LECTURE-24

Quantitative proteomics: Stable isotope labeling by amino acids in cell culture (SILAC)

TRANSCRIPT

Welcome to the proteomics course. In today's lecture we will talk about Quantitative proteomics: Stable isotope labeling by amino acids in cell culture (SILAC). The complexity and dynamic nature of proteome present major technological challenges. Mass spectrometry advancements have improved the high-throughput identification and quantification of proteins. The mass spectrometers have advanced significantly and now offering opportunities to understand the human diseases and discover biomarkers. MS has been proven to be an extremely powerful tool to analyze the protein complexes. However, MS is not a quantitative technique to begin with and peptide ionization efficiency is unpredictable. So usefulness of this data for quantitative purpose remains limited. Much advancement in MS during the last decade have provided new ways for protein analysis and facilitated the study of proteomic analysis of various biological systems. Advancement in MS include development of highly sensitive mass spectrometers, fast scan rates, automation, nano-flow liquid chromatography as well as new techniques and methods to quantify proteins its abundance for quantitative proteomic analysis.

So in today's lecture we will first discuss about MS-based quantitative proteomics, *in vivo* labeling methods, I will introduce SILAC and then we will talk about the SILAC experiment and step-by-step workflow. We will then discuss the merits and demerits of SILAC method.

So let's start with MS-based quantitative proteomics. Protein labeling with stable isotopes are new effective methods for quantitative proteome profiling using MS. These isotopic labels can be introduced *in vivo* or *in vitro* and provide relative abundance of proteins for proteomic analysis. The isotopically labeled peptides are chemically

identical and generate similar specific signal intensities in MS. The relative labels of isotopically labeled peptides are determined by comparing the signal intensities of the paired peptides.

So as I discussed MS is not a quantitative technique by itself due to the peptide ionization efficiency which is unpredictable. So the differential stable isotope labeling is used to create a specific mass tag. There are different types of mass tags which are very promising to various quantitative proteomic applications including ICAT, ITRAQ and SILAC as well as various other tags. These tags can be recognized by MS and provide the basis for quantification.

These mass tags can be introduced by various methods into the proteins such as metabolic labeling, chemical means of labeling, enzymatic methods or by using synthetic peptide standards. The accurate quantification in MS can be achieved by the use of stable isotope-labeled standards.

So there are different ways of quantitative proteomic analysis, *in vitro* and *in vivo* labeling methods. In today's lecture we will focus on in vivo labeling methods. So most of the quantitative proteomics approaches by MS utilize isotopic labels as a reference for the relative or absolute quantification. The labels can be introduced in vivo, for eg., by growing an organism in a media enriched with specific isotopes. The labels are also introduced by performing tryptic digestion in presence of heavy water. Many methods using isotopically labeled reagents that react with specific amino acids or the protein N-terminals are also developed. For comparative and quantitative proteomic analysis the development of stable isotope tagging methods can allow the quantification of relative levels of proteins. This differentially labeled peptides with stable isotopes can be distinguished by characteristic mass shift in MS.

There are different ways of in vivo labeling such as enrichment of ^{15}N media, culture derived isotope tags (CDIT) or SILAC. Although we will discuss briefly about ^{15}N media and CDIT and then we will focus on mainly SILAC method for rest of the lecture.

So these stable isotope tagging methods use isotopic nuclei such as ^{2}H, ^{13}C, ^{15}N and ^{18}O. These stable isotopes are incorporated in place of natural abundant isotope in heavy standard. By using the stable isotope labeling methods one can determine the relative expression level of proteins in two samples.

Let's briefly discuss about ^{15}N labeling methods. Yeast or bacterial cultures which are grown in two separate media, one containing ^{15}N. The cells are pooled together and proteins are extracted from these pooled cells and quantified by using MS.

Although ^{15}N labeling is an easy method but there are various disadvantages. The protein incorporation in your control and the treatment could be unequal due to these stable isotopes. Therefore the labeled and unlabeled peptides exhibit variable mass shift in MS spectra. The mammalian system very poorly incorporates these stable isotopes. Furthermore this method is difficult and expensive. Due to these limitations researchers have tried other labeling methods as well.

Now let's discuss about CDIT. In this method the cells which are cultured in the stable isotope-enriched medium are mixed with the tissue samples which serve as an internal standard. The synthetic unlabeled peptides can be used for the quantification of the corresponding proteins which are labeled with stable isotopes. The proteins can be extracted and digested prior to the MS analysis. The ratio between the two isotopic distribution can be determined by MS.

After having discussed some of the less commonly used such as ^{15}N media and CDIT, now let's discuss about SILAC. So the MS-based quantitative proteomics is an increasingly popular approach to study the changes in protein abundance in various biological samples. SILAC which is a metabolic labeling strategy to encode whole cellular proteome is one of the very widely used method for the quantitative proteomics. In SILAC method the cells are grown in a culture medium where natural form of an amino acid is replaced with a stable isotopic form such as arginine bearing ^{13}C atoms. Incorporation of the heavy amino acid occurs through the cells grown protein synthesis and turn over. The SILAC method allows for light and heavy proteomes to be

differentiated by MS while avoiding any chemical derivitization and associated purification.

So SILAC is a metabolic labeling strategy which uses the stable isotope labeled methods in the growth medium. This experiment depends on the cellular protein synthesis to incorporate the stable isotope containing amino acids into whole proteome, for eg., arginine or lysine which contains six ^{13}C atoms.

The SILAC experiments are performed by incorporating the stable isotopically labeled amino acids such as L-Arginine containing six ^{13}C through the natural protein turnover and cell growth. The cells are cultured in 2 separate media for the light and the heavy forms. So the light medium with the amino acids of natural isotope abundance whereas the heavy medium contains the SILAC medium amino acids of choice. These are commercially available.

So the labeled analogues of these amino acids are supplied during the growth of these cells which are incorporated during the protein synthesis in all the newly synthesized proteins. After number of cell divisions, each instance of a particular amino acid is replaced by its isotope labeled analogue and finally MS can be used for determining the relative protein abundance by the intensity of light and heavy peptides.

Now we will discuss about SILAC experiment and the workflow to perform the experiment. In the workflow we will discuss the SILAC protocols and how to incorporate SILAC labels into any given experiment.

So let's have a look on the workflow for performing a SILAC experiment. First step is preparing SILAC labeling medium. Second, adaptation of cells from the normal DMEM medium to SILAC labeling media. Third, the differential treatment application to the SILAC cells. Fourth, cell lysis and protein estimation. Fifth, MS analysis and quantification.

So let's look at the workflow stepwise. The first point, the preparation of SILAC labeling medium. So in SILAC experiment a defined media with known sources of amino acids

can be adapted for labeling. The amino acid for labeling can be left out from the media formulation to ensure that the light and heavy amino acid stocks used in the media preparation are the only available source of amino acids in cells.

So there are two widely used media. One is the Dulbecco's modified Eagle's medium (DMEM) and RPMI-1640. The arginine , lysine and methionine are removed from the standard formulations. The light medium is naturally abundant isotopic forms of amino acids whereas the heavy medium is the same medium which lacks the desired amino acids (for eg. arginine and lysine) and they can be substituted as the heavy isotopic forms.

Here I am providing you an example of stable isotope-labeled amino acids. $^{13}C_6$ is a stable isotope of $^{12}C_6$ L-Lysine. It gives 6 Da difference in the MS as compared to $^{12}C_6$ L-Lysine. $^{13}C_6$ is another isotope of $^{12}C_6$ L-Arginine form. It is again 6 Da heavier than $^{12}C_6$ L-Arginine .

So for the preparation SILAC labeling media one need to add certain supplements such as serum, antibiotics. So the recommended percentage for serum would be 10% dialyzed fetal bovine serum and 1% of antibiotic and glutamine. But these percentages can be optimized depending upon the type of cell cultures.

In the SILAC experimental workflow, let's discuss about the second point. The adaptation of cells from DMEM to SILAC labeling media. The cells need to be adapted due to slight differences in media formulations.

If the cells are grown in DMEM medium it should be first split into two culture dishes containing light and heavy SILAC medium. So first, take out 10-15% of cells from he original culture and allow it for doubling. Then subculture the cells in respective SILAC medium, the light and heavy medium, and allow it for atleast five to six cell doublings. I'll discuss about the significance of this doubling process in the next slides.

So for the adaptation of cells from DMEM to SILAC labeling media, at the end of this adaptation phase lyse the cultures by adding 6 M Urea, 2 M thiourea and then extract

the proteins, reduce disulfide bonds by adding 1 mM of dithio tritol (DTT). Then add 5 mM of iodoacetamide (IAA) to alkylate the cysteine residues. Add trypsin for digestion, overnight at 37°C, with the enzyme and substrate ratio of 1:100. Then these samples can be analyzed by LC-MS or LC-MS/MS. But first thing one needs to ensure that SILAC amino acids are fully incorporated.

So the adaptation of cell lines in heavy medium is the first step in SILAC-based experiment. Cells are adapted in heavy medium for atleast 5 or 6 doublings to be fully labeled. The digested proteins samples can further be analyzed by MS. These are representative spectra showing m/z ratio and the relative intensity of light, medium and heavy forms.

So in SILAC experimental workflow let's discuss the third point, differential treatment application to the SILAC cells. Differential response between the control and experimental populations can be introduced by treatment with drugs, growth factors to one cell population. Proteins from the cells adapted to the light and heavy SILAC media can be analyzed and distinguished by MS.

Once the SILAC labels are incorporated in the cell culture then differential treatment can be applied. This treatment could be the external stimuli exposure, drug treatment, growth factors, immunoprecipitation, and the comparison of differentiated and undifferentiated cells or it could be some other thing. While we are doing this labeling it is a very good idea to repeat the experiment to ensure that the fold changes are uniform and also a reverse labeling should be performed to check for the reproducibility. Because if reverse labeling is also showing similar trend then it means that experiment is unbiased due to any labeling issue.

During SILAC experiments one also need to check for arginine -to-proline conversion and this could be manually adjusted for the expereimental conditions. One may need to reduce the arginine concentration or add proline to the medium to overcome such issues. There are some software's which are designed that can be used for counting the

arginine-to-proline conversion and then experimental conditions can be adjusted accordingly.

In the SILAC experimental workflow, let's discuss the fourth point; cell lysis and protein estimation.

So once the cells have incorporated the labels these cells can be harvested using any standard protocols which one uses in the tissue culture. One needs to obtain a small aliquot of each cell lysate and estimate the protein concentration. We have talked different methods of doing protein quantification earlier. One can use Bradford assay or some other type of protein quantification assays. After the protein concentration determination the protein concentration should be normalized prior to mixing both heavy and light lysates. Because we want to do quantification later on so prior to mixing of both cell cultures, it is very important we are staring with equal protein amount in both light and heavy forms. This process can be done by normalizing by diluting the cell lysates with lysis buffer.

In the SILAC experimental workflow the last and most important point is MS analysis, protein identification and quantification.

As we discussed earlier, prior to the MS analysis, one needs to do DTT treatment to cleave the disulfide bonds, IAA for alkylation and trypsinization for protein digestion. These treatments should be performed and then the combined digested mixture and desalted peptides can be further used. The desalting can be performed by using C^{18} columns. By using the raw MS data files, extract sequwnce-specific MS/MS peak list. Then it can be used for the identification of peptides and proteins using various databases.

The quantification of SILAC-labeled peptide pairs, which are the light and the heavy peptide pairs. So these peptides, for eg. containing heavy arginine are heavier than 6 Da of the normal ones. So these fold abundance ratios can be determined by different methods. Here two ways are suggested to calculate the ratio of the intensities of each

peptide from individual MS spectrum or the ion chromatograms of light and heavy peptides eluted from reversed-phase columns and then it can be used to determine the ratio of area under curve.

In this MS spectra the pairs of chemically identical peptides of different stable isotope can be differentiated due to their mass difference. The ratio of peak intensities of such peptide pairs demonstrates the population ratio for 2 proteins. So here I have shown the light and heavy form separation with 6 Da separation.

Having discussed the workflow for performing SILAC experiments, let's have an overview of the protocol. In SILAC two different populations can be grown in DMEM media containing $^{13}C_6$ stable-labeled form of arginine in the place of normal or light arginine . After allowing 5 or 6 doublings in each protein, arginine is in heavy form now. The cells can be combined and further lysed prior to further proteolytic steps.

So continuing into the protocol, once we have combined the heavy and light populations then these can be separated on SDS-PAGE gel and each of these bands can be further digested by adding trypsin. After digestion of isolated proteins to peptides by trypsin the arginine containing peptides will be 6 Da heavier than their light counterparts which can be analyzed by using MS.

Now I will show you an animation of SILAC. Animation So let's discuss about SILAC method.

SILAC is a simple method for in vivo incorporation of a label into proteins for quantitative proteomic purposes. Two groups of cells are cultured in media that are identical in all respects except that one contains a heavy, isotopic analog of an essential amino acid while the other contains the normal light amino acid. The essential amino acids which are obtained from the cell culture medium are incorporated into the corresponding newly synthesized proteins during cell growth and replication. Medium containing the heavy amino acids will give rise to heavy, isotopic proteins. After a number of cell divisions, all instances of the particular amino acid will be replaced by its

isotopic analog. The grown cells are then combined together and harvested. Centrifugation of the mixture will result in the pelleting of cells which can then be used for further analysis.

The grown cells are then lysed using a suitable lysis buffer and the proteins degraded using a proteolytic enzyme like trypsin. This results in a mixture of light and heavy peptide fragments which can be quantified suitably by MS. The complex mixture of peptide fragments is further separated by SDS-PAGE to simplify the analysis. Each band of the gel is cut out and re-dissolved in a suitable buffer solution. These simplified peptide fragments are then used for further analysis.

Further purification is carried out by liquid chromatography wherein the sample is passed through a column containing a packed stationary phase matrix that selectively adsorbs only certain analyte molecules. Reverse phase and strong cation exchange chromatography are the most commonly used. The eluted fractions are further characterized by MS. The purified peptide fragments are then analyzed by MS/MS. Peptides containing the heavy amino acid show higher m/z than the corresponding light peptide fragments. The pairs of identical peptides can be differentiated due to the mass difference and the ratio of peak intensities can be correlated to the corresponding protein abundance.

The MS/MS data analysis shareware has some extra inputs such as Quantitation, MS/MS tolerance, peptide charge, instrument etc. in addition to the fields for PMF. They require inputs from the user regarding the experimental parameters used such as enzyme cleavage, protein name, modifications etc. and the desired search criteria like taxonomy, peptide tolerance etc. Commonly used protein databases against which the MS information is processed to retrieve sequence data include NCBI, MSDB and SwissProt. The data file generated from MS is uploaded and the search carried out.

After discussing the SILAC technique and watching this animation of this entire process now let's discuss about advantages and disadvantages of SILAC method and compare it with some other tagging methods. As we have discussed and realized that SILAC

method is very simple and robust and it labels the entire proteome with out chemical derivitization and less sample handling steps. The labeled samples are mixed at the stage of whole cells which makes SILAC approach ideally suited for the quantitative proteomic experiment because there will be less handling variations and manual artifacts for each population separately. Studies such as the subproteome analyis, for eg. the cellular organelles or complex purification protocols, can be performed with very good quantitative accuracy by using these methods.

So let's discuss about some of the advantages of using SILAC method. In SILAC no chemical difference is observed between labeled and natural amino acid isotopes. The cells are grown in tissue culture medium and they behave exactly like the control cell population grown in the presence of normal amino acids. So this method is very efficient and very reproducible. It has been observed that the isotope labels are 100% incorporated and one need to ensure that by doing the doubling for different generations. After 5 or 6 generations it has been observed that it has been 100% incorporated.

So samples can be mixed prior to the MS analysis. This method eliminates some of the handling errors. SILAC experiments have demonstrated that they have 5-plexing capability to compare 5 different states in one experiment. So there various applications of quantification of proteins by using SILAC method which we will discuss in the next slides.

Although there are many advantages of SILAC, there are some disadvantages due to the inherent nature of this method. SILAC is applicable only to the culture cells. It can not be used for tissue or body fluids. That is one of the major limitations of this method. The tissue culture process is always time consuming and requires very meticulous and efficient work. There are very limited forms of heavy amino acids available. Due to this limited number of states can be compared using SILAC.

The metabolic conversion of arginine to proline is one of the commonly observed issues in SILAC experiments which results in tryptic peptides containing heavy proline.

Now there are various experimental and biological solutions which can be used to reduce the interference from incompletely labeled peptides.

Let's compare SILAC with radioactivity labeling methods. The SILAC methods seeks to replace the labeled amino acid which is unlike the radioactive labeling that uses ^{32}P or ^{35}S-Methionine. SILAC ensures that the labeled amino acids are fully incorporated in the cells; however, small percentage of labeled amino acids can be detected. In the radioactive-based labeling the full incorporation is not necessary. In SILAC one readout the signals by using MS, whereas the radioactive detection is possible using scintillation counters.

Slide 44: Let's now discuss few applications of SILAC briefly.
The SILAC method is very promising for any cell line. So this method can be applied for any cell line whether it is HeLa cell, C127, HEK293, etc. However, the media formulation and the growth optimization is required individually for each cell line. SILAC applications have been demonstrated in different applications such as cell signaling, studying the induced protein complexes, studying temporal dynamics, identification of kinase substrates and studying differential membrane proteomics. We will have a look on some of the applications now.

So, Ong et al. published a paper in 2002 which was the first demonstration of SILAC application where they used the relative quantification of changes in protein expression during the time course of myoblast differentiation in mouse C_2C_{12} cells.

Researchers have reported various unique metabolic-labeling strategies. For example, by using tyrosine; identification of tyrosine kinase substrates using ^{13}C tyrosine. Labeling is also performed by using methionine. The *in vivo* methylation sites by heavy methyl SILAC.

There are numerous studies based on the global protein expression profiling using SILAC method. I am just highlighting some of the early studies which set the path for performing these protein expression profiling. So the study by Everley et al. in 2004

analyzed the expression levels of more than 440 proteins in the microsomal fractions of prostate cancer cells with varying metastatic potential. Another study by Gu et al. investigated the early stage of apoptosis by inducing the p53 up-regulated modulator of apoptosis.

SILAC has also been used for functional assays to study the protein-protein interactions. Study by Blagoev et al. used the differential labeling of proteins in EGF-stimulated versus unstimulated cells. In another study by de Hoog et al. did quantification of proteins interacting in an attachment-dependent manner with focal adhesion proteins. These are just examples of studying the functional assays and performing protein interactions using SILAC.

The identification of proteins which are enriched in specific cellular structures. A study by Foster et al. used the first proteomic analysis of rafts and they showed that the specific detection of proteins depleted from the rafts by cholesterol-disrupting drugs.

SILAC has been widely used for multiplex analysis to compare the cellular states. Anderson et al. showed the quantitative analysis of the proteome of human nucleoli. Blagoev et al. performed a temporal analysis of phosphotyrosine-dependent signaling networks to compare the proteome of three cell populations. Kratchmarova et al. analyzed the divergent growth factors in mesenchymal stem cell differentiation. These are just few examples of multiplex analysis. Now if you look into literature, there are many studies which have used SILAC method for comparison of cellular states.

SILAC method has also been used to study the protein turnover. Study by Pratt et al. used the rate of breakdown of individual proteins by analysis of mass shifts in tryptic peptide fragments. The analysis of abundant protein in glucose-limited yeast cells which were grown in aerobic chemostat culture at steady state was performed by using SILAC method.

SILAC has been used for identification and quantification of protein posttranslational modifications. Study by Ibarrola et al. identified and quantified phosphorylation sites.

Another study by Ballif et al. also identified and quantified the phosphorylation sites. There are many studies which have used SILAC method for studying posttranslational modifications.

Interestingly, SILAC method has been used in different organisms. Bacteria and yeast, are the more used in SILAC method due to the growth in the cell culture. But there are some studies on Arabidopsis, a plant, as well as in mice that has shown that SILAC can be applied for a wide variety of organisms.

So SILAC applications are straightforward and it only requires some initial attention to the cell culture conditions. Let's discuss the SILAC applications in an animation.

Animation - Let's now discuss the applications of SILAC.
SILAC is a useful quantitative approach that has found applications for several proteomic studies. SILAC provides an in vivo strategy to label and monitor quantitative differences at protein level in different conditions, which has been successfully employed for differential profiling & biomarker identification. Temporal dynamics of cell signaling pathways that transmit information through various PTMs, most commonly reversible phosphorylation, have been efficiently studied by SILAC coupled with MS. Quantitative proteomic studies using SILAC have been carried out with yeast, which serves as a model organism for eukaryotic cells in providing insights into biological processes. Methylation, which is one of the most common PTMs having various biological roles, has been successfully studied using isotopically labeled methionine residues. One of the more recent applications of SILAC include the identification of protease substrates using differentially labeled bacterial cell cultures. Cellular functions are mediated by several protein complexes that interact with one other. SILAC has be been applied for quantitative determination of such complexes and their interacting protein partners. Signaling pathways involving kinases are employed in cell growth and differentiation which play a major role in cancer development and progression. These pathways and effects of inhibitors on them have been studied using SILAC. SILAC

allows for labeling and monitoring of dynamically changing proteomes of sub-cellular organelles which are involved in several activities during apoptosis in cells.

Now let's take a case study from de Godoy and colleagues which determined the fold change of peptide pairs between haploid and diploid yeast cells using SILAC. Labeled lysine residues were used to grow the diploid yeast cells while haploid cells were grown in normal lysine medium. The cultures were mixed, proteins extracted and analyzed by LC-MS/MS. Protein ratios between haploid and diploid cells were determined with high accuracy. Comparison revealed that 97.3% of the proteome changes less than 50% in abundance.

After discussing about different types of applications of SILAC method, now let's summarize what we have studied in this lecture. SILAC in a typical experiment, a control and treated cell lines are grown in different media. One was enriched with light form and another with heavy isotope containing amino acids. The peptides from heavy and light form are mixed in 1:1 ratio and proteins were extracted, purified and digested. The peptides were analyzed by LC-MS/MS. The signal intensity of labeled and unlabeled peptides was able to provide the quantitative information. There are some limitations of SILAC method. The limitations include its usage only in the cell culture and the metabolic convertion of arginine to proline. But overall SILAC is a very promising technique and its applications are published in several papers. I hope by today's lecture you will be able to appreciate different types of *in vivo* labeling methods, the SILAC experiment and the stepwise workflow of performing a SILAC experiment, the merits and demerits of SILAC method. Thank you!