Cell Biology: Cellular Organization, Division and Processes

Dr. Ruwanthi Gunawardane

Director of the Allen Institute for Cell Science

Lecture 31

Stem cells Part II

Using gene-edited stem cells and live imaging to study Cell Biology.

(refer time: 00:24)

(refer time: 00:30)

Hello everyone, my name is Ru Gunawardane and today I am going to tell you a little bit about how we are using gene edited stem cells and live imaging to study cell biology.

(refer time: 00:47)

So, what we are doing with these technologies is to create what we call a Google map of the cell because we believe that understanding where the parts of the cell are, like the major organelles, can lead to a better understanding of basic cell biology in normal cells as well as diseased cells. And our approach here is that instead of studying all the genes in the cell or all the proteins in the cell, because there are so many of those, our approach is to look at the major structures of the cell, which there are much fewer and it makes both the measurement as well as the interpretation of the data much simpler. So, the ultimate goal is to understand the genotype to phenotype connection but in terms of doing live imaging, we are focusing on the major organelles so, that we can understand how the cell is organized and how it changes.

So, that is what I mean here, we are trying to understand the physical organization or the map of the cell. And one of the reasons we are doing this is because we believe a lot of cell behaviours or cell function can be understood by looking at the major structures of the cell. For example, in this image on the bottom right here you are looking at two fields of view from a microscopy image where you are looking at here two cells that are in interphase, looking at the blue is the nucleus, the DNA, and green are the microtubules. And by looking at the microtubules, the image on the right, you can see the cell is dividing right, this is because we believe structure, when I say structure I mean the organelle here the microtubules for example, is related to the function. So, if we can look at the cell

by seeing the different parts of the cell we can start to not only understand how the cell is organized but how what it might do, what its function might be?

So, for the rest of my lecture here, I am going to touch on three different technologies to study cell biology. One as I mentioned just now is microscopy the other is human inducible pluripotent stem cells, which I will get into a little bit more in a few minutes, these are the stem cells we are using and then CRISPR-Cas9, this is a technology that allows us to introduce a tag of fluorescent protein into the stem cell so we can visualize these various structures and CRISPR-Cas9 is the gene editing technology that won the Nobel prize last year for Jennifer Doudna and Emmanuel Charpentier. So, very briefly, our cell model system that we are using to study cell biology here is the human iPSC and this again was a huge revolutionary finding where Shinya Yamanaka won the Nobel prize in 2012 where we can now use an adult cell like a skin fibroblast and reprogram it to be more like an embryonic cell, which has all the potential to develop into different cell types. Usually this is only possible in embryonic stem cells or the stem cells that are in our in our tissues, but now we can do this in vitro, which is why it is exciting and revolutionary for research.

So we are really interested in understanding the rules of the cell, how the cells know to organize themselves in a certain way, which is why we are interested in first mapping the cell and then understanding how that changes when one when a cell goes from one cell type to another like the iPSC, the stem cell, to let us say a heart cell or a brain cell or a liver cell. If we can understand these rules we are basically getting a better picture of fundamental cell biology that when the cell goes wrong in disease we can start to understand why and how to treat it. So, for our model system, we are studying things like mitosis, differentiation and mutations, because these are opportunities for cells to change and for us to observe how the cells go about these changes by looking at the major structures of the cell within the cell.

And eventually what we want to do with all these data is to be able to look at any given cell, whether it is a normal cell or a diseased cell and predict what it might do in the future. So, this is kind of a long-term goal but with all the computational methods out there like machine learning, you can teach these computational models to understand these relationships and maybe even predict based on a cell shape or where the organelles are, what it might do in the future and what it might not.

(refer time: 05:54)

So, where do we start? So, a few years ago we started this institute by asking let us really map the organization of the cell, where are the major landmarks of the cell like the nucleus, the ER, the Golgi etc. We think we know where they are if you look at any cell biology textbook you can see these different structures nicely in this cartoon. But let me tell you that no single cell probably looks like this, and no two cells look exactly alike even in the lab.

So, we wanted to start asking that question, what does a normal cell look like if we actually look under the microscope and what if we look at thousands and thousands of cells? Do they look the same, do they look different, what are those rules, can we map it? So, that is what we have been doing for the last few years; we have been using live 3D high resolution imaging and we have been doing it in replicas, lots and lots of replicas. So, that we can develop a much better model than what

we have seen in cartoons, what we call the integrated cell where we integrate all the structures into one cell and at the end of this lecture I will show you our first models of what this looks like that came really from our own data, data-driven model of a cell. And of course, you can do this through the various cell behaviours or functions that a cell, especially a stem cell, can take.

And it is very important for us to really understand what is normal, what is normal variability, because if we want to understand what happens in disease, we need to get a really good picture in cell biology what the variability is, that is natural variability and what are the relationships between the different organelles as the cells are varied.

And we do understand that just looking at cell organizations not going to give us the whole picture, genes are a big part of it going from genotype to phenotype. So, we plan to introduce gene expression studies, along with the cell structural studies. So, what can we learn from doing these measurements? We believe we can understand the rules or the principles of cell organization, what are the structures that drive organization and how? How do the cells go from one state or cell type to another and what are those correlations, what are those transformations? So, that we; can understand how we can modify that or even fix it if we are talking about a diseased cell.

(refer time: 08:33)

So, the way we went about this was to create this a collection of cell lines that we can use in the lab so that we can see all these major structures of the cell. For example, the adhesion complexes, the microtubules, the nucleus, which we you know the 30 to 50 major structures in the cell that we know are very important for the cell, every cell has them. And then we identified a gene that expresses a protein that is needed for that structure.

So, let us take microtubules as an example; we know the microtubules one of the protein components is alpha tubulin. So, how we did this to visualize alpha tubulin as kind of the reporter for that structure is to tag that gene in the genome. This is where CRISPR-Cas9 came into play. So, CRISPR-Cas9 is a two-component complex where the Cas9 is a nuclease, it goes and cuts the DNA where you like it to; the CRISPR guides the protein to that location, for a double strand cut. And then we can introduce a mutation or in our case a fluorescent molecule so that we can visualize that protein. So, we are creating a recombinant protein in a way inside the cell. So, we use our WTC11, that is the name of the stem cell line we are using, it was derived from a healthy individual and it was reprogrammed to create this stem cell line. So, we use CRISPR-Cas9 then to go in there and introduce a tag into this gene of interest.

So, we have done this for now all the 30 to 50 structures. So, if you have a collection of cell lines, we are calling the Allen cell collection, that is available to everyone and if people are interested in learning actually the technical details of how we did this, I really welcome you to visit our website allencell.org and in there we have some video tutorials you can watch. So, we go into the genome and introduce in this case I am showing an example where the gene we are tagging, the C-terminus or end of the gene, with this MeGFP tag.

So, we provide the template for the cell to repair itself after the double strand break. So, we are kind of tricking the cell to introduce this tag into the gene. And then we make clonal cell lines because it

is important to us to create robust and consistent data and to do that we create one clone that results from this gene editing experiment and then we use it for our live imaging studies as shown here on the bottom.

But we also differentiate these cells into other cell types like the cardiac heart cells and we make it available to the community. So, it took us a couple of years to figure out how to do this in a consistent way to develop the editing strategy, the workflow, which I will show you in a minute, and the quality control measures. So, we know for each line we are checking the right boxes in terms of the cells tolerating this gene edit and behaving as they should, as if they were not edited so that we can do the live imaging for all the major structures to the cell as well as the cells that we differentiate from these iPSCs.

(refer time: 11:59)

So, first I want to show you what these stem cells look like if we look under the microscope. So, they grow as a colony here on the left here. So, this is a colony where there are probably hundreds of iPSC cells, they are dividing, they are moving. They are grown on a substrate called Matrigel, which contains many, many proteins that the cells attach to, and they are very happy and so, this is a movie of three hours sped up. So, they do not like to live alone and they like to move. So, it is a very dynamic epithelial sheet. So, it is a real good opportunity for us to understand epithelial like cell biology with these stem cells, and on the right I am going to show you a fluorescent image, actually it is also a series of images going from the bottom of the cell to the top; these cells are pretty tall and we have stained these cells with two dyes. one to see the nucleus-the DNA and the other is to see the cell periphery of the cell membrane.

So, I will click that and you can see going from the bottom of the cell to the top, you can see the nucleus-the cyan is very big compared to the cell, and you can also see some of the cells dividing some of the DNA here. So these are the cells we are starting with for our imaging where we want to see the individual structures within the cell without having to add a dye to see them.

(refer time: 13:28)

But we had to overcome some challenges to do it. First, stem cells are hard to work with, they require a lot of careful maintenance to make sure they retain their stem cell-like properties so that they do not always differentiate into some other cell type, which is called spontaneous differentiation. So, we want to keep its pluripotency, its ability to remain its stem cell features and differentiate into other cell types, and that is associated with this beautiful morphology, that is seen in the colonies that I just showed you.

Then the editing is complicated as well. We want to make sure we edit the right locus and we do not introduce any other mutations while doing so. So we do whole genome sequencing and we also do a lot of parallel experiments because we do not know at the start of the experiment, which kind of CRISPR will work, even though it is based on the sequences and the delivery.

So, we have to deliver three components into these stem cells to get this editing to work; we need the Cas9 protein, which does the double strand break, we need the CRISPR RNA, which guides the Cas9 to the right location, for example the tubulin locus versus some other locus, and of course our template, the GFP. So, we can trick the cell to introduce the GFP, while it is repairing that double strand break.

And of course, this is not a very efficient process; it is only at best five percent of the cells get edited, most likely less than one percent cells get edited, because it requires this repair mechanism called homology directed repair. So, we have to enrich the cells. So, we use fluorescent activated cell sorting to do that and that is also complicated because the cells have to go through this traumatic process during enrichment. So, again this was part of the couple of years we spent to develop methods that were robust; they were consistent and we could do for many, many thousands of cells and across the 50 different lines we have created.

(refer time: 15:40)

So, we have generated now a workflow to do this, which is pretty standard, pretty routine in the lab. This is where we design these various components for each locus and we introduce them into the cell, which is called the transfection. And then we have this edited population of cells where only a few are fluorescently tagged but because the few that are edited are fluorescent we can use FACs to enrich for them based on the fluorescence signal. That doesn't always mean the editing was accurate. So, we make about 100 lines, 100 clones, and then do genotyping, PCR and sequencing, to look at the locus and make sure the edit was correct. And then we have several other assays or experiments we do to make sure the cells are okay, which I will show you in the next slide including do the cells have the ability to differentiate into other cell types, which is a hallmark of a stem cell. So, we check for that. So, we have done this now for over 50 different loci or genes and initially we were so excited to see that this gene editing CRISPR-Cas9 method works so well in our stem cells that allows us to edit very very essential genes in the cell but of course, the efficiencies depending on which gene and which CRISPR we use. And we are only trying to introduce this tag into one of the alleles, not both because we believe editing both the alleles would probably be harmful to the cell. And of course, we do extensive amount of quality control to make sure these cells are okay are happy and behaving as they should because that is the whole point for our studies.

(refer time: 17:27)

So, I want to compare a traditional method that cell biologists have used for many, many years versus this endogenous tagging that we are doing, which is introducing the tag to the gene. So, that it is only expressed when the gene is expressed. Most of the time what we have done in the past is called transient transfection where we introduce, basically overexpress the protein, so that we can see it. And on the left and on the right are just the two different editing methods for mitochondria, the gene is Tom20. On the left you see overexpression has a lot of artifacts only some of the cells get transfected and also it is not really endogenous meaning natural, whereas on the right when we do CRISPR-Cas9 gene editing, we can pick cells or populations of cells that are additive. So, it is

homogeneous and it is natural expression of the tag of the protein. So, it is more normal. So, I think it is a game changer for cell biology, versus over expression and transient transfection.

(refer time: 18:36)

So, as I mentioned we do a lot of quality control. So, we ask several questions while doing this. We ask is the tagging precise? So, is the location of the tag correct, is it at the right locus, have we only introduced one tag and have we introduced any other double strand breaks or mutations on either allele? So, we do a lot of sequencing to make sure we do that; again the details you can find in our paper, which I will show you at the end or on our website. We ask is the tag protein expressed and localized to the correct structure? For example, if we tag the microtubules we do not want to see the fluorescence all in the nucleus right. So, we compare this with unedited cells and we also do western blots to make sure the expression level is correct and we ask if the protein is stable by growing the cells for many, many passages or many, many life cycles and ask if the protein is still there, the tagged protein. And we also make sure the genomic stability is okay, meaning all the chromosomes are still intact and that is by doing an experiment called karyotyping.

And we do things like cell cycle, growth rate and dynamics of the tagged structure, this is just basic cell biology and of course we want to make sure the stem cells retain some of the stem cell markers, the morphology is the same and have the potential to become other cell types. And this is done again with some imaging as well as flow based assays to look for expression of proteins that are known to be expressed in stem cells and making sure they are expressed at a very high rate still, like over 90%.

(refer time: 20:26)

So, again this clonal cell line generation and qc takes a long time; it takes us about four to six months to do this for each line of course we do many lines in parallel and here are the 17 different assays we do. I do not plan to go over all of these except to say that we start with 96 clones and we end up with the one or two clones, then we release one to the public. But the initial assays, the genotyping assays, which use PCR and sequencing help us rule out a lot of clones that have imprecise or inaccurate editing.

So, at the end of the day, the clonal recovery, we have certain metrics that we track; once we pick the clones that have edited correctly and genomically, the rest of the assays usually end up working pretty well but we do want to make sure that that is true, that our final clone is behaving as we would like and expect.

(refer time: 21:28)

So, we share all these methods and data on our website as well as a couple of papers we have published and there are some video tutorials where you can see the scientists go about their day doing the experiments, along with protocols that go along with that. And we have a catalog where

we can go in, anyone can come here click on their favourite structure, whether it is the nucleus or the mitochondria, and look at some movies that we have taken of these tagged structures both in 3D and over time.

And if you are working in a lab and want access to these cells there is a way you can get them to your lab as well. But if you just want to look at the data and play with it, there is a 3D viewer where you can rotate the cells, pick the structures you want to look at and explore, and even do some basic analysis of the cells. And some of these tools, what we have learned, some of the cell stories are all available on our website as well and we call it the cell explorer. So, we can explore ourselves.

(refer time: 22:39)

So, we have created about 50 lines so, far. So, we think most of the major structures are tagged now and this is called again the Allen cell collection and they are isogenic because the background cell line is the same for all these lines; the only thing that is different is the fluorescent protein is associated with a different protein. So, that we can visualize the structure that that protein localizes to and all of these have undergone the rigorous QC we do and there are a few details here that I have included that I mentioned already and both the plasmids we use to generate the line as well as the cell lines are available to everyone including the papers, the methods, the tutorials. But I want to add that for people who are not using stem cells but using human cells, the plasmids we have generated can be used to generate similar lines. So, very exciting.

(refer time: 23:40)

So, I want to spend a few minutes showing you what these cells look like because this is where we are discovering the cell biology that most of which we knew already but now we get to see it in real time but other things that we did not know that we are discovering. So, here is a movie of an epithelial stem cell colony tagged with the microtubule protein, the alpha tubulin. So, you can see a lot of cell division happening here right, this is a sped up movie of course but you can see how dynamic this colony is. The colony is also growing as the cells are dividing and this is using confocal microscopy, a spinning disc confocal microscopy, this is time lapse and then here's a higher magnification image and I will play this and you can see this one cell dividing, it is just beautiful.

So, because the cells are endogenously tagged and not over expressing this fluorescent protein, it is really clear. And since all the cells have similar amounts of the protein it is also easier for us to see and extract this information.

(refer time: 24:58)

I will show you a couple more, here is the nuclear envelope, the lamin B1 cell on the left is the colony and you can see a cell that has just divided here on the left and on the right I will show you a time lapse. So, you will see those two cells kind of 5 o'clock and 11 o'clock, going through cell division. And I will play a couple of times and what was surprising to me here is that the nuclear envelope

does not completely break down during cell division something a lot of people do not appreciate that they do not completely go away. So, again you can see how clear this is with the endogenous tagging.

(refer time: 25:40)

And I will show you three more because it shows that the different structures within the cell do different things just using cell division as an example. And what we are excited about is that you can do similar studies on any biological process you are interested in whether it is that cell death or stress or something else. So, here is a nucleolar protein, fibrillarin, which completely disappears or melts away during mitosis and comes back and you can see some of it disappearing in this movie and coming back here.

Here are the tight junctions. So, the tight junctions are at the apical surface of the cell, this is what retains that epithelial sheet because it does not actually go away during cell division even though you can see the colony is moving. And on the right is our plasma membrane cell line, we have just pseudo-coloured, the pink is the bottom of the cell, you can see the cell is exploring its substrate that it is growing on and the green is the top of the cell. Again, this was not editing difference, this is just pseudo colouring.

And you can see as the cell is dividing it rounds up and the cell the green kind of retains that morphology. So, to summarize this endogenous tagging is a game changer in cell biology because now we can visualize what is happening in the cell, the parts of the cell how they are organized and how they are different as they go through a process.

(refer time: 27:13)

And one of the big challenges that we have had to overcome is to extract this information from each cell in 3D in live images. So, none of these structures look the same. So, extracting this information, what we call image processing, and segmenting out those structures. So, we can analyze the data was a major challenge, just even segmenting out the different cells, because they grow so close together in an epithelial colony, was a challenge. And I will just show you this kind of in 3D, just to show you how they are all very, very different, and the technical challenge that was posed, that we are working through, they are not all done, but we are working through it. So, that we can really in a mathematical way extract that information.

(refer time: 28:08)

So, what can we learn from doing this? I will give you two examples again we are at the beginning of really analyzing this data and gathering insight. One is very excitingly we can compare what the different structures do at the same time in the cell cycles on the left here. This is just mitosis separated out into various categories based on what it looks like, the DNA looks like. And now we can ask what is the Golgi doing and what is the ER doing and how are they related?

And on the right here is a newer study that you can read in depth using this reference, where we have integrated 25 structures in 200000 cells. So, that we can look at natural variation in a colony and understand do these different structures change in different parts of the colony. How much of each structure is in each cell and what is the relationship between them? And this is kind of the exciting new insight we are starting to have by doing this kind of imaging and visualizing each structure individually, and integrating them together in one single cell.

(refer time: 29:26)

And what we are doing next is to kind of take those, what you can say, the principles we learn. And ask do they apply as the stem cell goes from being this very pluripotent epithelial cell to losing that epithelial morphology as it goes through things like the epithelial to mesenchymal transition during differentiation? So, this is one of the main reasons we use stem cells because they can go into these different lineages, which also happens during early development.

Although of course this is not going to truly mimic development, because this is not in a 3D tissue, this is in in vitro, but some of those principles might still apply and that is why it is so, exciting to do this kind of live imaging and to do it at scale which you can do in the lab. So, we want to create not only a map of the stem cell but also a map, the Google map, of these different cell states or stages as they go through differentiation.

So, that we can understand how the cell changes its behaviour and how that is related to changes in organization and so that we can understand the rules or the principles of cellular morphogenesis. So, we have also started to do this with heart cells that we can make from the IPSCS.

(refer time: 30:59)

So, there are methods to do this in vitro by using different chemical compounds and growth factors, hormones, that make the stem cell stop being a stem cell and differentiate into other cell types and again these details and protocols are on our website. So, we have done this, we do this as part of our quality control when we make a gene edited cell. So, I will show you two examples, it is pretty amazing to watch how those colonies I just showed you before, the stems cell colonies become this beating tissue on a dish in the lab.

On the left and on the right are just two examples of this, which is about two weeks after we add these molecules into the media, the growth media that these cells the differentiation media these cells grow in. So, on the left is a sheet of cells that used to be colonies that are now just synchronously or wave-like way beating in a dish. And on the right these cells are beating so fast they are almost coming off the dish that is why you see those holes in the middle and they almost look like it could be a heart but of course this is very, very early muscle cells that we are making in the lab using these human-inducible pluripotent stem cells.

So, this is kind of an exciting time in cell biology where we can compare and contrast different cell types and look under the microscope and see these amazingly different behaviours. But then go inside the cell with that tag structures and see how the cell is organized differently.

(refer time: 32:49)

And that is the exciting part of what we are doing. So, in terms of gene editing, we have gone and also edited substructures of the muscle unit the sarcomere, which is very important for the cardiomyocyte and we have created additional tagged lines. So, we can visualize these pieces of the sarcomere as the cell becomes a cardiomyocyte, because we can now understand the cellular heterogeneity in the cardiomyocyte as well as the stem cell. But we can also watch as these parts come on in a way, as they are expressed and assemble a very important organelle or structure the sarcomere and follow that during maturation. So, in vitro in the lab what we grow cells that look very, very early, they are not like the adult heart at all, they do not have all the other cell types but we can start to ask questions like what if we introduce a mutation that is associated with heart disease? What do these cells look like, does it change how the sarcomere forms assembles and mature.? So, this is again a tool we can use in vitro to better understand disease and see how that is different from what a normal cell looks like.

(refer time: 34:09)

So, I will show you two examples where we have tagged again the white is the fluorescence of the sarcomere; on the left is a Z-disc which is one component of the sarcomere and on the right is titin which is another component of the sarcomere and these are aged cardiomyocytes on the right, day 50. So, again the point is to show that these cardiomyocytes are a beating, which makes imaging a bit challenging but you can see how regular these white patterns are, the railroad-like patterns, which is a classic feature of the sarcomere because it is a repeating unit of the muscle fiber.

(refer time: 34:55)

So, another aspect of our studies, like I mentioned right at the beginning, is not only to look at how the different parts or the pieces or the landmarks of the cell are located, and the relationship between them, but to ask what about gene expression? We know gene expression also changes as the cell changes. Is there a way to connect that to how the cell is organized?

So, we asked that and we asked whether we could build a framework for mapping both gene expression and structural organization because one of our goals is to see if we can predict one from the other so, that we do not have to make all these measurements in a single cell. So, we did this study in the cardiomyocytes that I just talked about, and I would recommend you to take a look at that because it is actually quite a number of pieces we had to put together to connect the dots between gene expression and cell organization.

So, what we did was to do RNA-seq, a single cell RNA-seq study, where we looked at the changes in gene expression as the stem cell became a cardiomyocyte. And for this lecture I am not going to go into details but say then we took those genes that changed the most in the cardiomyocytes, that was one aspect of it, but then we asked whether those changes could be related to the changes or the variability we saw in the sarcomere, which I just showed you in the previous slide in different aged cardiomyocytes.

And the short answer is no; there was really no strong correlation of these genes that change the most, to the structural changes we saw in the cardiomyocytes, which is interesting, it just shows you that maybe you cant just look at gene expression alone and nor can you look at organization alone. I think you have to look at both if you really want to understand what a cell is doing.

(refer time: 37:03)

So, what we were able to do is to take these images of structure, which is again showing the sarcomere here, and kind of pseudotime organize them from the least organized to the most organized to the right where you see the regular aligned sarcomere patterns and the blue is the nucleus, the white again is the tagged actinin2, which is localized to the sarcomere. We took the same cells and asked what about expression of the two most common genes found, differentially expressed genes, found in cardiomyocytes at this stage, which are two myosin genes myosin heavy chains 6 and 7.

And we did fluorescent in situ hybridization or FISH, RNA-FISH to look for gene expression. So, on the bottom panel here there is the same cell as the one we imaged for the GFP; you can see pink is the younger gene and the green is the older gene. So, what we expected is to see it go from pink or magenta, to green to the right, right? If the gene expression changes were associated with structural organization changes, and you can see yes maybe it is changing a little bit from pink to green but there are cells here that look very disorganized that have the older gene, most of the cells have the older gene expressed. And that is the point here, that you cannot assume because the gene is changing and as a population there are changes in gene expression, that any single cell you could predict what its structure might look like, based on gene expression alone, and vice versa. If we had just looked at this structural organization, I might have predicted that this cell would mostly be pink in terms of gene expression, but it is not.

So, what we are understanding again from our kind of analysis is that both types of measurements should be done in the same cell and this should be done in a single cell analysis.

(refer time: 39:10)

So, we hope to do that in our studies as we explore cell state changes during differentiation. So, we are kind of studying the middle now, we have studied the stem cell and the cardiovascular cell ,now we are trying to approach the middle, how do you get there how do you get from point a to point b. And the last part I want to mention here is that a big part of cell state changes that we want to associate in terms of cell biology, is the change in the nucleus. Because that houses the DNA and we

know the 3D architecture of the genome within the nucleus goes through huge changes as the cell goes from a stem cell to a different cell type, and we want to visualize this we do not want to just sequence it, we want to visualize it.

(refer time: 40:01)

So, to do that we are now starting to do gene editing to have cell lines that represent the different parts of the nucleus that regulate that 3D genome. And to do that, we are creating a collection of nuclear lines, about 20 of them, that mark the major landmarks of the cell. We have done some of these already like the nuclear lamina and the nucleolus, but we are adding others and this is where we are also creating lines where multiple substructures are tagged; we are also doing things like chromatin, RNA polymerase and centromere. So, we get a really broad and diverse understanding of the changes to the nucleus. And here are some examples of the cells we have created. So, far I am not going to show you movies, these are just stills but if you look at just the replication complex PCNA it is very exciting because what we are finding is that we know we can tell what part of S phase each of these cells are based on the pattern of PCNA and RNA polymerase 2 as well so, and the histones.

(refer time: 37:03)

So, this is a really exciting time in cell biology because we can explore cell states in terms of the cytoplasmic structures as I showed you at the beginning of my lecture like the microtubules, the actin, the Golgi, the mitochondria. But now also the nuclear structures and then relate that to sequencing information; there is a whole host of nuclear information that we can get from things like ChIP-Seq and Hi-C that we can connect with this to get the spatial information. And of course start with the stem cell and watch as the cell changes during differentiation.

(refer time: 41:57)

So, to summarize, I talked to you about how we are creating this visual model of stem cell states and phenotypes. Starting with the physical organization of the cell, I showed you some examples from mitosis and differentiation. We are just starting to introduce mutations and other perturbations to explore these stem cell states and how we are starting to build these models where we put things together to understand the relationship between different structures to each other, as well as gene expression to structural organization.

(Video Starts: 42:39)

So, that we can go from these cartoons to much more realistic cell models and I want to show you one example. So, today I did not go over the computational pieces that allow us to take one structure and combine it with another structure. So, here is one example, these are two cells from a

colony that were sitting next to each other. So, this is real data in terms of the morphology and the cell but the only thing that was tagged in those two cells is what you are seeing in green here.

So, one cell is in mitosis you can see and the other cell is not in mitosis, all the other colours you are seeing the purple is mitochondria, the blue is the Golgi, the yellow is actin and the red is a nucleolar protein, were from what we call label free methods, which is basically we have trained a computer model to learn based on cell shape and nuclear shape where those structures are, and that is because we have trained that model from looking at thousands and thousands of cells.

So, it is learned where they are, and they can tell from just a bright field and a non-fluorescent image where those structures are. So, we can kind of superimpose it. So, this is one way we are integrating many structures into the same cell, because technically you cannot tag 20 loci in the same cell right now; maybe one day we could. So, this is one approach we are using to create these models and there are actually three or four others that we are working on.

(Video Ends: 44:10)

So, what we are working towards is combining images, fluorescent images to the left, where we follow the dynamics of how the cells change with RNA-seq type of sequencing information. So, we know what are the genes that change? And then visualize them in the same cell like I showed you for the cardiomyocyte and integrate all those data, I only showed you two genes in that study, what if we can do 50 genes 100 genes. And what if we look at all 50 structures? What are the relationships we can build and can be using that information on thousands and thousands of cells be able to predict what a cell might do in the future. And what we hope to gain is to get at better signatures of cell types and states, better definitions of what they are. So, we can come compare normal to abnormal and also learn in terms of cell biology how do cells change, when it comes to its components whether they are the genes or the structural organization?

(refer time: 45:18)

So, we share all these tools with the community again. The cells we have created are being used around the world right now to study various tissues in the dish. And we are very excited to see what they will find and what I did not show you today, I alluded to, are all the image processing and image related computational tools we have developed so that we can extract that information, gain insight and apply them and combine them with many many cells and data types.

(refer time: 45:53)

So, all this is available again for you to explore yourself, whether it be for educational purposes or just for fun. I hope you do that.

(refer time: 46:06)

And I want to end by thanking the cell science team; we are a team of about 70 people from different backgrounds and expertise, all the way from molecular biology, cell biology, coding, imaging, animation that have come together to do this big science approach, team science approach of studying cell biology using the three technologies I talked about today, which is using iPSCS, doing live imaging and using gene editing. As well as all the computational methods I did not talk about today.

(refer time: 46:46)

And I would like to end by thanking our founder Paul Allen for his vision, encouragement and support. And with that I will thank you, and I will end now. Thank you, everyone.