commercially available software, so as per predesigned analysis part you will get the data output, but you cannot managed to get information what you want from your samples according to your experimental requirements, so when you have such a high end data set,

(Refer Slide Time: 02:13)

Ramesh Uthappa
Chemical Biology
2/24/2018 | Page 27

RPPAnalyzer - Tool for data analysis and visualization

RPPAnalyzer Tool box workflow

raw data quality control, data normalization

box plots

differential regulation of protein expression

signaling dynamics

fitting mathematical model to dynamic data - systems biology

heat maps

cluster analysis of proteome profiles annotation of heat maps

Wenzinger HA, et al. *Bioinformatics*. 2010 Sep 1;26(17):2202-3
 Kishin S, et al. *BioTechniques*. Vol 57 No 3 pp 125-135

Ramesh Uthappa
Chemical Biology
2/24/2018 | Page 29

Normalization of RPPA data

Loading control (LC) and variance stabilization of reverse-phase protein array (RPPA) data have been challenging. Different normalization methods were developed for LC and variance stabilization.

- Super curve normalization
- Total protein normalization (The loading control)
- variable slope normalization
- Robust Z normalization (Z')
- House keeping gene normalization
- Invariable protein set normalization (Inv)
- Global median centering

PLoS ONE 9(12): e97213. doi:10.1371/journal.pone.0097213

PLoS Inform. 2014; 13: 109-117.
 PPTSL

then you need to worry that how experiment is correct, how much experiment is correct or not, so one of the sensitive point that Dr. Sanjeeva touched in the morning very nicely and gently suggested you to follow this normalization methods particularly in RPPA data normalization is

very, very difficult because when you do the protein estimation and when you extract all the steps cannot be implicated for thousand samples in the same manner, it is very, very difficult if you do 10 samples we will have 10 different values in terms of protein concentration, in terms of method you followed and all, so to avoid that there are different normalization methods have been developed to apply for RPPA data, so loading control and various stabilization of RPPA data have been challenging, so different normalization methods were developed for LC that was loading control and variance stabilization.

One among them is a super curve normalization, total protein normalization, the variable slope normalization and robust Z curve normalization methods, one of the method what we follow in western blot is housekeeping normalization we all favorite of doing gap DH more than the target proteins very often, in variable protein set of normalizations, in variable protein set of normalizations means in all the thousand samples we know that one particular protein concentration will not change, so that will set as a landmark.

I think some of you might have done, sorry 2D gels in the proteomics approach in the earlier, in the beginning days, so we use to set landmark spots, so in that way those proteins expected to not change across sample set will be set as a landmark spots based on those concentrations we can normalize, how a normalized data look different from the un-normalized data you can clearly see that these bars, each bar represents one protein across the controls and treatment conditions then you can get a heat map like this, but after normalization so there is no difference within the samples and nicely you can get a spread of the signals here.

So if you don't apply any normalization method you end up identifying the more differential express proteins across the sample set you have analyzed. So how we can do?
(Refer Slide Time: 04:35)

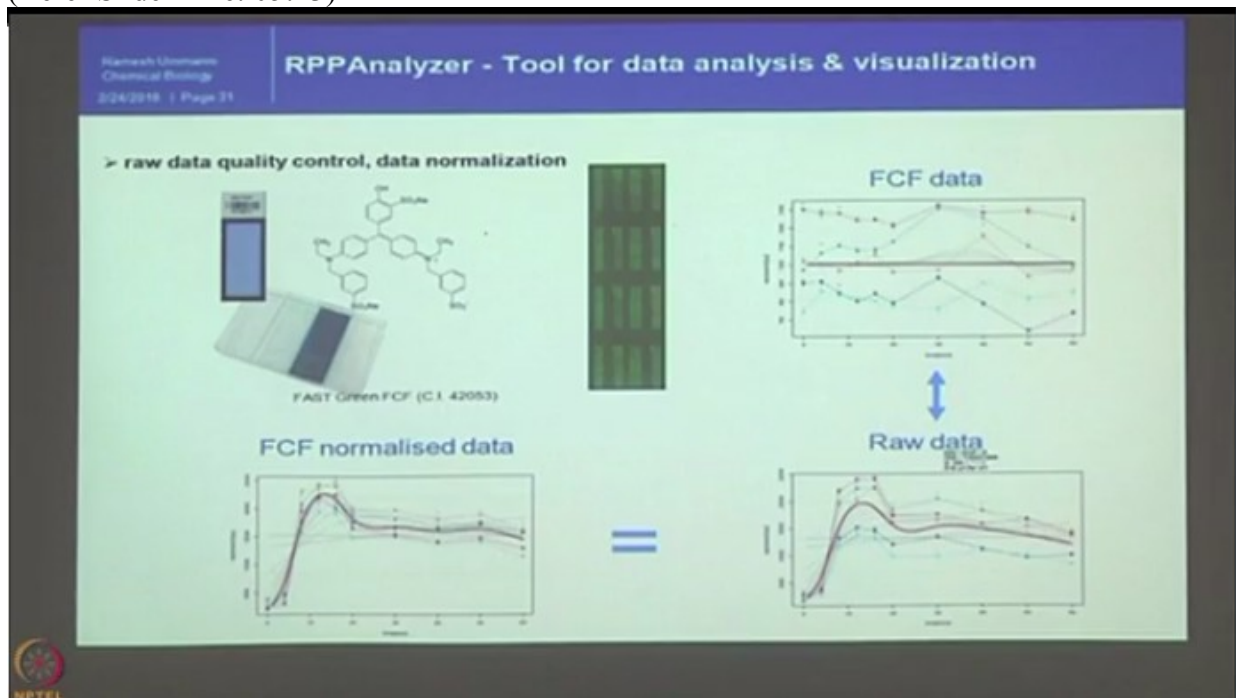
Normalization Using Total Protein

- **Why normalization?**
It is essential to account for differences in total protein concentration between each sample so that antibody staining between each tissue sample on the array can be compared directly.
- **How do we do it?**
One slide/printing is stained for total protein using a total protein stain such as Sypro Ruby™ blot stain or a colloidal gold stain and visualized on a Fluorchem™ imaging system (Alpha Innotech).
- **Reproducibility**
Normalization is based on the total protein per microarray spot. Normalized intensity values are calculated by dividing the measured intensity value of the antibody by the corresponding measured intensity value of the total protein. This allows normalization of sample to a known analyte, maximizing reproducibility.

Ramesh Unnani
Chemical Biology
2/24/2018 | Page 30

Among the methods previous slide explain at least we use to follow total protein normalization method which is more easy to do and also reliable to do, so it is, how do we do that? Always we print lot of slides, one slide after every 10 slides will be used for this normalization methods, every 11 slide, 21st slide, 31st and 41 slides will be used as a normalizing slides.

This will definitely give the reproducibility, this normalization based on the total protein per microarray spot, how we do? So we extend this line with fast green FCF, this is called FCF, so this will directly bind to the protein and we can screen,
(Refer Slide Time: 05:23)



sorry we can scan the slide at 800 channel in infrared imaging system, this green colour indicates the total protein, we can quantify this same as like antibody signal quantified, and we can see that, so you can already see that in this plot, there are 5 or 6 different samples going it different way depending on time, and this signals can be normalized after normalization, once you plot this normalized raw data you look like this, to get a more smoothen data once we normalized this with FCF can get a signal like this, you can clearly see that already how nicely that data can be, data can be collected to make it biologically sensible.

(Refer Slide Time: 06:09)

Negative Control Slides

- Why negative slides?

Serve as controls for non-specific binding of the secondary antibody to the array

- How is it done?

Arrays are probed with the labeled secondary antibody (biotinylated anti-rabbit or anti-mouse IgG) in the absence of the primary antibody against the analyte of interest and processed as all other slides in the experiment



So with this and also morning he suggested to have what kind of controls to include and how to plan experiment and all, so it is important to include negative slides, so these negative slides serve as non-specific binding of the secondary antibody to the arrays, these array, every, one of the slide out of 50 will be probed with a only secondary antibody, and those signals will be deducted from the original signal and if you see any signal for the secondary antibody alone, so those antibodies definitely need to be dropped, there are some antibodies which, there are some secondary antibodies labeled for infrared from different companies, behave completely different.

(Refer Slide Time: 06:51)

RPPA Vs other protein detection methods

Protein analysis platform	Advantages	Disadvantages
Western Blot	Separation of proteins according to molecular weight	Work-intensive, high amounts of protein lysate required, low- or medium-throughput
ELISA	Quantitative, very sensitive	High amounts of protein lysate required
IHC	Cellular localization of protein of interest	Semi-quantitative, sensitivity often not sufficient to detect phosphorylated proteins
Mass spectrometry-based technologies	<i>De novo</i> discovery platform, highly multiplex, protein isoforms can be distinguished, analysis of thousands of proteins, no protein binding reagent required	Complex sample preparation, poor analytical sensitivity compared to immunoassays, low-throughput
Forward Phase Protein Arrays	Many analytes can be measured in parallel in a single sample, quantitative	Two highly specific antibodies are needed for every assay, high amounts of protein lysate required
Reverse Phase Protein Arrays	Robust quantification, low amount consumption, high-throughput, highly sensitive, detection of phosphoproteins possible	One highly specific antibody is needed for every assay, special devices needed

So now to summarize here how RPPA is advantageous or not advantageous over the different methods for detection of proteins, first you look at the western blot, advantages, separation of proteins according to molecular weight, so now you know that what molecular weight you are target of interest is, those santa cross will give 10 different lines but you know which molecular weight you should look for.

Then ELISA, but the disadvantage is very work intensive, high amount of protein is required, and it is very low throughput, maximum you can do medium throughput but which is not sufficient to analyze thousand subsamples, the ELISA which is quantitative, very sensitive, but it requires high amount of lysate.

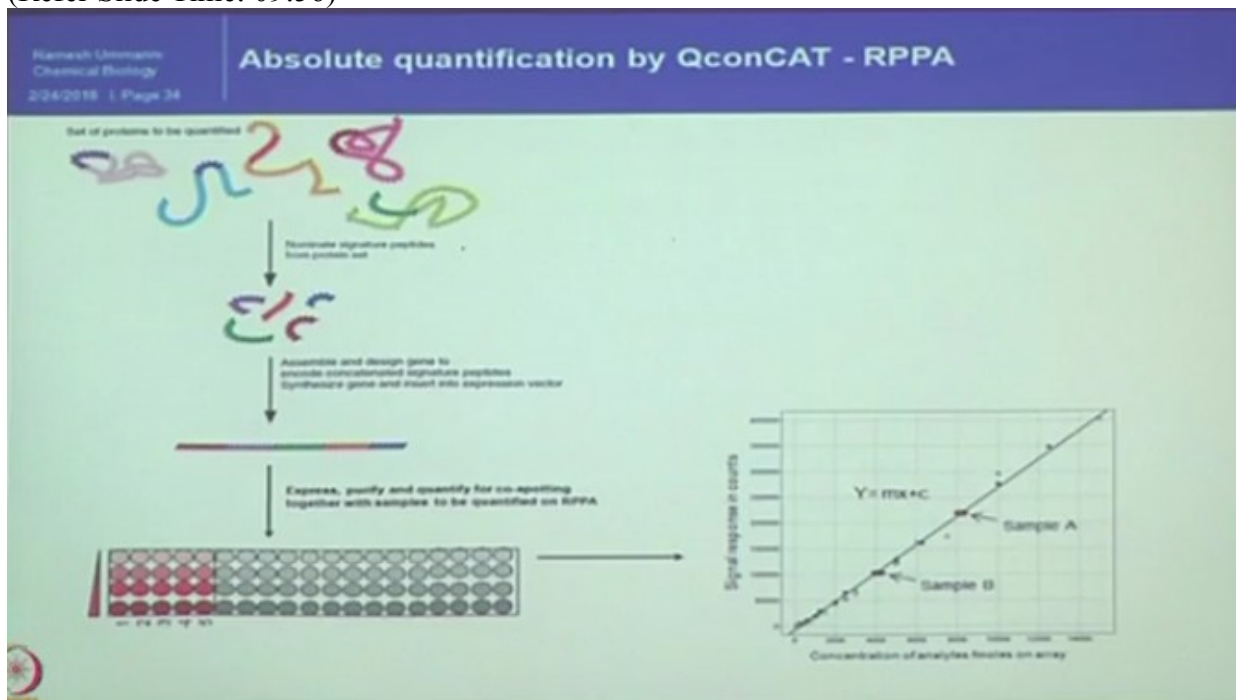
Then IHC is a cellular localization of protein of interest, it is also semi-quantitative and sensitivity often not sufficient to detect phosphorylated proteins very, very few copies of phosphorylated proteins will be there in the cell, and then it is very difficult to look at them in the IHC.

And mass spectrometry based technologies I'm not saying that it is not good, in fact it's a De novo discovery platform but highly multiplexing can be done, protein isoforms can be distinguished, analysis of thousands of proteins and no protein binding region is required, it's a just direct injection to some, this is the main advantage, but disadvantages if you look for complex, sample preparation, pore analytical sensitivity compared to _8:28_ definitely is a low throughput compared to the array platforms, I'll not say that it is very low throughput, it's a compared to array platforms it is little low throughput.

And forward protein arrays many analytes can be measured in parallel in single sample and it is quantitative. The disadvantage is that too highly specific antibodies are required there also because it's a antibody captures the un light, right, so capture antibody also should be highly

specific, but reverse phase protein arrays robust quantification low amount of consumption, high throughput, highly sensitive and detection of phosphoproteins, it's not that I'm saying that it will meet all the other disadvantages of the methods mentioned here it also has a limitation that it has a highly, it requires highly specific antibody and every SA need to be optimized and special devices required for this technique as well.

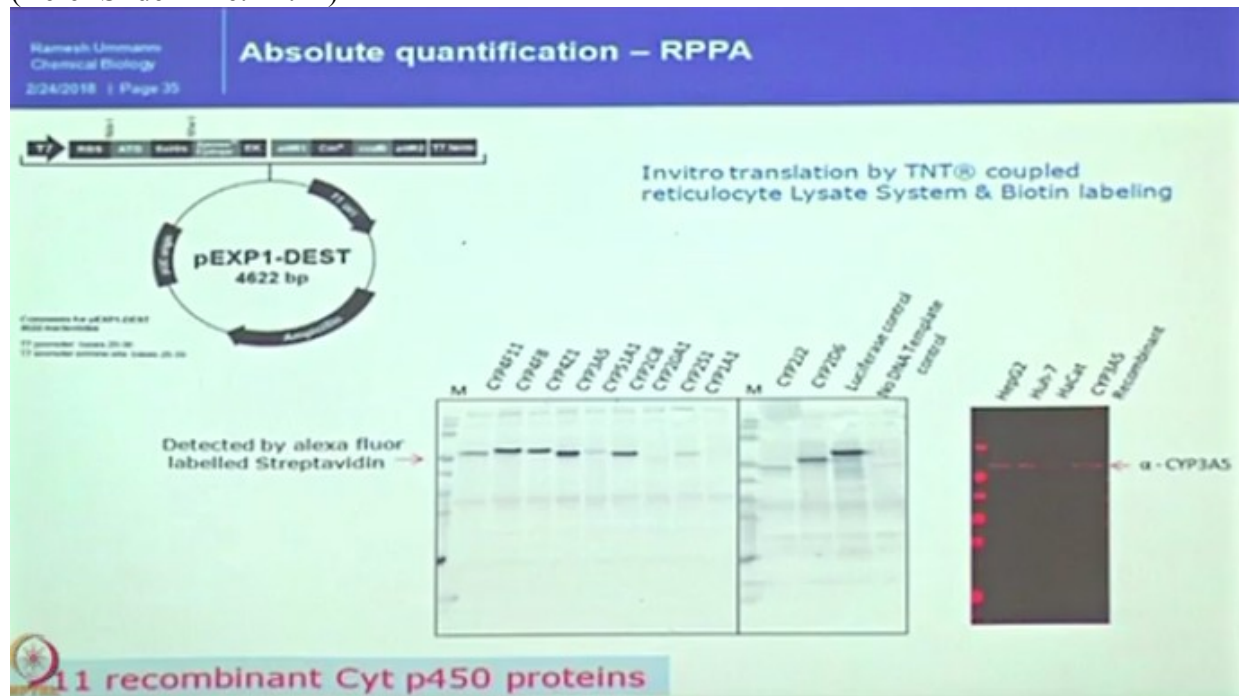
So this is how the RPPA works, and also you might have seen that always I'm talking about qualitative, now mass spectrometry evolved much higher level that even can perform the quantitative, absolute quantification even the number of copies of proteins present in per cell and all, so to some extent here also RPPA is useful to make absolute quantification of proteins, so this QconCAT approach that means quantitative, concanavalin, protein quantitative approach, that's why it is called as like this.
(Refer Slide Time: 09:56)



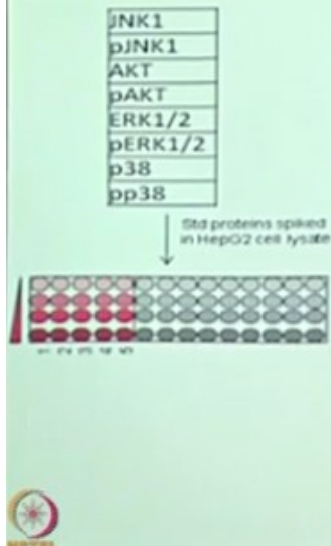
So let's say for example you can, you wanted to establish QconCAT approach, you select a protein of interest and you select epitopes recognized by the antibodies, so here you have a like and always when you buy antibody, that antibody will bind to the certain region of a protein, it will not bind to the entire region of a protein, right, so you can also develop a method called QconCAT approach by selecting those epitope regions of a proteins and you can express them as a fusion proteins, so this is definitely useful in all kind of protein array approaches.

And you can develop a single fragment like this containing different epitopes, and those proteins can be printed as a standards here in a concentration dependent manner, and then you can do the RPPA along with unknown samples here, then you make a graph like this, concentration of analytes on array and the signals on Y axis and you can also extrapolate the signals of samples, for example A and B on this plot and you can extrapolate and calculate the concentration exactly.

If you know the numbers of help you have taken to prepare the sample B, you can also say that this many copies present for this many population of cells, this is a way you can quantify, I'll show you one example,
 (Refer Slide Time: 11:24)



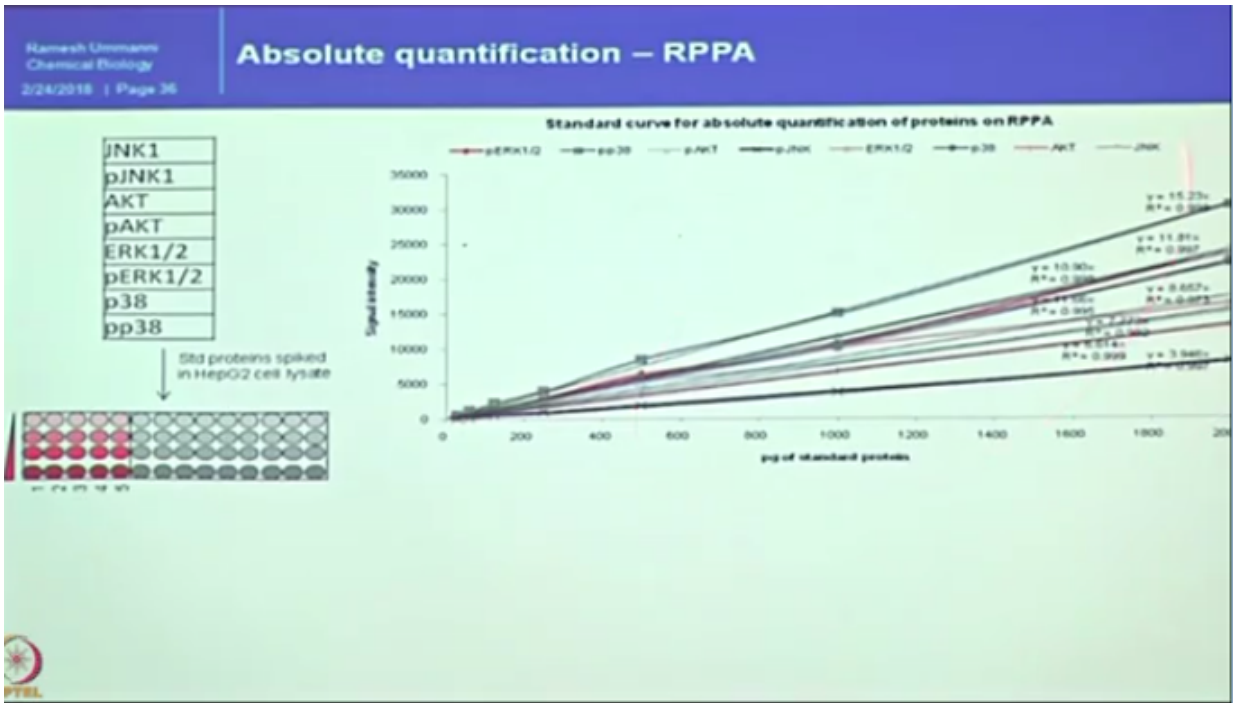
11 recombinant Cyt p450 proteins
 in this case at least for nash project nonalcoholic steatohepatitis project we tried to express all 11 recombinant cytochrome P450 proteins as a fusion proteins and using this vector system we developed the proteins and we purified them and we check the antibodies specificity and use them as a standard proteins to quantify the cytochrome P450 protein levels in a certain population of cells, so if 1 million cells contain how many copies of cytochrome P450 proteins like this 11 different recombinant proteins.
 (Refer Slide Time: 12:03)



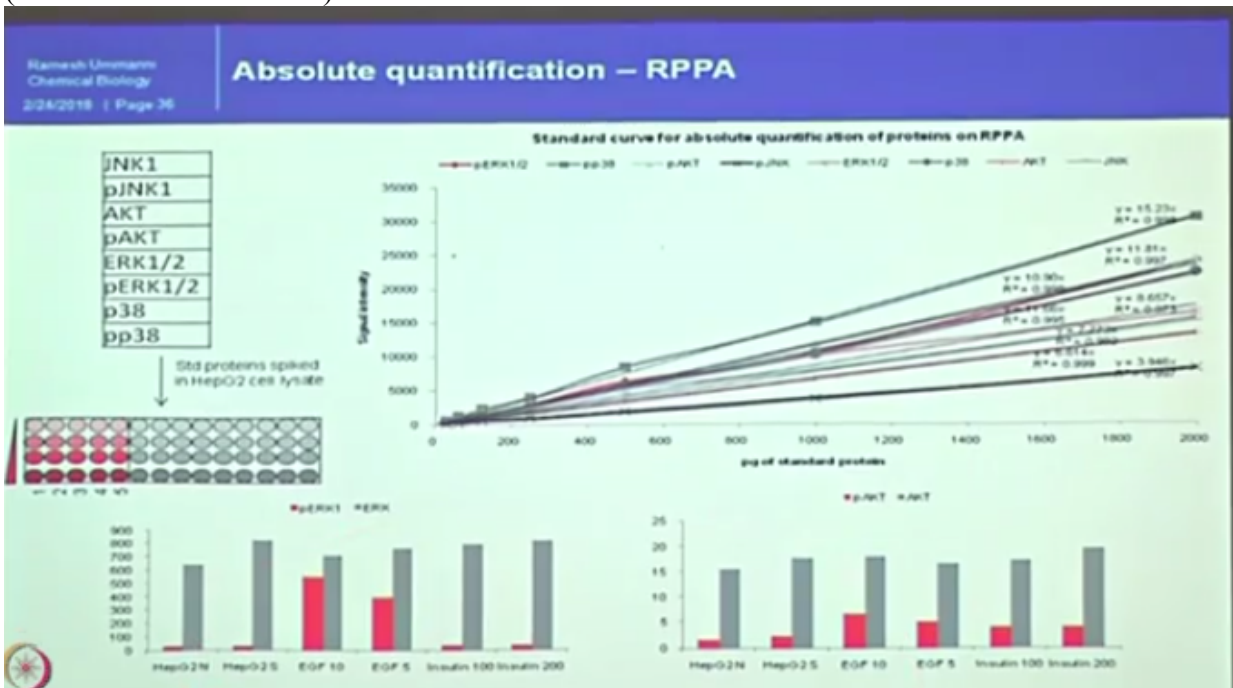
And also nowadays it is very, very important to do quantitative approach for phosphoproteins for example, if you take AKT protein not all AKT copies are phosphorylated when it gets activated, only certain population or certain copies of AKT proteins gets phosphorylated, how can we do that? So what we did is we have expressed the targets of interest in this case JNK, AKT, ERK and P38 expressed, purified, and invitro translated for protein modifications, sorry invitro phosphorylated for protein modifications using that specific carnages.

And those phosphorylated proteins we spotted in the concentration dependent manner, and then later on we did assess and we can show that phosphorylated proteins, unphosphorylated protein detection,

(Refer Slide Time: 12:56)



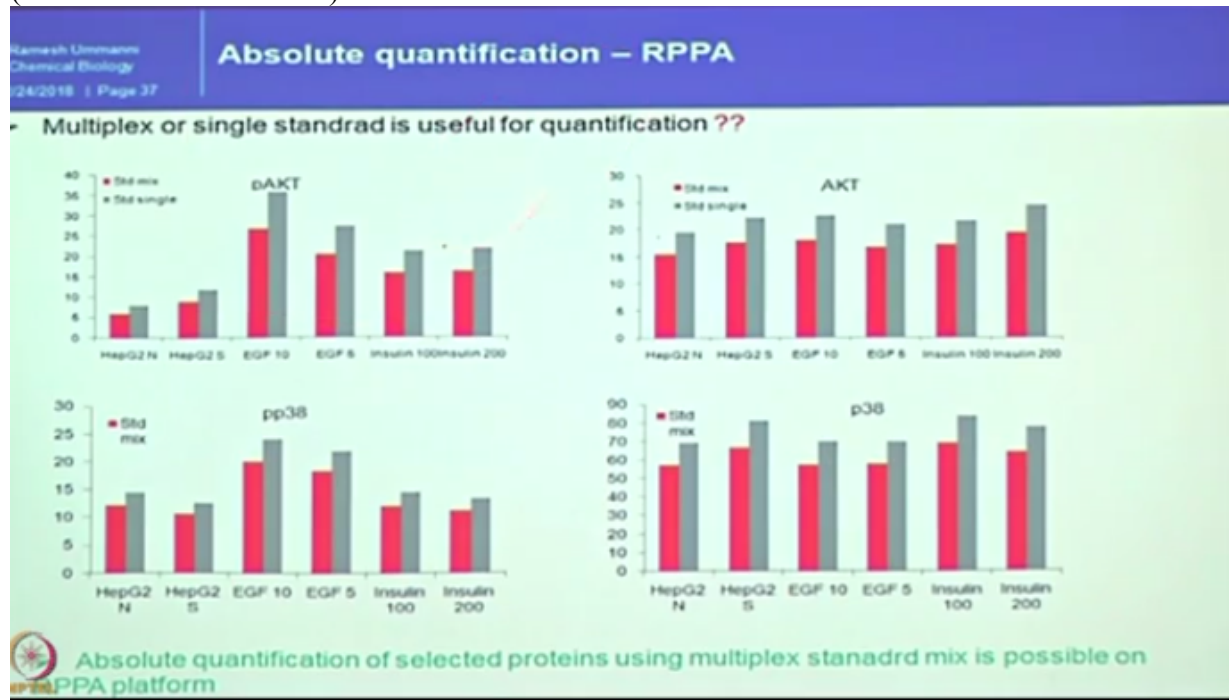
the linearity range and above point 9,
(Refer Slide Time: 13:02)



and then we looked at the signals detected here different cell lines and different stimulated samples, we can nicely detect total values are like this and up-regulated down-regulated copies are mentioned here.

And the similar way for phospho AKT and AKT here, I'm sorry the scales are different, here it is into 10, here is a total number is shown here, so in this way we can perform.

(Refer Slide Time: 13:24)



The other one is that absolute quantification can be done in multiplexing manner as well, so on the previous experiment we have spotted separately then we mix all of them, spotted as a single spots on a concentration dependent manner, and then when we look is a single, spotting is important or that can be multiplex is possible, so you can also, already see that little bit signal is down, but the behavior is same for all the different proteins, so you can say the power of RPPA is reached to a different level that we can also use this method for absolute quantification of a particular protein of interest in a larger population of samples, larger set of samples.

(Refer Slide Time: 14:09)

Applicability of RPPA

Various type of samples: 2D/3D cell cultures, xenograft, primary tissues, tumor biopsies, sections and LCD material etc.

High sample throughput: Hundres to thousands of samples simultaneously on a single nitrocellulose coated slide

Low sample and highly sensitive: Comprehensive signaling patterns from as low as 20 µg protein

Key nodes of signaling: validated antibodies cover total & phospho proteins in signaling networks

So now just imagine that, this is a technique is in your hands, how you imagine to apply into your own area of research is up to you, so to highlight some of them you can say that 2D or 3D, what kind of samples can be analyzed? 2D or 3D cell cultures, xenograft, primary tissues, tumor biopsies, tissue sections, LCD materials, laser capture microdissection materials, so definitely high throughput is low sample, components a signaling pattern from as low as 20 microgram protein which is not even sufficient to have one protein and one housekeeping gene on the western blot approach.

That key nodes of signaling, validated antibodies to cover total and phosphoproteins so that you can reconstruct the signaling networks when you stimulate at the cells for, with certain factors or that drugs and so on,
(Refer Slide Time: 15:08)

Ramesh Ummann
Chemical Biology
2/24/2019 | Page 39

Applicability of RPPA

Various type of samples: 2D/3D cell cultures, xenograft, primary tissues, tumor biopsies, sections and LCD material etc.


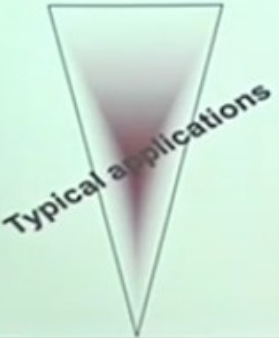

High sample throughput: Hundres to thousands of samples simultaneously on a single nitrocellulose coated slide

Low sample and highly sensitive: Comprehensive signaling patterns from as low as 20 µg protein

Key nodes of signaling: validated antibodies cover total & phospho proteins in signaling networks

Typical applications

- ❖ Disease mechanisms
Signaling pathway profiling
- ❖ Biomarkers discovery and validation
- early diagnosis, prognosis, prediction
- treatment efficacy
- ❖ Drug discovery and development
- On and off target activities of new drugs
- Right combinations of drugs
- ❖ PK/PD studies
- biologically active dosage of drugs
- biologically relevant dosage of drugs
- PD markers
- ❖ Clinical research
- Patient selection of personalized treatment

so to typical applications you can use this approach for disease mechanisms, signaling pathway profiling, bio-market discovery and validation particularly, for early diagnosis, prognosis, prediction treatment and efficacy.

And drug discovery projects like and on and off target activities of new drugs, when a new drug is coming into market, definitely nowadays regulators are asking what are the off target effects of your drug or what is the target of your drug, even if we don't, you don't know off target effects and right target of this molecule, what are the possible signaling mechanisms or signaling nodes targeted by the molecule this can be definitely achieved by RPPA approach and also right combination of drugs, so when you are looking at the personalize medicine, so if you using a cell based models and by reconstructing the signaling pathways you can choose a right combination of drugs, so always sometimes combination of combinatorial treatment or more effective than the single treatments, okay.

So PKPD studies you can identify them biological active doses of drugs, because some certain signaling pathway should be shut down or shut on, so what is a concentration of drug is required to switch on and switch off, so our reading or the end point reading, read out is the switching on and switch off by using the drug molecule, this switch on and switch off signaling pathways can be definitely done by RPPA in a high-throughput fashion.

So clinical research definitely patient selection for personalize treatment, I will touch up on later on, so using this technique we have done different projects, at least I have done this different projects, so I will touch only non-small cell lung cancer and prostate biomarker a little bit.

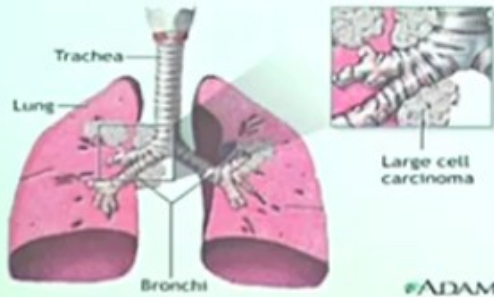
(Refer Slide Time: 16:59)

Predictive & prognostic markers - NSCLC

Lung Cancer :

uncontrolled growth of epithelial cells in

- one or both lungs
- tracheo-bronchial tree



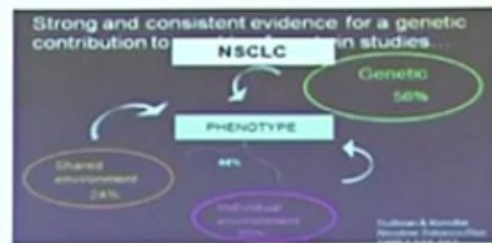
Risk factors: Genetic defects

Smoking (90% of all cases)

Second-hand smoke

(25% of non-smoker cases)

Occupational/environmental



Non-Small Cell Lung Cancer (NSCLC)

Comprises 85% of all lung cancers

Small Cell Lung Cancer (SCLC)

- Comprises 15-20% of all lung cancers

So lung cancer I don't want to go more details in to the lung cancer and definitely has two different types non-small cell lung cancer and small cell lung cell. Non-small cell lung cancer about 85% only 15 to 20% the small cell type, but today cancers are difficult to treat though having lot of drugs it is very difficult to treat and still mortality rate is high and due to the several reasons that the different subtypes of non-small cell lung cancers are there in the patient.

Now in the clinic even doctors are not able to segregate them before starting the treatment, so what kind of drug is good to give to them and all,
(Refer Slide Time: 17:40)

When genomic/transcriptomic is not informative...we are left to proteins

- Cancer is driven by hyperactive or defective protein networks (Pathways)
- Components of these pathways can be druggable targets



Patient A



Patient B

- Cancer in every patient is different. Patients will teach us how to treat them
- an approved drug works for one patient may not work for another one with the same cancer

so keeping that in mind at least we try to do some good experiment and how we can separate the patients based on their molecular subtypes using RPPA approach, okay.

I will show only one or two examples, so why it is required? So when genomic and transcriptomic is not informative, we are only left to proteins, so when doctors don't have any genomic information or molecular diagnostic information by analyzing the proteins they can take a decision what kind of drugs can be given to the patient.

So cancer in every patient is different, definitely patients will teach the doctors how, what kind of treatment they should receive. In this case patient A and patient B you can see that red line is passing differently in two different diagrams, but both of them are lung cancers assume in that way, but the signaling pathways are responsible for non-small cell lung cancers are different, but how doctor knows this? Unless and until he analyzes the signaling pathways responsible lung cancer only he will be able to understand how we did this,
(Refer Slide Time: 18:46)

Predictive & prognostic markers – why ?

- Cancer is driven by hyperactive or defective protein networks (Pathways)
- Components of these pathways can be druggable targets



<http://www.nature.com/reviews/drug-discovery>

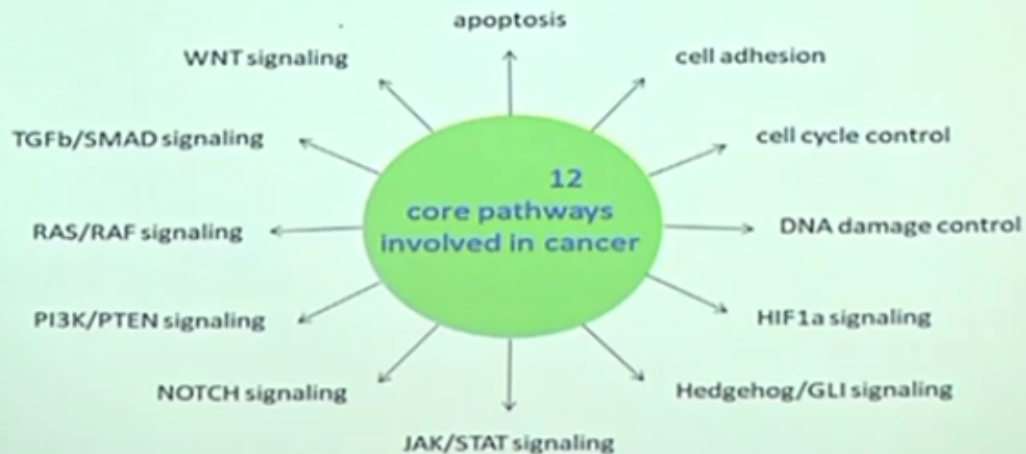
- Cancer in every patient is different. Patients will teach us how to treat them
- An approved drug works for one patient may not work for another one with the same cancer

in this case so again predictive and prognostic markers are required whether the disease is getting cured or not cured, so I will show with some examples.

(Refer Slide Time: 18:50)

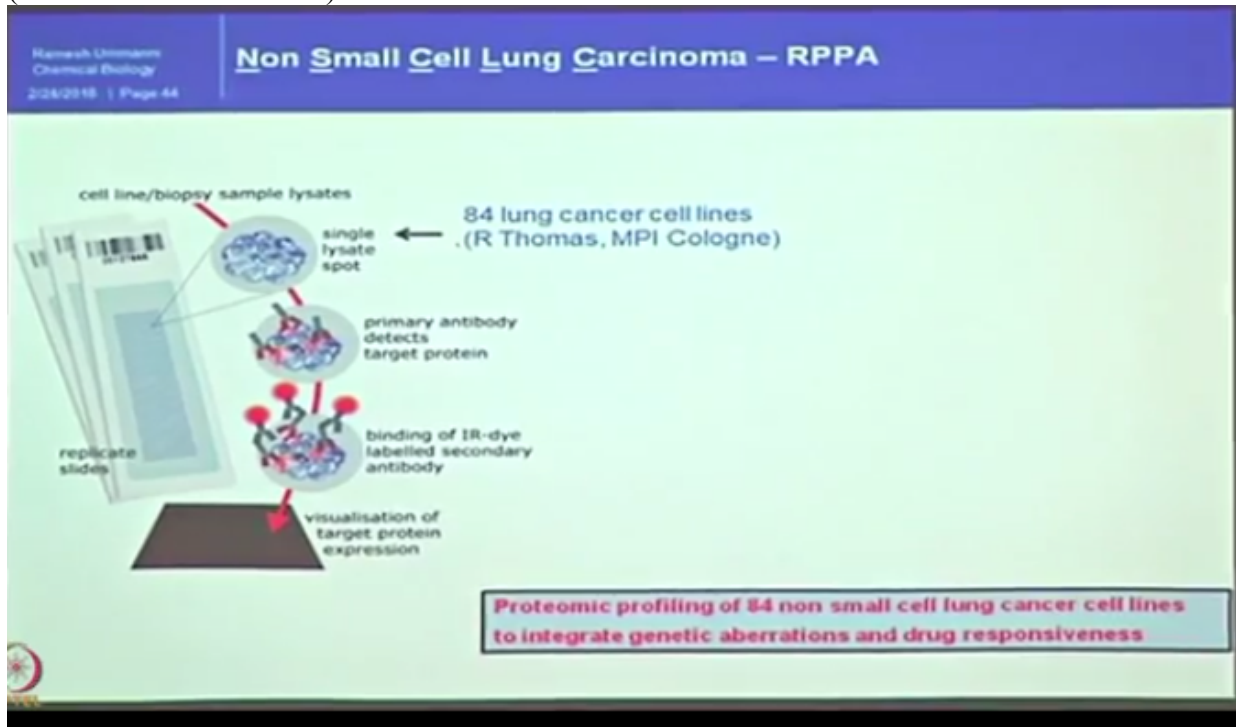
Antibody selection for signal transduction pathways profiling by RPPAs

Although a wide range of genes are mutated in cancer, the apparent driver mutations identified so far fall into components of 12 core pathways - Bert Vogelstein



Again as I said in the first couple of slides antibodies selection is a very, very critical and important to establish this method, although wide range of genes are mutated in cancer along with lung cancer, the apparent driver mutations identified, so far are components of 12 core pathways that means all the different mutations which caused lung cancer map into one of these pathways, that means if you analyze this pathways activation, inactivation in lung cancer

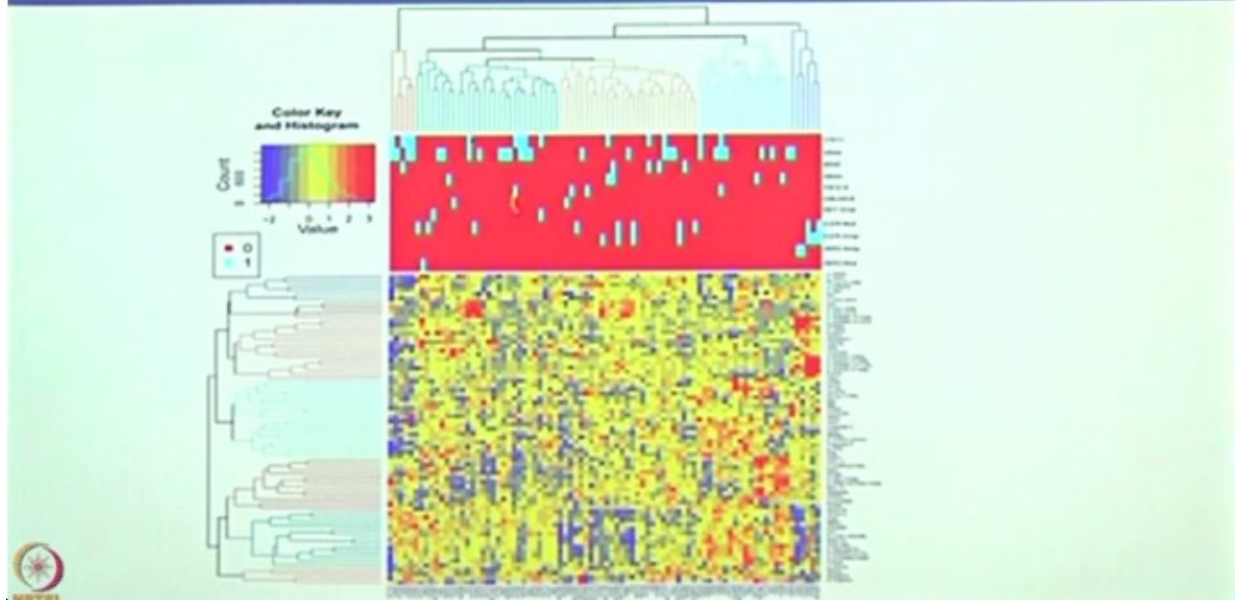
patients, based on the activation status of the pathways we can choose the drug to be given to the patient, this was the motto.
(Refer Slide Time: 19:47)



We decided or we arranged a panel of antibodies which are measuring the activation status of this pathways to proceed forward, of course it is very, very difficult to get patient samples with all possible molecular subtypes, but one of the group in Cologne Roman Thomas group has established 84 lung cancer cell lines from different molecular subtypes of the patients, so we know what kind of mutation they have, all the genomic data, genomic fingerprints are available and those cell lines were established but what are the different pathways activated/inactivated is not known.

So in this case these are the genetic aberrations which are causing lung cancer today and these 84 cell lines have this 11 different mutations and we perform the proteomic profiling again using the 82 antibodies which are from the 12 different pathways,
(Refer Slide Time: 20:31)

Un supervised hierarchical clustering of NSCLC cell lines using protein expression profiles

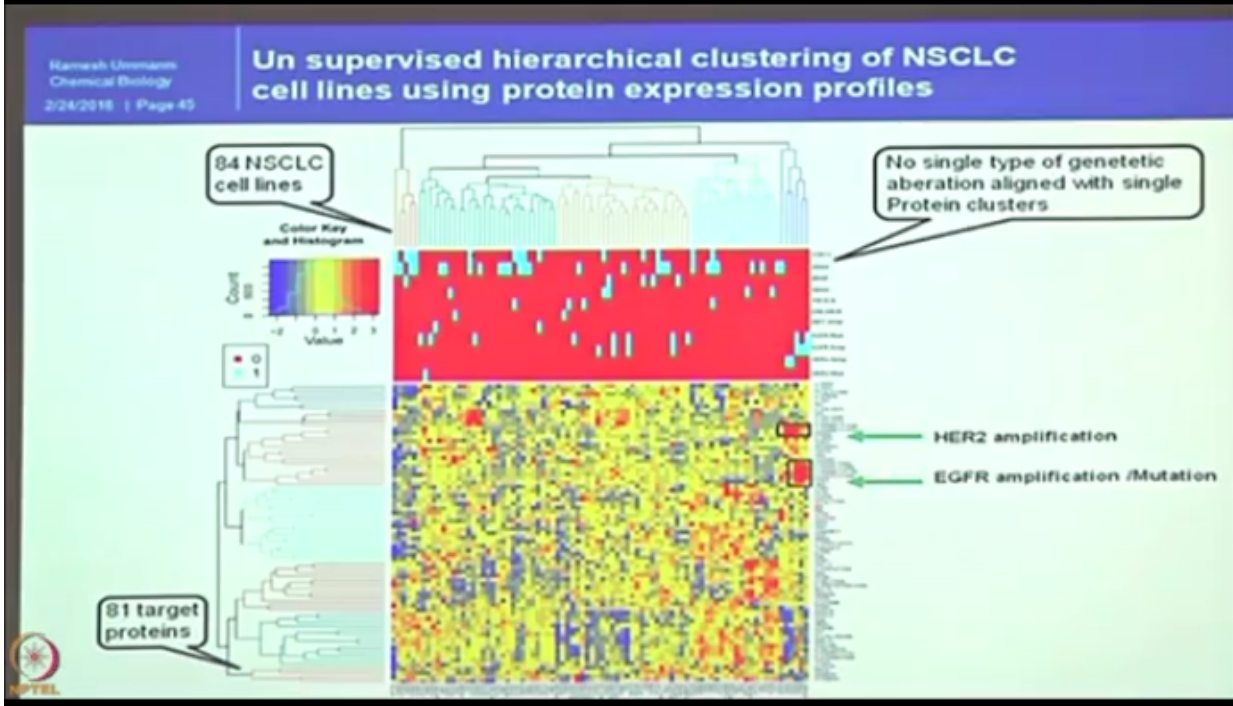


in the previous slide ultimately we get a heat map like this, and you can see this some of the blue areas concentrated, some of the red areas concentrated.

And on the X axis we can see the cell line and Y axis of the different proteins we measured, and this proteins are divided into almost 5, 6 clusters, all the proteins are divided or expressed as different clusters, here the cyan colour indicates a presence of the mutation and red colour is not presence of the particular mutation.

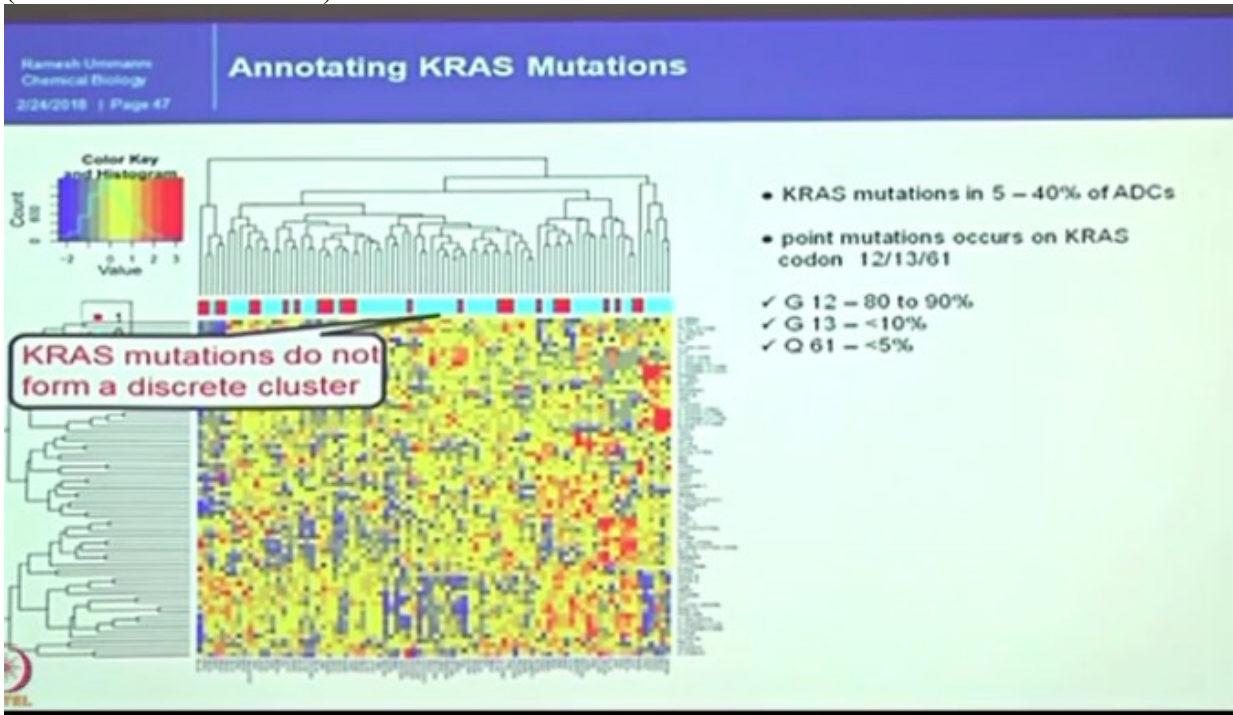
When we tried to align this, any of this concentrated areas or not aligning with cyan colours, so wide spread is there you can see that, but here we need to take decision here, what we have done is,

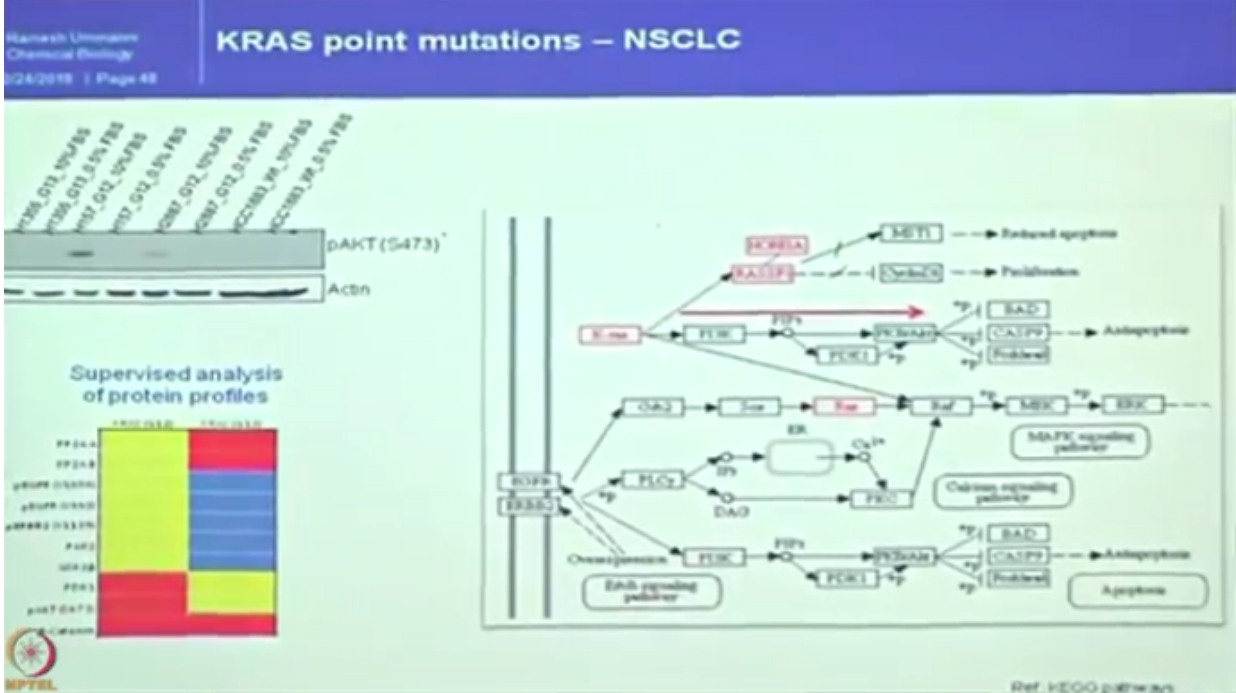
(Refer Slide Time: 21:15)



did the closer look here we can see that HER2 amplification and EGFR amplification or mutation signals, this activation of EGFR pathways are nicely aligned with this three places here, you don't say any sions here, right, this will looked at again, this is some of the AKT pathway proteins are inversely related to here.

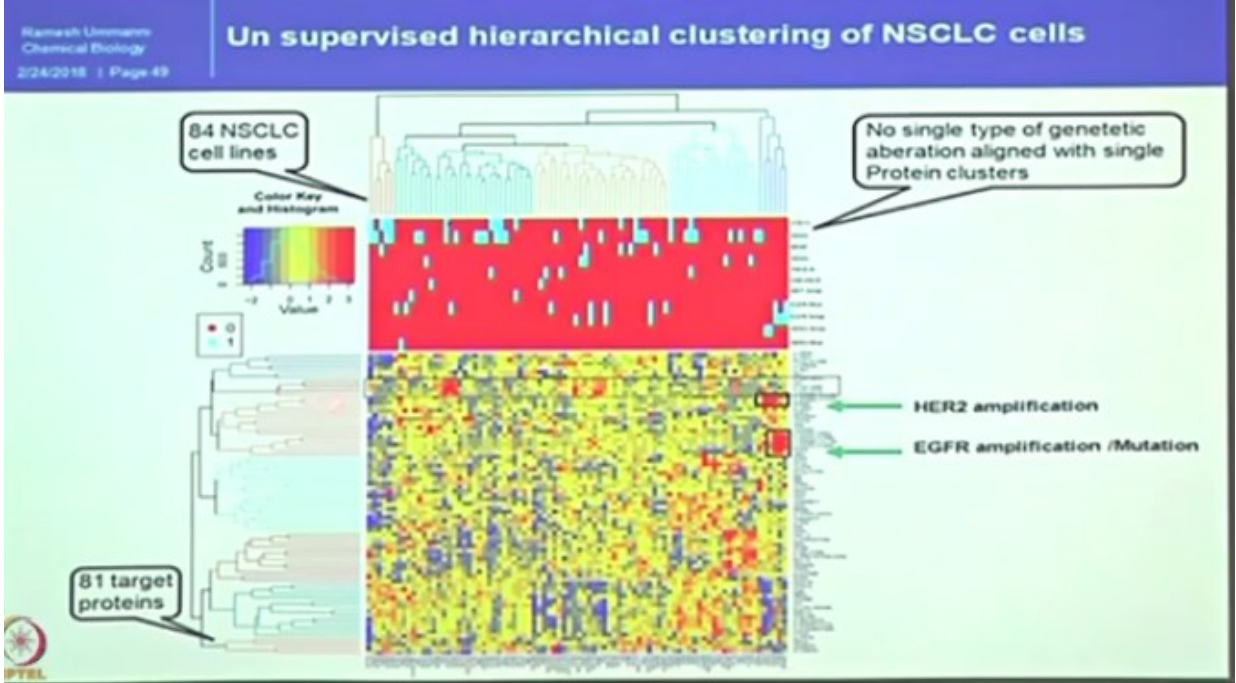
Red means up regulated, blue means down regulated, yellow means no change indeed, so this blue colour is inversely related, inversely proportional to the red colour observed here, (Refer Slide Time: 21:46)



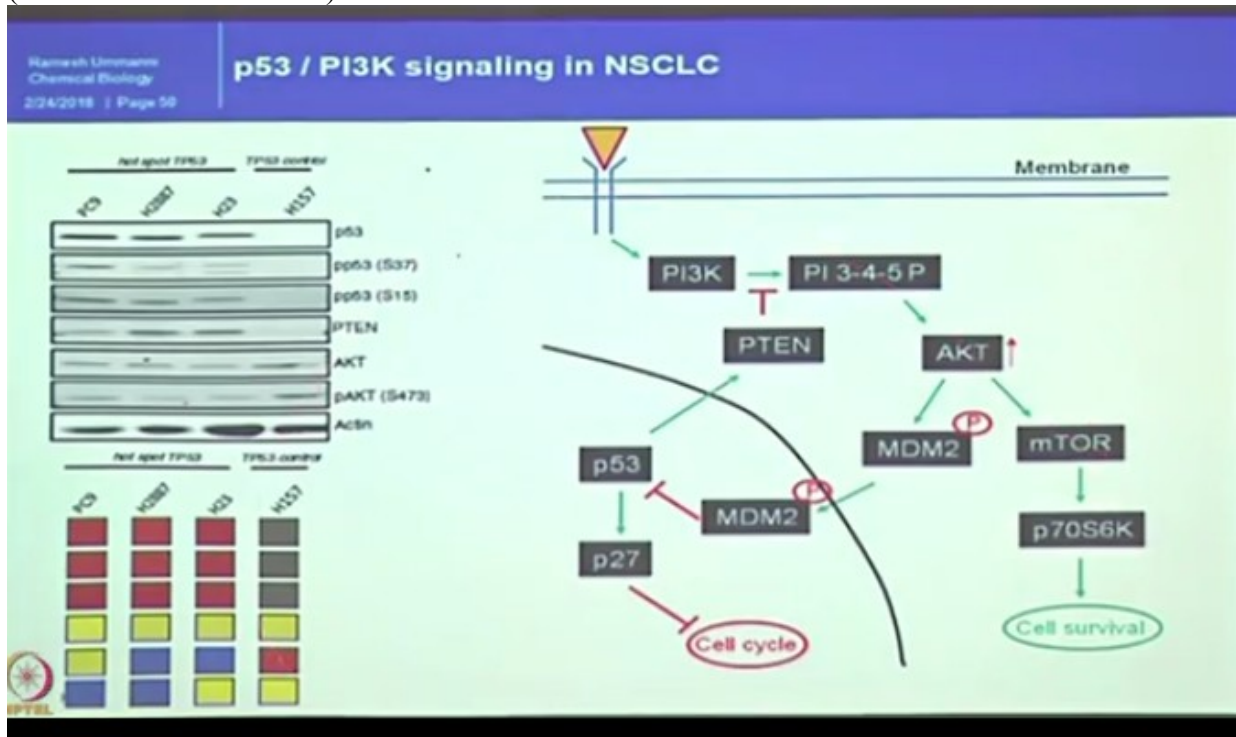


we proposed that, so in the KRAS mutations patients usually treated with RAS RAF inhibitors, so we also proposed based on this expression pattern that a set of patients with the K 13 mutation they should be treated with AKT inhibitors, not the RAS inhibitors, okay, so if the doctor knows that KRAS only then you will give the RAS pathway inhibitors, definitely set of population with K13 mutation they will not respond here, so this kind of information need to be managed arrays in the form of arrays and bring into the clinic to help the patient community here.

(Refer Slide Time: 23:38)

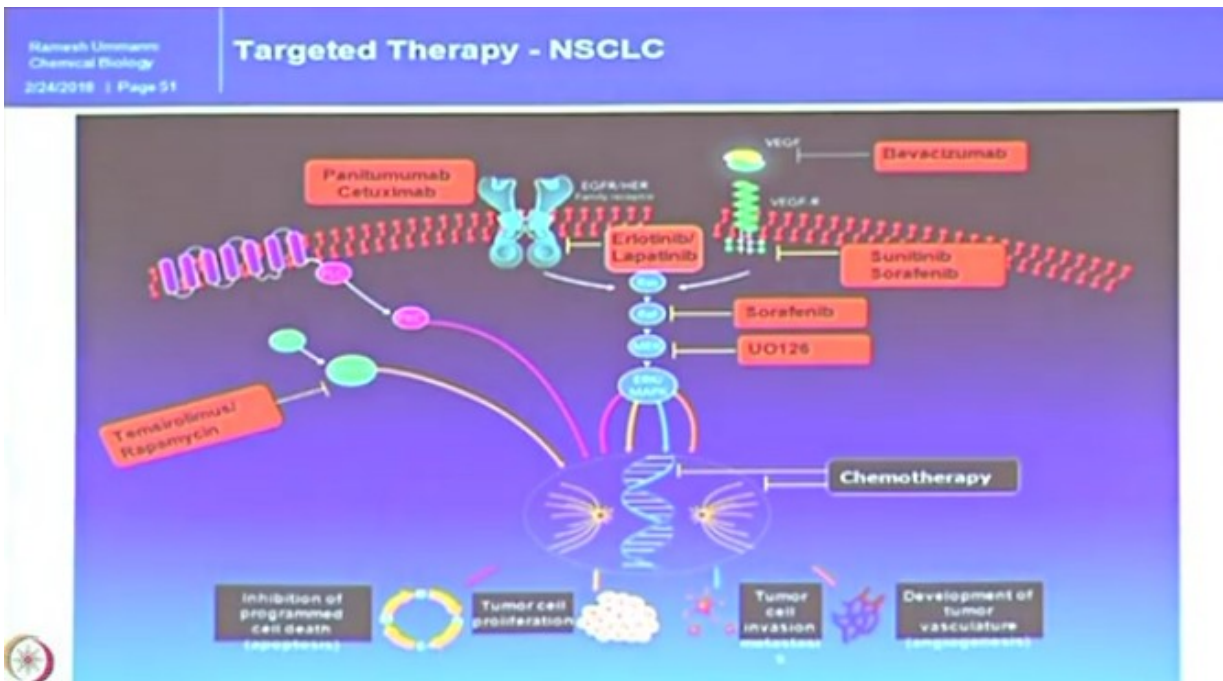


So in the similar way we also did for closer look for another set of small cluster here, you can see that this is the kind of small cluster with red patches here and here, this is basically the P53 pathway, and then we looked at the P53 hotspots or this kind of pathway we reconstructed, (Refer Slide Time: 23:50)



so which was not known for P53 patients, that means some of the P53 patients were never receiving AKT inhibitors earlier they were always treated in a different manner, so those P53 patients with activation status of AKT should be treated with AKT inhibitors rather than the other inhibitors, this information was missing earlier.

(Refer Slide Time: 24:13)



So this way we can use reverse phase protein array approach to categorize the patients for treatment approach.

And another way is that targeted therapy, so we have a drugs can be predicts whether the patient is responsive or nonresponsive, in that case so these are the different drugs available in the market,

(Refer Slide Time: 24:30)

Ramesh Ummann
Chemical Biology
2/24/2018 | Page 52

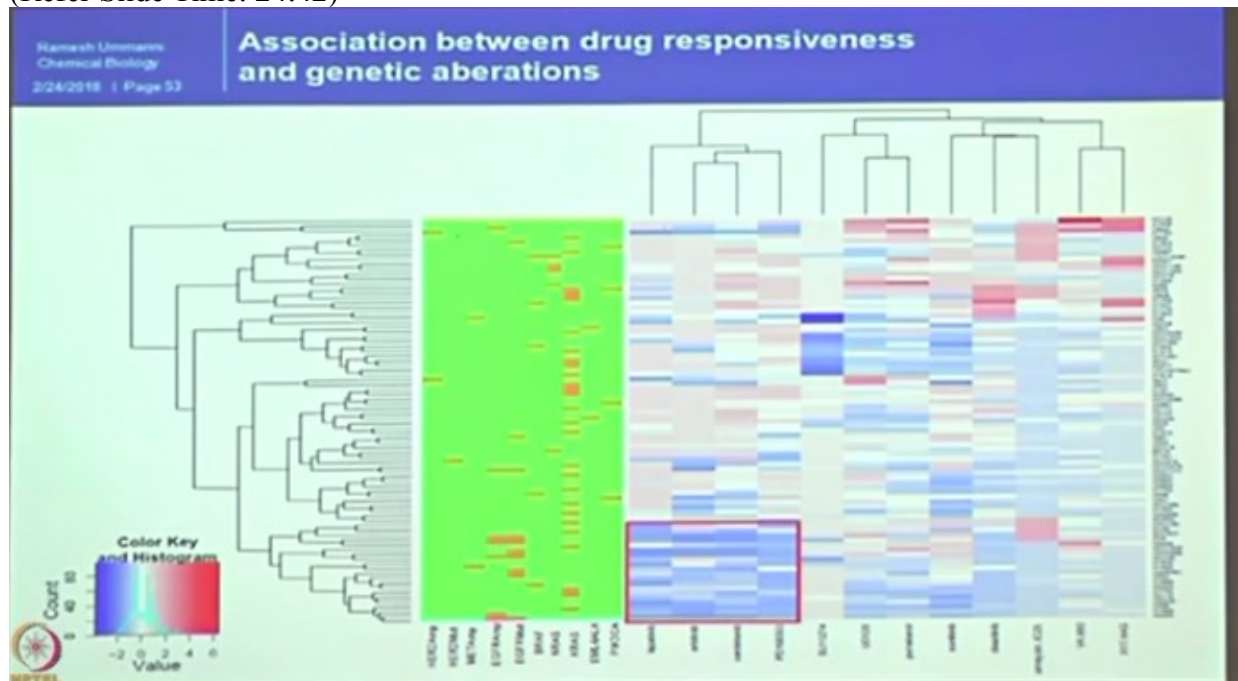
Targeted drugs tested in the NSCLC cell lines

Name	Target proteins
erlotinib	EGFR
sunitinib	Multiple receptor tyrosine kinases (RTKs)
vandetanib	EGFR and VEGFR
X17.AAG	HSP90
purvalanol	Cyclin-dependent kinase inhibitor
SU11274	MET
rapamycin	mTOR
VX-680	Aurora kinase
UO126	MEK
dasatinib	BCR/ABL, SRC, c-KIT, Eph kinases
PD168393	EGFR
lapatinib	EGFR and HER2/neu

> IC50 values of all above drugs was determined for all NSCLC cell lines

we have establish the IC50 values for all the drugs, again is the 84 cell lines, then we integrated the data using RPPA approach,

(Refer Slide Time: 24:42)

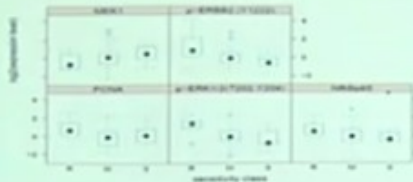


you can see nice cluster here, all this cell lines having the aberrations for EGFR amplification and DGFR mutations, the responding to the EGFR targeting drugs, that means already established data we can observe in this, whatever the data we generated here is a new, that means this is the kind of case control for us, or can we establish some of the new prediction markers using the other data set, so using this we also establish of that datasets, in this case it is nicely aligning here, further we established a set of markers for all the drugs we have screened for their IC50 values against 84 different cancer cell lines, so this is the way we have established.

(Refer Slide Time: 25:26)

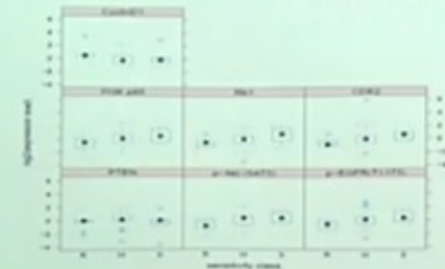
Proteins correlating with with drug sensitivity

UO126 (MEK inhibitor)



Protein	0	10	20
ERK1	Red	Yellow	Blue
p-ERK1/2	Red	Yellow	Blue
p-ERK2	Red	Yellow	Blue
PKA	Red	Yellow	Blue
MAPK3	Red	Yellow	Blue

VX680 (Aurora kinase inhibitor)



Protein	0	10	20
Aurora A	Red	Yellow	Blue
PKA	Red	Yellow	Blue
p-ERK1/2	Red	Yellow	Blue
PKB	Red	Yellow	Blue
ERK1	Red	Yellow	Blue



Prostate cancer – Proteomic analysis

Protein analysis of prostate biopsy material using RPPAs

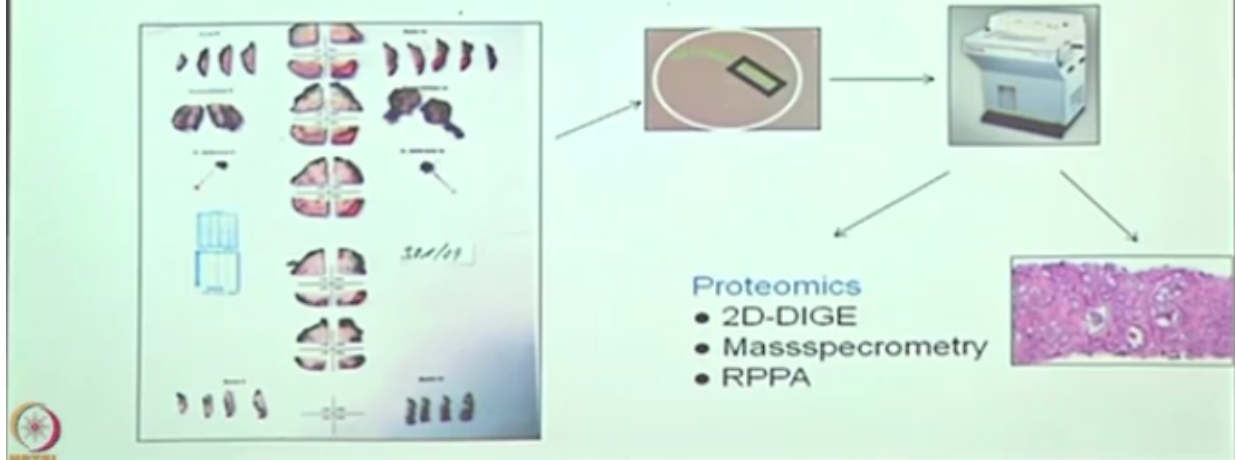


Now last couple of slides I will sum up now, protein analysis of prostate biopsy material using RPPA approach, then I said in the beginning that any type of sample can be analyzed here, so the capacities high-throughput fashion in the applicability of RPPA towards the biomarker discovery approach.

(Refer Slide Time: 25:45)

Prostate cancer – Proteomic analysis

Protein analysis of prostate biopsy material using RPPAs



One of the study I did during my post that time, so we have got the patient sample and I start with the pathologist sections, do the sectioning and get the protein extracted, ultimately I end up with a very, very small core which can give 50 microgram of sample with a real tumor cell population, with 50 microgram of sample I should identify the differential express proteins, I should validate them, show the proof that they are really differential express and then write a paper and published to get my PHD, okay, so that was a challenging.

But again here we did a traditional 2D DIGE, nowadays 2D generally is not so encouraging for the people, not very attractive method, those days it was possible to do and did mass spectrometry finally we did RPPA, so these are the different proteins which are differentially expressed,

(Refer Slide Time: 26:41)



Prostate Cancer, TP6 - Identification of clinically relevant protein markers of prostate carcinoma - NGFN PLUS

specifically in cancer patient compared to normal or benign prostatic hyperplasia in the surrounding tissues.

But again this many list is around 112, to validate 112 proteins I don't have a lysate with me, I wanted to analyze and establish the signature from the same set of samples, right, (Refer Slide Time: 27:04)

Ramesh Ummann
Chemical Biology
2/24/2018 | Page 59

Identification of clinically relevant targets systems biology network platform

PLOS one P1-S-10-02345R1
Identification of clinically relevant protein targets in prostate cancer with 2D-DIGE coupled mass spectrometry and systems biology network platform

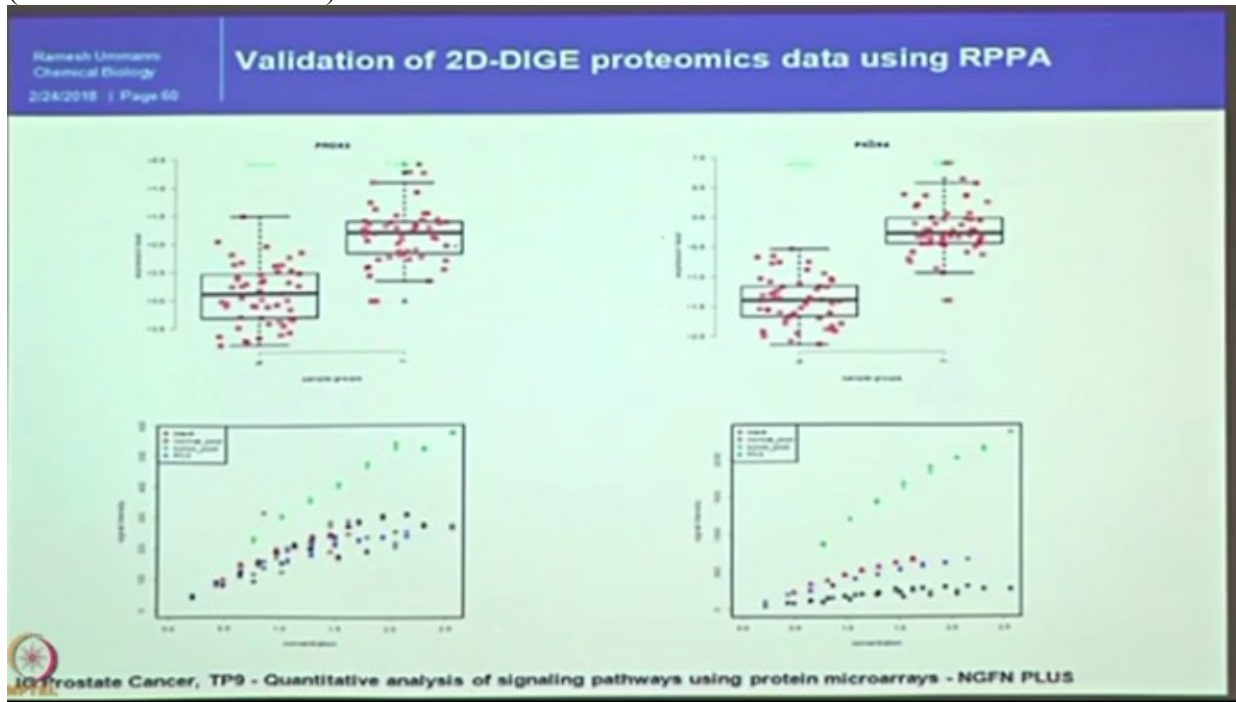
Ramesh Ummann¹, Frederike Misch¹, Birke Probst¹, Susanne Vogt^{1*}, Christian Schorf^{1*}, Christian Bauer¹, Maria Fahn¹, Jens Kallmann¹, Erhard Walter¹, Thorsten Schönm¹, Guido Sauer¹, Carsten Bokemeyer¹, Holger Schönm¹, A. Schoppert¹, Tan H. Brunsendorf^{1*}, Stefan Batschauer¹

¹Department of Oncology, Hematology and Bone marrow transplantation with section Prosenology, Herbertus Wald Tumor Zentrum (UCCH), University Hospital Eppendorf (UKE), Hamburg, Germany

Protein	3972 T	3972 N	4167 T	4167 N	4455 T	4455 N	36126 T	36126 N
GGAH1	+	+	+	+	+	+	+	+
eIF4A3	+	+	+	+	+	+	+	+
ARQ2	+	+	+	+	+	+	+	+
Prdx3	+	+	+	+	+	+	+	+
Par4	+	+	+	+	+	+	+	+
FKBP4	+	+	+	+	+	+	+	+
GAPDH	+	+	+	+	+	+	+	+
Prdx4	+	+	+	+	+	+	+	+
UCHL1	+	+	+	+	+	+	+	+
GAPDH	+	+	+	+	+	+	+	+

so this network platform we built using systems biology based approach and we enriched the marker analysis using the software tools, and using the known information these are the different targets at least I have highlighted to proceed further, this listed is also not small, so almost 10 proteins are there to go proceed forward, can see that difference between the tumor

and normal very nicely here, to proceed forward then I did this reverse phase protein array approach,
(Refer Slide Time: 27:29)



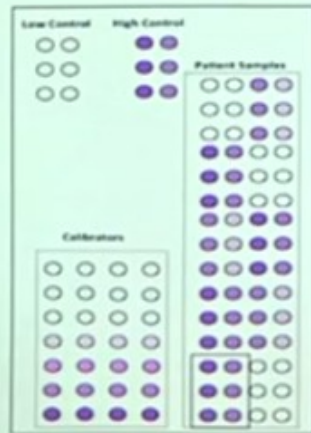
I could see clearly that different for example I'm showing peroxiredoxin 3 and 4, you can see each dot represents one patient and it is up regulated in patients samples compared to the normal or control samples.

And also we can see the signal dynamics depending on the serial dilutions and so on, so based on this approach we highlighted, we could identify that these are the different proteins are really differential regulated in cancer soft tissue compared to the normal tissue, from 50 gram protein we ran a 2D gel which obviously requires high quantity of protein, we performed the validation of 10 proteins and still 2, 3 micrograms left over in the tube, so this is a power of RPPA residues.

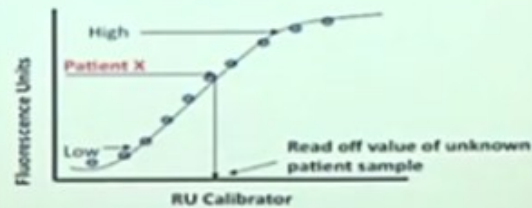
(Refer Slide Time: 28:13)

Where RPPA is today for Clinical application

RPPA CALIBRATED ASSAY FORMAT FOR CLINICAL LABORATORY IMPLEMENTATION



- Elimination of long dilution curve: lysates printed in large number of technical replicates: better precision and accuracy
- Built-in low and high controls
- Built-in calibration curve: Constructed with constant total protein concentration but with known and varying analyte concentration
- CLIA-oriented format provides ability to compare and integrate data across experiments and time
- Values reported as quantitative RU value or absolute quantitation (e.g. ug/ml) obtained by interpolation of sample value to calibrator curve



Rehan Akbani et al. Mol Cell Proteomics 2014;13:1625-1636

Now where are we standing today? And as I said that, that time not many people were accepting that it can go into clinic and all, so now in 2014 RPPA construction meeting was there, people were talking about how can we bring it? So here once we know the level, cutoff level let's say for example blood glucose level we know it is 80 to 120, below 80 is like a hypoglycemic, then above 120 hyperglycemic, so in that manner so if we know the levels we could spot some of the known protein, known concentrations as a low control and high controls, and here we can have them as a calibrators based on the concentration dependent, we can construct a plot like this, the moment if we get a patient data in this panel and we can extrapolate on the graph and you can say that what is the quantity, if this quantity is below or between the low and high or above high or below low then we can decide that, so this patient is having X problem or Y problem and so on.

So now people are talking about this kind of mini-expressed RPPA platform to bring into the personalize medicine.

(Refer Slide Time: 29:25)

Summary & Conclusions

- RPPA is suitable for simultaneous study of 100s to 1000s of samples for expression and/or activation of a protein of interest
- RPPA correctly identified cell lines with amplified EGFR & ERBB2
- Single point mutations did not align with activated pathway specific clusters
- RAS (G12) mutation results in active RAS – PI3K signaling rather than RAS – Raf pathway in NSCLC cells
- RPPA revealed a cross talk of p53 activation and AKT/PKB signaling axis
- RPPA data integration with IC50 values of NSCLC cells provide additional prediction markers for assessing their drug responsiveness
- RPPA is useful for validating proteomic data obtained from limited sample of prostate biopsies to identify clinical markers
- Absolute quantification of selected proteins (total/activated fraction) can be achieved

So with this I will just sum up, it is suitable for simultaneous study of 100s to 1000s of samples for expression and activation of a protein of interest, it's correctly identified the cell lines with EGFR, single point mutations did not align with activated pathways, but at the closer look for mutations on G12 and G13 this is how RAS PI3K signaling rather than RAS RAF pathways.

RPPA revealed a cross talk between P53 activation AKT which was not known earlier, the RPPA data integration with IC50 values have provided additional markers, it is also useful for validation of proteomic data obtained from limited sample, and absolute quantification of selected proteins either it can be total or activated fraction can be achieved or sell in fact, so this kind of experiments can be done using RPPA approach, but all in all now you might have understood that this basic steps, necessary steps to follow to printing and then hybridization and all steps are same, only difference between forward arrays and reverse arrays is that here we spot a just screwed lysate, you don't have to follow any separation of samples and so on.

So on one experiment, on a one slide you can understand or you can estimate quantity of a particular protein of interest across the large number of samples.

(Refer Slide Time: 31:02)

Acknowledgements

Center for Chemical Biology, CSIR – IICT

- Divya Duscherla
- Chandra sekhar
- Karthik Reddy
- Sirisha
- Supriya
- Bhavani

Max-Planck Institute, Cologne, Germany

- Roman Thomas
- Martin Sos

Quantitative Proteomics, DKEZ, Germany

- Ulrike Korf †
- Heiko Mannsperger



† 30-03-2016

So this is what, so with this I will just like to acknowledgement some of the people in my lab and most of the wet lab data generated with the help of Ulrike Korf and she unfortunately passed away 2 years back, so she is the one who got me into the reverse phase protein array area in fact, so with this I thank you all of you for patiently listening for almost one and a half hour probably I'm very opened to discuss with you, so in this case again and again I wanted to emphasize that I may not be able to show you how the slide looks like and all, it is exactly the necessary steps you are following is 100% similar, unless and until if I tell you that this is reverse forward you will not know yet, that's it.

So I hope you will be convinced to do, to do the experiment by virtually imagining the steps I have highlighted you. Thank you very much.
(Refer Slide Time: 32:22)

Points to Ponder

- The overall workflow and various steps involved in an RPPA experiment is similar to any other microarray experiment
- RPPAnalyzer is a popular tool for RPPA data analysis and visualization
- Absolute quantification can be achieved by QconCAT-RPPA
- RPPA has been applied extensively in the context of clinically relevant problems



Sanjeeva Srivastava: So this now concludes the session on reverse phase protein arrays, I hope you have learnt good work flows and basics of reverse phase protein arrays, setting up the experiment, leading towards the data analysis. By now you know that the remaining applications which can be performed on reverse phase protein arrays platform which Dr. Ummanni has explained in contexts of clinical relevant problems.

I just like to add that now the technologies are really progressing well and we have you know better printed and the arrays to make this kind of chip platforms we have sophisticated software tools which can do that analysis better, we have much more high-throughput capability, nevertheless the reverse phase protein arrays have been in used in the clinics from long time, not as the high-throughput technology platform, but even to test out abundance of a given protein, clinicians have been using it from long time and this also shows the need for having reverse phase protein array based platforms for the clinical applications.

So by knowing that the basics as well as the possible applications, I hope now you are excited to really understand this technology in much more detail which can be really, really helpful if your goal was to delivered something to the clinics for translational research. Thank you.
(Refer Slide Time: 33:58)

Next lecture....

Antibody signatures defined by high-content peptide microarray analysis



MOOC-NPTEL

IIT Bombay

Prof. Sridhar Iyer

**NPTEL Principal Investigator
&
Head CDEEP, IIT Bombay**

**Tushar R. Deshpande
Sr. Project Technical Assistant**

**Amin B. Shaikh
Sr. Project Technical Assistant**

**Vijay A. Kedare
Project Technical Assistant**

**Ravi. D Paswan
Project Attendant**

Apoorva Venkatesh

Teaching Assistants
Shalini Aggarwal

Nikita Gahoi

**Bharati Sakpal
Project Manager**

**Bharati Sarang
Project Research Associate**

Nisha Thakur
Sr. Project Technical Assistant

Vinayak Raut
Project Assistant

Copyright NPTEL CDEEP, IIT Bombay