RO-BID-Speaking of Molbio Podcast Episode 9-Global_EXT5617

Audio snippet-15 sec- https://thermofisher-my.sharepoint.com/:v:/p/lhea_lecy/Ea9-Zg05FupKkSqudn537fABtcLrYicDrqFTk-Sv7Grcnw?e=DOZYg3

Full length video(does not follow transcript): Full-Length Video - Ep9 Full length

TRANSCRIPT

Full length audio (follows transcript) https://thermofisher-my.sharepoint.com/:u:/p/lhea_lecy/EUNrUqjFcvtNvs0AIDM5FmsBg1kPAKITFjxAs3nO1kj7YA?e=oHNOHJ

Steve Lewis 00:09
Welcome to Speaking of Mol Bio, a new podcast series about molecular biology and its trending applications in life science. I'm Steve Lewis,

Dr. Gabriel Alves 00:19
and I am Dr. Gabriel Alves.

Steve Lewis 00:22
In our first season of Speaking of Mol Bio, we're focusing our conversations on four exciting application areas: CRISPR cell engineering, multiomics, exosomes, and single - cell analysis. And today we're diving into single cell analysis with Dr. Mandovi Chatterjee.

Dr. Gabriel Alves 00:40
Mandovi is the director of the Single Cell Core at Harvard University Medical School. She earned her PhD from the University of Virginia in 2014 and has become an expert in various single cell techniques throughout her time in Cambridge. We hope you enjoy our conversation.

Mandovi Chatterjee, PhD 01:02
My PhD work focused on a system that involved cells from patients suffering from an early aging disease called progeria. In these patients' cells, the nuclear transport was impaired, and I the mechanism of nuclear transport using that those cells as the system. I was initially interested in studying aging and continue my research in aging. And I started my postdoc in a lab at Brigham and Women's Hospital, which is a Harvard-affiliated institute. And I was studying muscle development, because in aging, there is muscle dystrophy. And it was back in 2015, and that is the time when high throughput single cell RNA sequencing was kind of becoming popular. And during that time I also learned about this core, Single Cell Core at Harvard Medical School, which started out of a lab in Harvard Medical School. That lab actually developed one of the high throughput single cell RNA sequencing technologies called interrupts. So the Core was established in 2016, and I started at the

Core as a postdoc in 2017. And so from those very early days of single cell RNA sequencing, and I kind of like, since then, I am working in single cell RNA sequencing. And I have seen the development of the field. It's incredible to see how the field has grown over the years, how it has developed and incorporated so many different state-of-the-art technologies to answer many, many biological questions.

Dr. Gabriel Alves 03:08
You've been involved with the single cell research for quite some time now and even helped pushing the edges in this field. I would like to know from you, what are some of the challenges you had to overcome? And alongside with the technology of reagents and products improving, how has this technology evolved throughout the years, and where do you see it going in the future?

Mandovi Chatterjee, PhD 03:30
In the early days, the single cell RNA sequencing, like all the single cell RNA sequencing experiments, used to be done in plates, where one would sort, single-cell, and put them in individual wells of a 96- or 384-well plate, and throughout the entire workflow that well identity for individual cell had to be maintained. And you can imagine that that can be very laborious, and that also, you know, gives way to, sort of, cross contamination, making mistakes, and so on. Because this workflow is not like one single-step process. It involves, you know, reverse transcription, then purification of the nucleic acid material, then PCR, and so on. The experiments used to be pretty expensive, compared to the amount of data one can get. Also, during this time, like a few very clever ideas were introduced. The concept of cell barcode was introduced. Then people realized the need for quantitating, or accurately quantifying, the RNA molecules within each cell. Which is when the concept of UMI, or unique molecular identifier, was introduced as well. But in 2015, as I mentioned earlier, there were two microfluidic-based droplet methods for high throughput single cell RNA sequencing published in the same issue of Cell: inDrops and Drop-seq. And when I say high throughput, it's actually like hundreds of thousands of cells. That's when the paradigm kind of shifted. So initially, it was only single cell RNA sequencing, which is looking at the transcriptome of the cell. And then people wanted to see, OK, what happens to the other modalities, for example, genome or epigenome? The scientists tried to innovate and incorporate other modalities alongside the transcriptome from the same single cell. So, this is called multiome. You know, being able to connect the transcriptome with the proteome, or, say, the chromatin accessibility, provided a very powerful tool to the researchers to look at, say, a certain disease state, more carefully. And now the focus is kind of going towards, you know, spatial transcriptomics where you can study an individual cell in its native location.

Dr. Gabriel Alves 05:44
Having the cell in its natural environment in a bulk tissue, getting data from that, is most likely more valuable. What are other technologies that you see coming out of there? Maybe also proteomics? Maybe, is it a possible way to see the proteins? We have talked about in this podcast about exosomes, maybe collect the exosomes from that cell and analyzing its content, the multiomics of its content. So, I would like to hear more from you.

Mandovi Chatterjee, PhD 07:14

It is mainly the cell surface receptors that are being studied in addition to the transcriptome from the single cell. And these are mostly antibody-based approaches, where you label the cells with certain antibodies, and a lot of antibodies, actually, a lot of cell surface receptors can be studied in one single experiment. In a way you can think about it as combining FACS sorting with single cell RNA sequencing—you are labeling the cells the same way as you would for FACS sorting. The difference is these antibodies are conjugated with a barcode—the barcode is a short oligonucleotide sequence. So because you're not relying on the detection of fluorophore, you can imagine that you can actually label many different cell surface receptors at the same time. This small oligo that is conjugated to these antibodies has a poly(A) tail. These high throughput single cell RNA sequencing technologies, they capture the mRNA, or these oligos, from its poly(A) tail, because the barcodes that is used for capturing either the mRNA or these oligos have oligo(dT). So, along with capturing the mRNA within the cell, these barcodes also capture these oligos, barcoding oligos, conjugated with the antibody that has the poly(A) tail. And then rest of the workflow can happen like usual. So, that's one approach. The other approach is that sometimes the barcoding oligo, or the barcoding bead, to which the barcoding oligo is attached, has two different kinds of sequences. One has oligo(dT) and the other one has a specific capture sequence. Now if your oligo that is attached to the antibody contains the complementary sequence of that capture sequence that is present in the barcoding bead, you can use that, too, to capture these antibodies conjugated with the barcode.

Dr. Gabriel Alves 10:03
So, each oligo(dT) will have a specific sequence of the barcoding labeling the cell, the single cell you're looking at. So yeah, it makes sense.

Mandovi Chatterjee, PhD 10:15
So the barcoding oligo that is conjugated to the antibody obviously has a barcode sequence of its own. And say you're using four different antibodies. So these four different antibodies will have four different barcodes. And adjacent to the barcode, there is either an oligo(dT) or this complementary sequence to the capture sequence that is present in the barcoding oligo. Then, during the reverse transcription, the captured sequences are reverse transcribed to make cDNA. And then they're amplified and purified from the prep to make the libraries. And because the barcoding step happens during the reverse transcription, these oligos conjugated with the antibody and the mRNA from the same cell has the same cellular barcode. And during bioinformatic analysis, you can find those cellular barcodes and you can tell that, OK, my cell type 1 has cell surface receptor A, you know cell type 2 has cell surface receptor B. It is a very quantitative process; you can exactly say how many cell surface receptors these antibodies bound to.

Steve Lewis 11:49
Yeah, that is very fascinating. I'm curious a bit about barcode design and how that area is having to develop alongside single-cell analysis and many other analytical techniques.

Mandovi Chatterjee, PhD 12:03
The barcode design is somewhat dependent on the technology that is being used for the single cell RNA sequencing analysis. So every single—there are many different commercial players in the field—and every single one of them has their own design of the technology, obviously. Say, for example,

you're using a microfluidic-based droplet approach, where by using microfluidics, you're basically generating thousands and thousands of nanoliter- or picoliter-sized droplets. And you can consider an individual droplet an individual reaction chamber, if you will. Inside that droplet, there is your single cell, there is the reagents necessary for lysing the single cell, for performing reverse transcription, and a barcoding bead. This barcoding bead has, you know, million copies of the same barcode. That barcode contains a cell barcode, an UMI, which will label individual RNA, and an oligo(dT) sequence, which will be used for capturing the mRNA or whatever it is, anything that has poly(A) tail. The goal is to get one barcoding bead with one cell in one droplet. Then within the droplet, there is lysing reagent present. So, the cells lyse, they release their RNA content within the droplet, and the oligo(dT) present in the barcoding sequence in the barcoding bead captures the mRNA. Then reverse transcription also happens inside the droplet. Then after the reverse transcription, you have a RNA/DNA hybrid. Once they're barcoded, you know, they're PCR-amplified, and downstream molecular biology steps happen to make them sequenceable by next-gen sequencing method.

Steve Lewis 14:43
We hope you're enjoying this episode of Speaking of Mol Bio. We wanted to take a quick moment to tell you about the Invitrogen School of Molecular Biology. It's a great educational hub for molecular biology, with rich and reliable technical content designed for new and experienced molecular biologists alike. Check it out today at thermofisher.com/ismb. And now back to our conversation.

Steering to a little bit of a different area, we have some listeners who may not have such a deep knowledge of molecular biology. When you're talking about the single cell droplets and everything happening within that, that term is, it's called encapsulation?

Mandovi Chatterjee, PhD 15:31
Yes, the droplet-based method is called encapsulation. But not, all the technologies are not droplet based. So the "encapsulation" term will not apply to them.

Steve Lewis 15:41
Got it. Okay. So it really is just a very robust technology, but it's also very new, it sounds like. I'm sure back in the 50s, they would have loved to have this technology.

Dr. Gabriel Alves 15:56
For sure. Talking about past and future, you already gave us a little bit of a taste of what is currently happening in the field of single-cell research, like the spatial transcriptomics and the labeling. What are some other technologies that you see appearing in the future, or that you wish that would appear in the future, to maybe help you overcome a current challenge in the field?

Mandovi Chatterjee, PhD 16:25
Spatial transcriptomics is in its early days right now. So, there are a few NGS-based approaches—NGS is next gen sequencing—based approaches. The way it works is there is, if you can imagine, an arrayed slide. This array is basically an array of barcodes. So every single barcode has a spatial coordinate, if you will. And then you lay the tissue over that array, then you barcode the mRNA using that array. And based on that spatial coordinate, you can then construct an image of the tissue. During

the workflow, there is a need of, kind of, H&E stain, or stain the actual tissue with antibody, that you can overlay with your reconstructed image to kind of align. So that's one approach where you would need to do sequencing of these barcoded mRNA to determine, you know, which cell type is present in which coordinate of that array, to build that image. Then there is another kind of approach where it's image based. You rely on the technology called FISH, fluorescence in situ hybridization, where you introduce or you kind of flood the tissue section with thousands of probes, FISH probes. They go and attach to your target mRNA, and then you amplify the signal of that probe by using a fluorescently labeled secondary probe, if you will. And then lots of images are taken during multiple rounds of the hybridization of these probes. And from the signal you then find out, you know, which mRNA is located in which area.

Dr. Gabriel Alves 16:27
I would imagine you could quantify as well, based on maybe...

Mandovi Chatterjee, PhD 18:57
Yes, yes, absolutely. These are all very quantifiable techniques. However, you know like I said, it's like early days of spatial transcriptomics. So, first of all, every single technology needs to be sort of tried and tested across many different tissue types. And there are challenges even in, you know, the challenge starts from the way you process your tissue, to data analysis. And you can imagine that data analysis is quite complex for every single method where, you know, data integration. So, when you are, first of all, studying one modality, say transcriptome, that is like annotation, and what kind of analysis or what kind of biological question you are actually asking dictates how complex the analysis is going to be. Then if you try to integrate other modalities, like cell surface receptor information, or chromatin accessibility information, or in spatial transcriptomics, you're now trying to integrate the spatial coordinate, then the analysis becomes a little more complicated than just simply looking at how many RNAs you're determining for a barcode. For these high throughput approaches, yes, you are looking at many thousands of cells. There is always a tradeoff, between, like, throughput and the sensitivity, I believe, because you also have to, you know, maintain the cost somewhat manageable. These are cost-effective, but still quite costly experiments.

Dr. Gabriel Alves 21:08
For those who single -cell analysis is, or single cell research is a newer field, or have never heard of it, what are some advantages of looking at a single cell than doing the bulk analysis?

Mandovi Chatterjee, PhD 21:20
So in a bulk analysis, you are basically looking at an average gene expression. So what you're doing is you're taking a tissue, and you can imagine that a tissue is very heterogeneous, meaning it comprises of many different kinds of cell types. In bulk analysis, what you're doing is you're extracting RNA from the entire tissue, and then sequencing it to look at what are the transcripts that are present in the entire tissue. And that information is very beneficial when you're comparing, say, a diseased versus a control tissue type. But what you're not able to understand from the data, from that data, is that whether any particular gene is overexpressed or repressed in any particular cell type within that tissue. Say cell type A, in a particular tissue, has 10 molecules of gene 1, and cell type B has 50 molecules of the same gene, you won't be able to determine the difference. What you're seeing is an average expression level

for that gene in your bulk analysis data. That's the power single - cell analysis gives you. By barcoding cells, then you can say that, OK, cell type A expresses 10 molecules of mRNA for gene A, whereas cell type B has 50 molecules of the same gene. And you can do differential gene expression within the same tissue. And you can look at the cell types. That we can also detect a rare cell that is present in a tissue, which is impossible to determine in a bulk analysis. There is a very popular analogy in the field to understand the difference between bulk and single cell RNA sequencing. So if you compare a glass of smoothie with the fruit bowl, in the smoothie, you're basically getting an average taste of all the fruits blended. You cannot tell the taste of individual fruits. Whereas in the fruit bowl, you can savor the taste of a strawberry, you can also savor the taste of a banana or a mango. So here, the glass of smoothie is analogous to bulk RNA sequencing, whereas the fruit bowl is like single-cell.

Steve Lewis 24:31
You mentioned earlier that you have a variety of sample types. Do you have any insight into some of the samples that you get, what might have been your most interesting case that you received?

Mandovi Chatterjee, PhD 24:43
Quite a few, actually; let's take this example of marine worm. They have to be grown and maintained in high-salt buffer—sea water, basically. We were using it for one of the microfluidic-based droplet methods, and these droplet methods, the reverse transcription happens inside of the droplet, which is why there is a restriction as to what buffer you can actually load your cells with. No component can be added or introduced which can potentially inhibit reverse transcription. So high salt can actually have an adverse effect on reverse transcription, or therefore the downstream molecular biology. But then if we load the cells in the regular buffer, which is usually, we use like PBS with a small percentage of BSA, the cells will burst open even before they are captured. So, we were using a technology called inDrops at that time—it was a homebuilt platform, and therefore had a lot of flexibility to kind of modify and customize the platform. At that time we were also making the microfluidic chip in our lab and also making the barcoding beads for inDrops. Someone came up with an alternate design for the microfluidic chip, where we could load these cells from the marine worm in seawater, and right before they get encapsulated within the droplet, they'll get, the seawater will get diluted at least three times, and then it will get encapsulated within the droplet. Because once it's inside the droplet, and then if it bursts open, it's not a problem. But we have to maintain the cellularity up until that point.

Dr. Gabriel Alves 27:11
So, by diluting slowly, you can…

Mandovi Chatterjee, PhD 27:15
Right before it was getting encapsulated within the microfluidic device. So yeah, that was a very sophisticated modification that we were able to incorporate.

Dr. Gabriel Alves 27:33
Very nice, thanks for your response. And as we're getting to the end of the interview here, I have a couple of questions that I have asked all of our guests. What advice would you give to the young researchers, young scientists, who are interested in a career in single-cell analysis? And the second one, what was the secret for your success?

Mandovi Chatterjee, PhD 27:59
Career in single - cell analysis: This is a very interdisciplinary field. Wet bench is as important as the bioinformatic analysis part. And by background and by training, I'm not a bioinformatic analyst; I cannot do, I cannot write, algorithm. But it is important to understand what is exactly happening in the steps for analysis, to perform a single cell RNA sequencing experiment. It is important for the biologists to understand the bioinformatics steps. It is also important for the bioinformatic person to understand a little bit of biology, to actually perform the analysis the correct way. So, it applies to both, you know, populations of scientists, that it's important to know how to communicate with each other. And that's something I found very important, and probably that's one of the things that I have learned on the job. Oftentimes I, when I consult a lot of scientists to help them correctly design a single cell experiment, I try to make them understand, you know, these individual steps. Only then they can make sure that even the biology is done properly. And the same applies for the bioinformatic analyst as well, where it's important for them to understand the challenges and the process of the entire steps before the data is handed to them. And I guess that is the success.

Dr. Gabriel Alves 30:12
That was Mandovi Chatterjee, Director of Single Cell Core at Harvard Medical School. If you're interested in hearing even more of today's conversation, you can view the extended video version of this interview by visiting the URL in the Episode Notes. And consider sharing this episode with a friend or colleague—we love sharing our conversations with more great people. This episode was produced by Matt Ferris, Sarah Briganti, and Matthew Stock.

ABSTRACT

Speaking of Mol Bio Podcast Series

Episode: S1E9

Guest: Mandovi Chatterjee

Thematic topic: single -cell analysis

Title: A look into the fruit bowl of biology

Quotes:

- *There is a very popular analogy in the field to understand the difference between bulk and single cell RNA sequencing. So if you compare a glass of smoothie with the fruit bowl, in the smoothie, you're basically getting an average taste of all the fruits blended. You cannot tell the taste of individual fruits. Whereas in the fruit bowl, you can savor the taste of a strawberry, you can also savor the taste of a banana or a mango. So here, the*

glass of smoothie is analogous to bulk RNA sequencing, whereas the fruit bowl is like single cell.

- *It's incredible to see how the field has grown over the years, you know, how it has developed and incorporated so many different state-of-the-art technologies to answer many, many biological questions.*
- *It is important for the biologists to understand the bioinformatics steps. It is also important for the bioinformatic person to understand a little bit of biology, to actually perform the analysis the correct way. So it applies to both populations of scientists that it's important to know how to communicate with each other.*

Episode Summary:

In this episode of Speaking of Mol Bio, we speak with Dr. Mandovi Chatterjee about single -cell analysis. Dr. Chatterjee is Director of Single Cell Core at Harvard University Medical School. She has spent her career exploring and mastering a wide variety of single-cell` ````` methods. In the conversation with Dr. Chatterjee, our hosts dive into the more technical aspects of this ever-evolving field. We also learn an interesting analogy to better understand the difference between bulk and single-cell RNA sequencing.

Episode Notes:

Dr. Mandovi Chatterjee started her journey toward single-cell analysis as a PhD student at the University of Virginia while studying nucleocytoplasmic transport in relation to progeria. A postdoc appointment studying muscle development in aging opened her eyes to the potential of single cell RNA sequencing. Since then, Dr. Chatterjee's work has focused on exploring and mastering single cell technologies. As Director of Single Cell Core at Harvard Medical School, she works with a team specializing in single cell transcriptomics, epigenomics, and multiomics.

Our hosts dive into the technical aspects of single-cell analysis with Dr. Chatterjee. We learn about the evolution of this field and consider where it is going. An engaging conversation, where we learn more about the evolution of this field, consider where it is going, and develop a better understanding of the benefits of single-cell analysis over bulk analysis. A great way to close out our first season!

Tune in today. Subscribe to get future episodes. Share with a friend or colleague that might enjoy too!

Visit https://thermofisher.com/ISMB to access molecular biology resources and educational content.

Keywords:

barcoding, cell, sequencing, reverse transcription, droplet, single cell
analysis, tissue, analysis, antibodies, cell surface receptors, mRNA, molecular
biology, oligo, technology, transcriptome, spatial, single, high throughput, individual, label