An Introduction to Proteogenomics

Dr. Sanjeeva Srivastava
Prof. Abijit Dixit
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay
Bio- Rad Laboratories (I) Pvt. Ltd. India

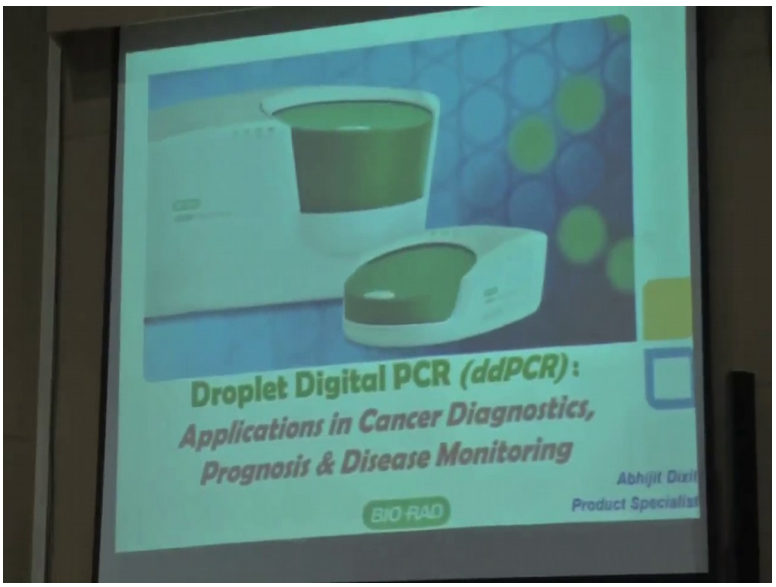
Supplementary – 6
Genomic Analysis using Droplet PCR – I

Welcome to MOOC course on Introduction to Proteogenomics. Till this week, you have heard Dr. Mani, Dr. Ruggles, Dr. Bing Zhang, that why study of mutations and gene sequencing is so important to understand the clinical conditions. In today's lecture, Mr. Abijit Dixit will talk about droplet digital PCR. He will talk about the three generations of PCR; quantitative PCR including the normal standard polymerase chain reaction, relative quantitative PCR including the real time PCR, an absolute quantification which includes droplet digital PCR. He will also talk about the basic principle and workflow of the technique.

He will then talk about how partitioning of templates facilitates the finding of real mutations in the gene with normal PCR due to the low abundance. He will also talk about how this technology platform enables us to a study the rare molecules even after the presence of the inhibitors. So, let us now welcome Mr. Abijit for his today's lecture.

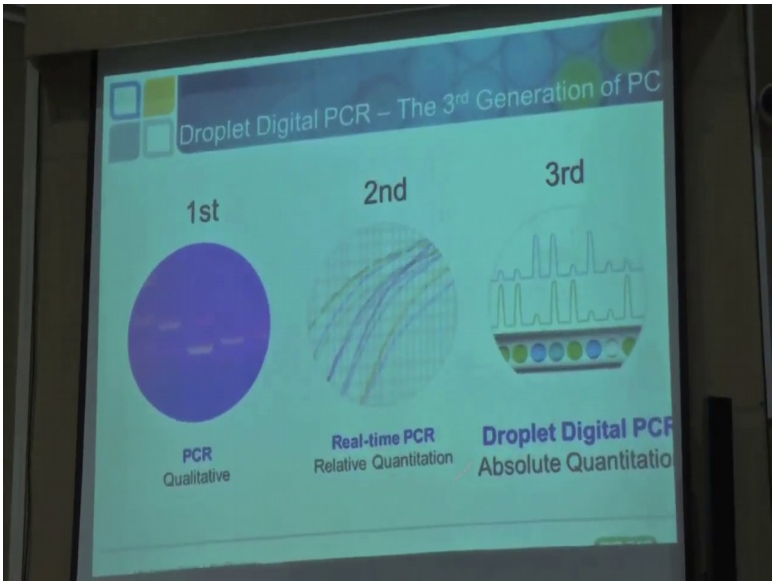
My name is Abijit. I work as a product specialist for Bio-Rad looking at the genomics portfolio. I know it is a really taxing and a long session you guys have had dealing with this part right. It would be my effort to drag you here right. I would sincerely try to get it through; if there is anything that is not clear or you guys are not able to take it up, please do let me know, it will be my pleasure to clear it up for you yeah. So, here we are today to talk about a technology called as the droplet digital PCR.

(Refer Slide Time: 02:13)



As I mentioned I guess most of you know what at least a PCR is ok.

(Refer Slide Time: 02:19)

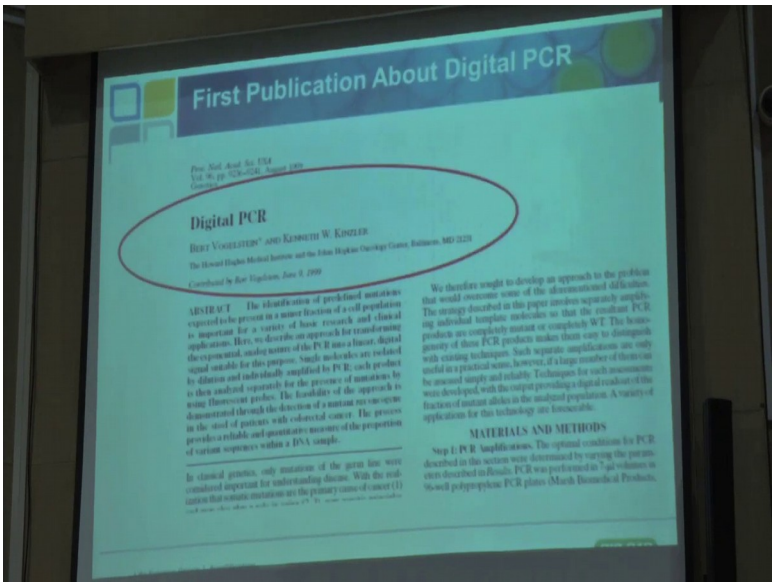


So, everybody here, I mean they know what a PCR is, this is what we do. The PCR is generally evolved over technologies ok. The first generation of PCRs where you put up a PCR to know a presence or absence of a particular gene, you have to run a gel right. A presence of a band would indicate the presence of a particular gene; a band is absent that means, the gene is not present ok, so that old traditional PCR is called as a qualitative PCR, you know the presence or absence of a particular gene.

Second as I went a step ahead to let you guys know what a real time PCR is ok. A quantitative PCR if you want to know the amount of DNA or RNA expressed for a particular gene right, you can go around and do a quantitative PCR. But for those of you who have done real time PCR you know, it is a relative quantification. To know the amount of nucleic acid, you guys have to run a set of standards or run a housekeeping gene, and then you compare it with that expression to know the quantity of DNA or RNA that is present, so that is why to call as a relative quantification.

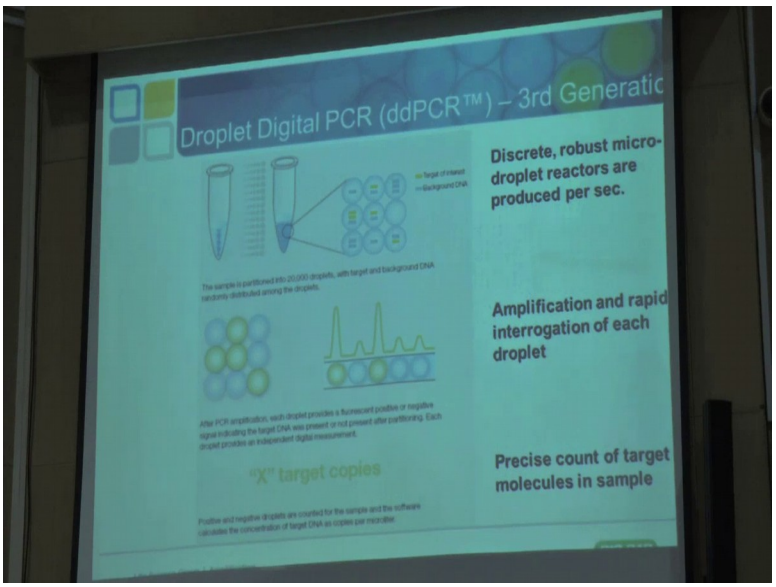
What we have come up with now which is called as a third generation of PCRs is the droplet digital PCR. The major difference being it is an absolute quantification, so that you do not need to run any set of standards or housekeeping genes to know the amount of your nucleic acid in your sample right. In our presentation through I will brush upon what the technology is, how does it work, how is it unique from the previous PCRs right and what are the advantages.

(Refer Slide Time: 03:57)



Again for those of you feel this technology is new, it is not. The first publication which had come with digital PCR was way back in 1999, but then it is been over one and half decade that took for the technology to get evolved and reach this level where it is today ok.

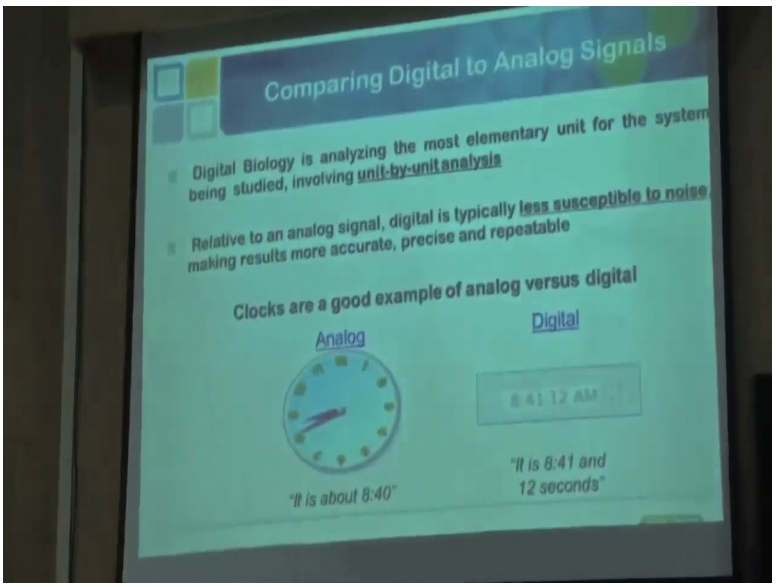
(Refer Slide Time: 04:15)



So, what happens generally in this technology? I mean obviously, a all of you have put a PCR know what a PCR is we make a same reaction mix right. This reaction mix will have your nucleic acid let that would be a DNA or RNA, your specific primers or probes, your enzyme right. We make a same thing here as well. But what we do is instead of putting this reaction for PCR, we split this 20 micro liter reaction into 20,000 droplets. These are oil in water emulsion droplets right, and then we put a PCR for those droplets, and read those droplets individually.

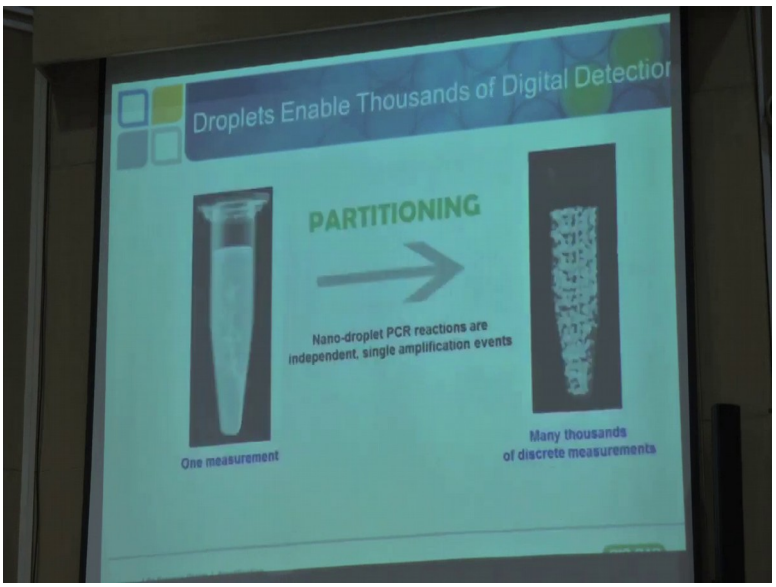
So, instead of one reaction which you put in a PCR and now you have 20,000 reactions going together that is going to give you 20,000 times in more specific and sensitive data as compared to a real time PCR technology. And again the data that you get here is directly in copies per micro liter also DNA. So, you do not have to do any manual calculations to get to a data, the software automatically analyzes and give you the precise amount of nucleic acid in your samples right. This is just an analogy to compare what a real time or a normal PCR versus this PCR is. This is the clock right or these both of them mean the same.

(Refer Slide Time: 05:29)



But what you see here is the time is somewhere around 8.40 right. But once you go here, this is the digital result that you get. A more in depth, a more precise, a more detailed analysis of your data; this is what a digital PCR or a droplet digital PCR does as compared to the traditional PCR technologies.

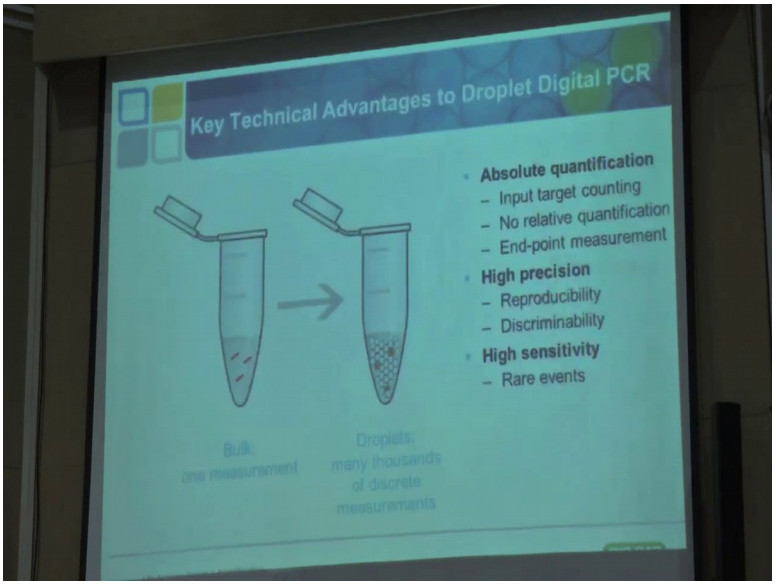
(Refer Slide Time: 05:47)



As I mentioned, the main principle of this technology is partitioning. You have one reaction otherwise which you put in a normal or real time PCR which is partitioned into 20,000

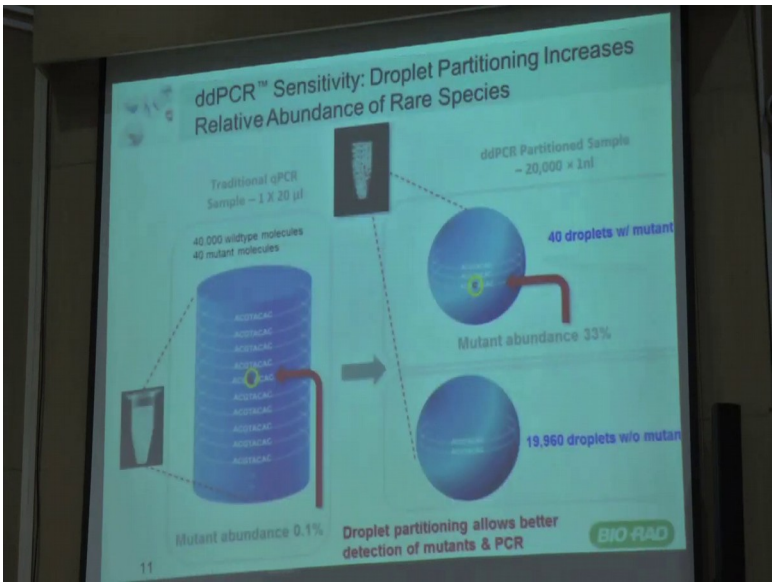
reactions, and then each of these small droplets works as a small micro reactor and undergoes a PCR individually, so that is a main principle on with this technology works partitioning.

(Refer Slide Time: 06:09)



And the benefits that you get out of it is as I told you it is an absolute quantification, you get high precision and high sensitive. I explain you how and why this happens.

(Refer Slide Time: 06:21)

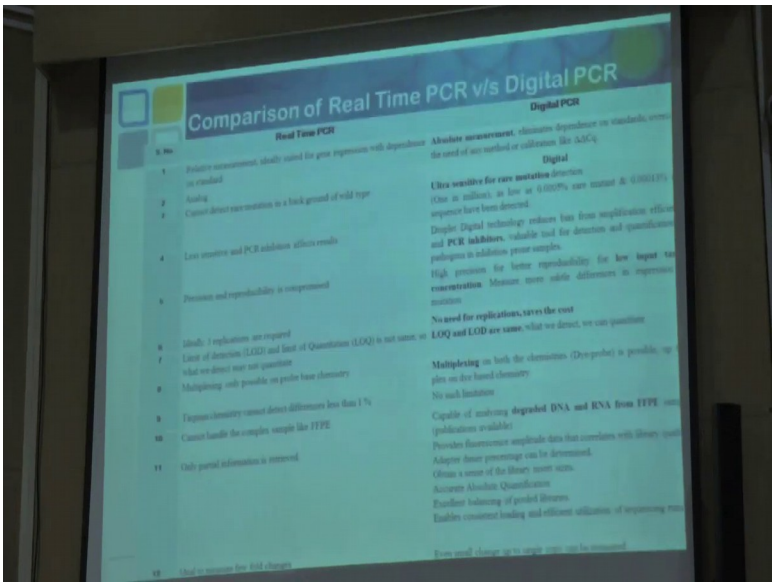


So, this figure is a very simplistic figure from one of our publications where people have used our technology. Now, imagine you are looking for a particular mutant, a mutation right which is 40 mutant molecules across your wild type molecules are 40,000 in a particular sample all

right. So, that accounts to 0.1 percent of a mutation which by a real time PCR or any other technology is practically impossible to detect, because you are searching for 0.1 percent of your desired product out of the 99.9 percent of your wild type.

What happens because of partitioning is, now when you split this sample into droplets right. So, each of these 40 mutant molecules would go into droplets, and then each of these droplets will undergo a PCR individually. So, now, this 0.1 percent chance of capturing your DNA becomes as high as 33 percent, because each of your droplets undergo individual PCR reaction and is scanned individually, the data is individual droplet wise. So, that is how the percentage and the precision increases for this technology.

(Refer Slide Time: 07:29)

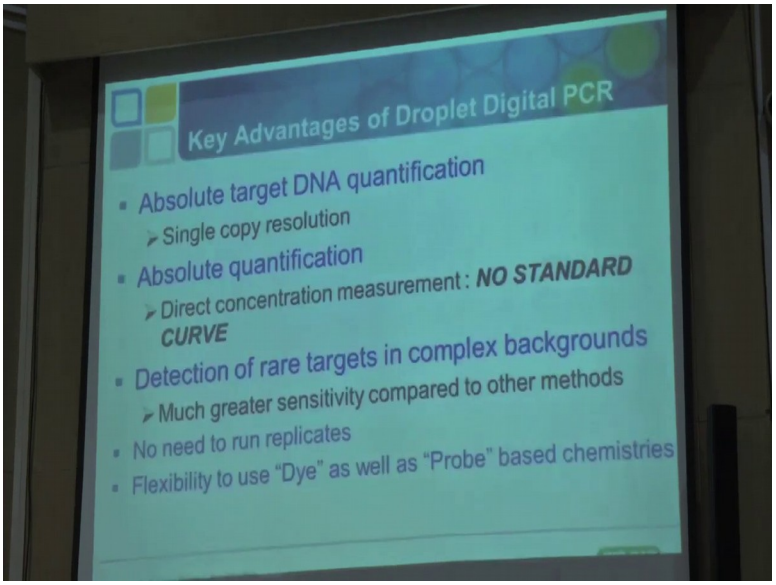


This is just a comparison that I have put up; I am sorry this slide is very busy as a real time PCR versus this technology. The first difference is yes, this is an absolute measurement, no relative quantification. This is ultra sensitive where you have to precisely pinpoint and detect the copies of your DNA or the mutation in your particular sample. Even if your sample a lot of inhibitors, the technology is good enough to capture your data which otherwise a normal traditional PCR would not.

Most important point here is LOD and LOQ, what this means is the limit of detection and the limit of quantification is exactly the same in this technology ok; otherwise in a PCR in a real time PCR they differ. What you can detect is a limit of detection, and what you can quantify

amongst that is a limit of quantification. So, this technology exactly counts everything that is amplified and gives you the precise data.

(Refer Slide Time: 08:25)



So, the major advantage is if I talk of this technology would be it is an absolute quantification method, you can go up to a single copy resolution. If you want to look at any of the DNA or RNA in a particular sample, you do not need to run any standard curve. So, it is an absolute quantification, you can detect your sample and in a complex background as well even if there are very high amounts of inhibitors in your sample or your DNA is not pure right. The system will precisely still amplify and give you the desired amount of your DNA. Can I even tell me why would this happen? It is a simple logic that is implied.

Why would this technology still help you even in a complex background that means, even if the sample has lot of inhibitors or something? Why would this technology still give you a precise data?

Student: By dividing or partitioning it

Perfect the main principle of the technology partitioning. So, when you partition your whole sample, even your inhibitors are getting partitioned. So, the inhibitory effect of your inhibitors is not there ok. Is everybody clear? Anybody has any queries on this? And last this technology is good enough to run with a dye based chemistry or a probe based chemistry. So,

I do not know if everybody is aware, there are two major chemistries by which you can only real time PCR.

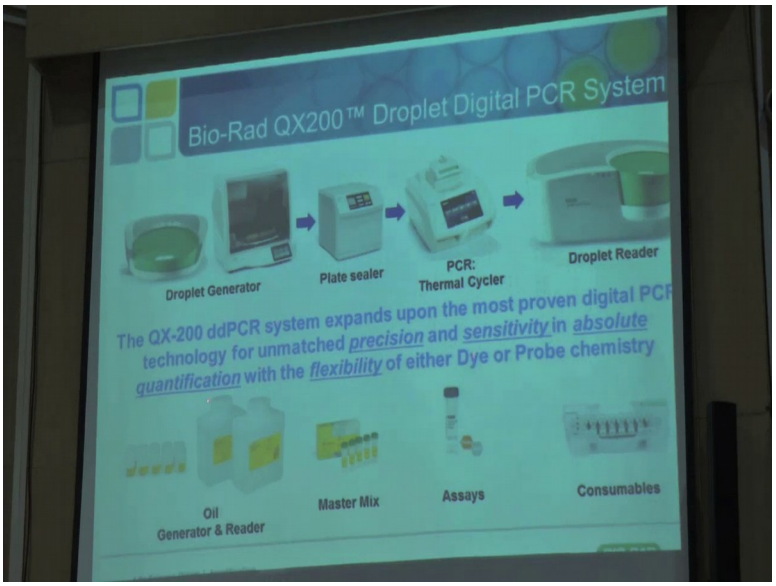
Student: SYBR green and TaqMan probe.

Correct SYBR green and TaqMan probe the example. So, SYBR green is which chemistry dye or probe?

Student: It is a dye based.

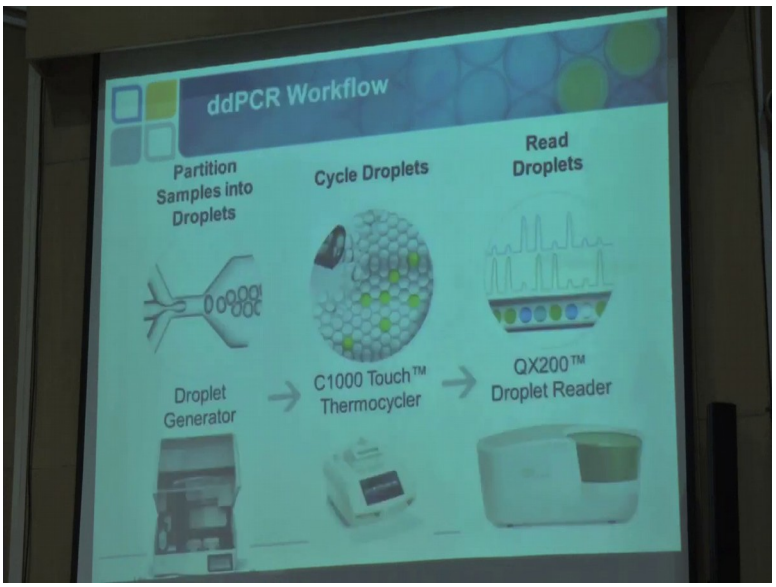
Correct, a dye based chemistry. And TaqMan is the probe based chemistry right. So, you can run your samples with either of the chemistries in the sample right. This is how the workflow goes. You first as I mentioned before generate the droplets we do a PCR and we read.

(Refer Slide Time: 10:05)



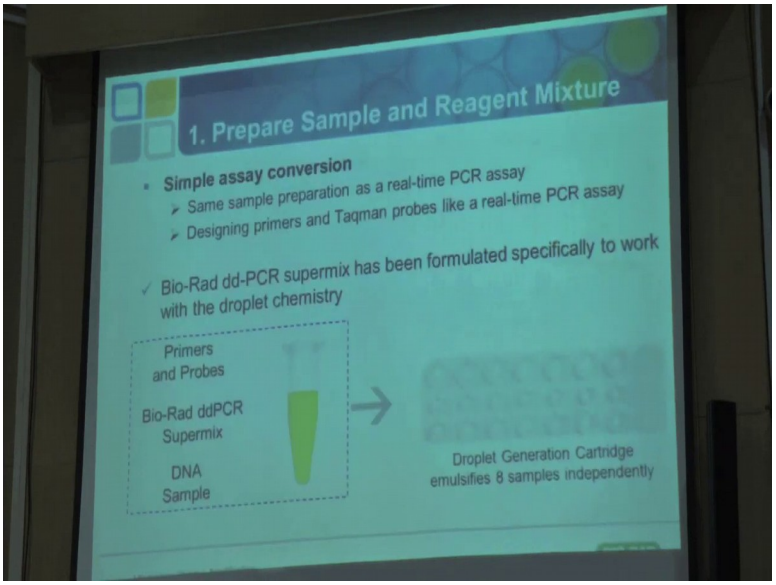
I will go through each step in details. This is just a workflow. So, in just if I have to explain this technology, it would be it expands upon the most proven digital PCR technology with unmatched precision and sensitivity in absolute quantification with the flexibility of using either a dye or a probe based chemistry. So, that is in a gist over the technology principally works on.

(Refer Slide Time: 10:27)



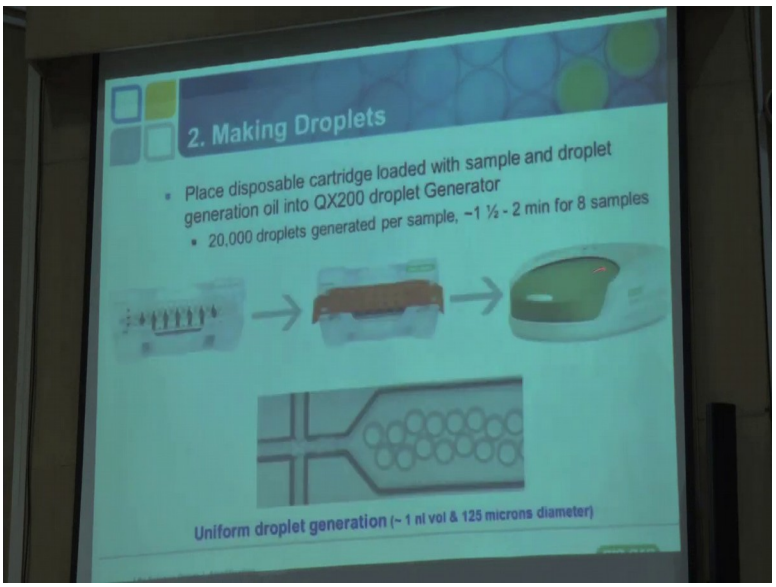
Now, this is a workflow we will look at the each step in detail.

(Refer Slide Time: 10:31)



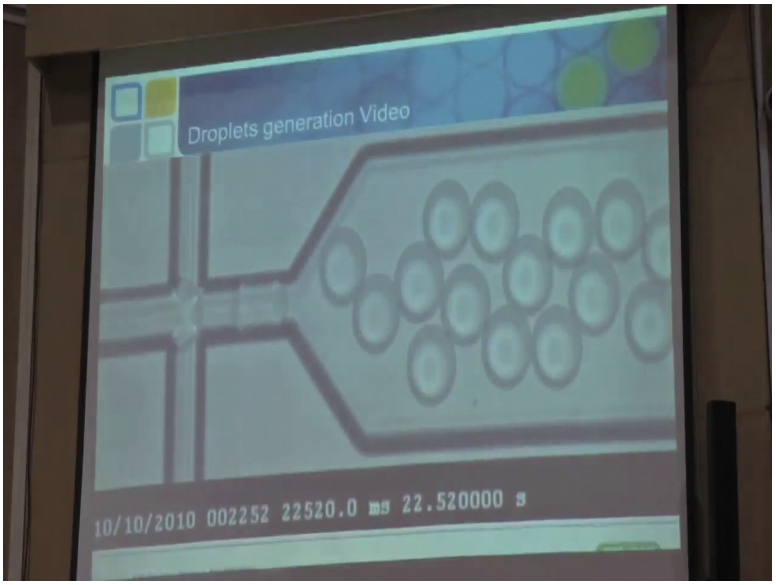
The first thing which I mentioned your partitioning your sample. You have a reaction mix which otherwise you make a same thing for your real time or normal PCR as well right. This is a cartridge which we have as you can see there are three rows right. In the middle row, you put your reaction mix that you are made with your DNA, primers, probes and your enzyme. And you add a oil here in the bottom most part. And you put this cartridge in a cut droplet generator. This is droplet generator.

(Refer Slide Time: 10:59)



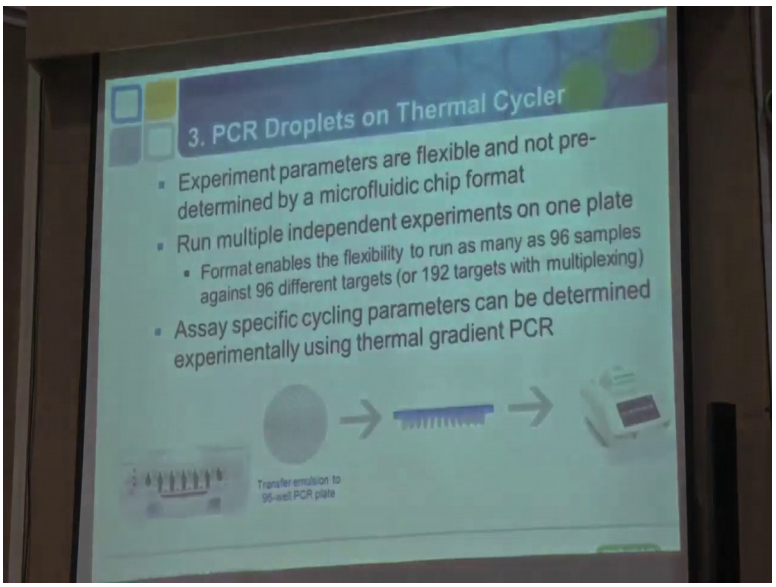
So, what the system does is this.

(Refer Slide Time: 11:01)



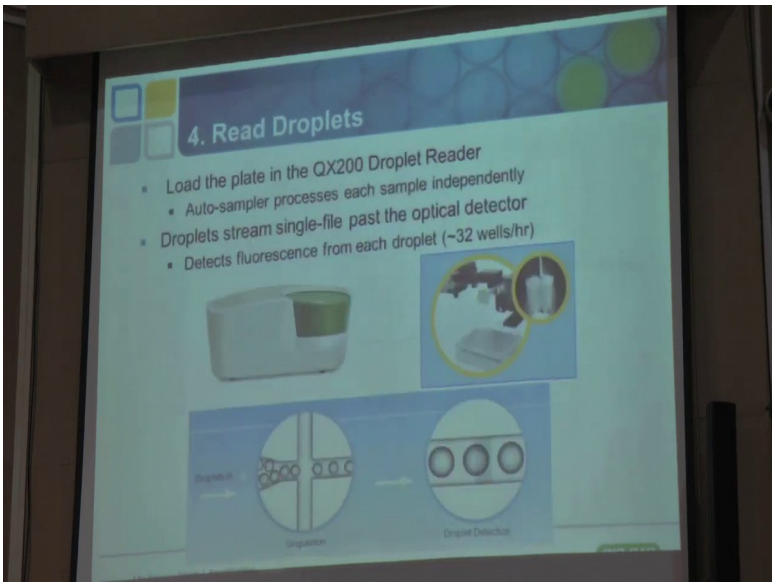
This is the actual video of how our droplets have being generated it combines your reaction mix along with the oil, and it generates these uniform size 1 nanometer volume droplets right.

(Refer Slide Time: 11:19)



And now these are the droplets which we put for PCR under a normal optimized condition which would have done for a normal PCR.

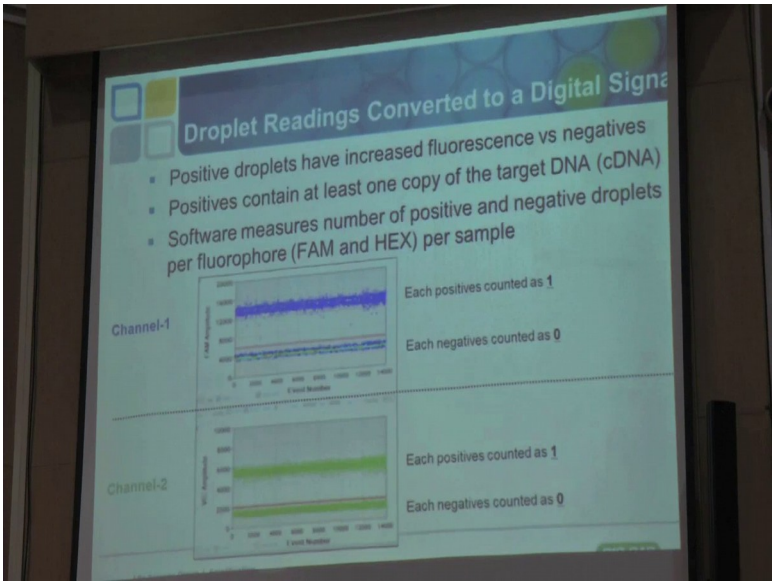
(Refer Slide Time: 11:29)



And then these are read. So, the reader here is like a flow cytometer ok. What the what it will do is this is the probe, the probe will enter each of your well, whether at 20,000 droplets each of the droplets will be pulled by vacuum right, they will pass in front of a two laser system. There are two lasers inside, and it will detect for a positive signal or a negative signal out of

each droplet. So, now there are 20,000 readings from one sample which are then combined together, and you get a data all right. So, this is how principally it works.

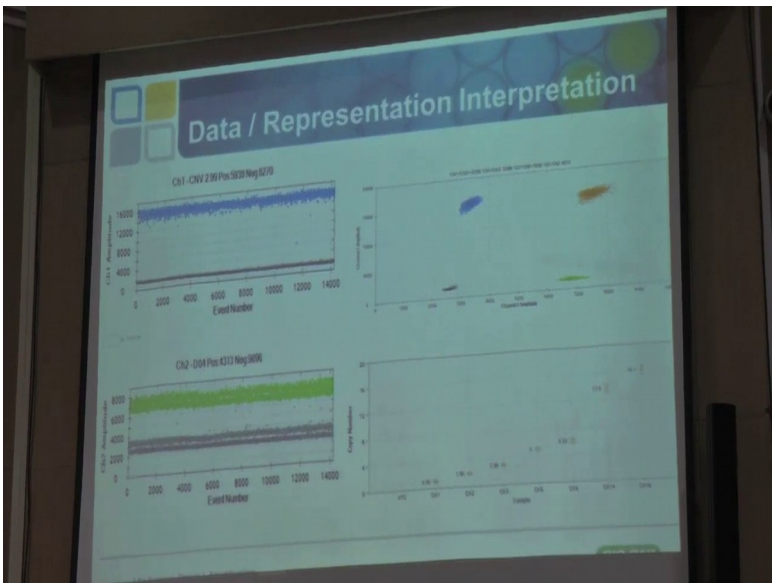
(Refer Slide Time: 12:01)



And the data that you get is something like this for those of you who have worked on a flow cytometer, you get a scatter plot right. So, this is similar scatter plot that you get. This is a data from one sample. So, each of this dot that you see here each dot corresponds to one droplet. So, there will be approximately 20,000 readings from one sample. The droplets which are amplifying something which has a DNA, there is a amplified will have some fluorescence emitted and they will fall here in this region. And the ones which do not have any amplification will fall here.

So, your y-axis is the fluorescence amplitude right. And the software then counts each of them individually and then gives you a data. And this is for your first channel. Similarly, if you are doing a multiplex as a if you are using two different fluorophores, you have a similar data for your second channel as well. So, these are also droplets which do not have any fluorescence which have no DNA which is amplified versus these are droplets which have your amplification of a DNA.

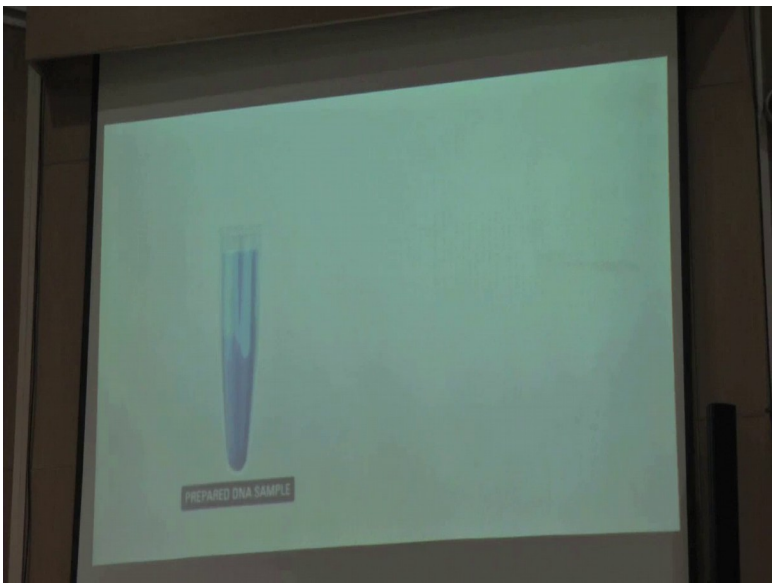
(Refer Slide Time: 13:03)



Similarly, we do get a 2D plot as well for this, where these are the first channel positive data, for the second channel positive data, this would be for the droplets which have both of your DNAs if you done a multiplex assay and these are your negative droplets. OK? Am I clear till now, how the technology works right?

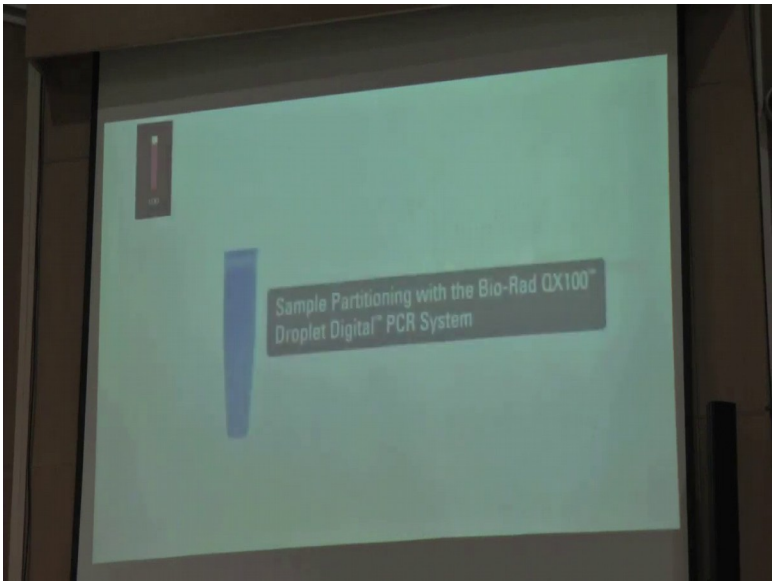
Next what I will show is I have a small video which explains everything that I have told till now you know graphical manner. So, in case there is anything that is left out you guys can understand it further better. I doubt if the audio system is working ok. I will just explain you along it goes.

(Refer Slide Time: 13:49)



This one right. [FL].

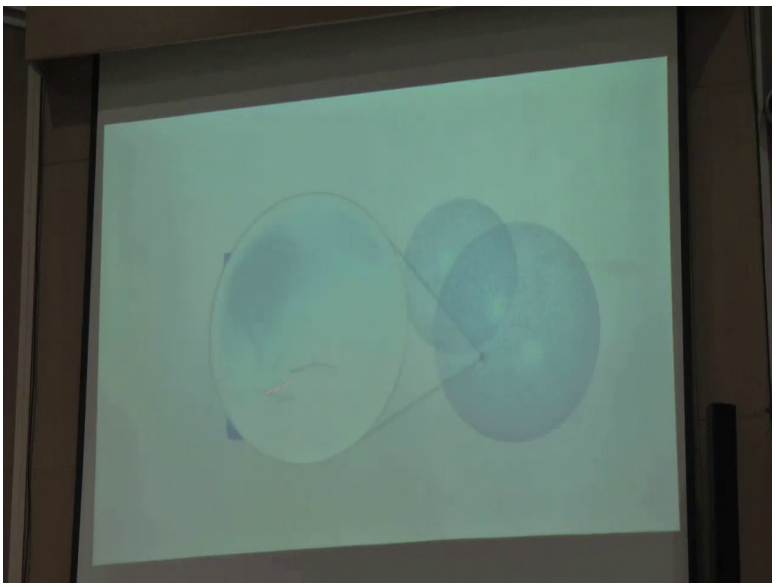
(Refer Slide Time: 14:13)



Volume full [FL].

So, I will explain you as it goes, there is no audio coming through right.

(Refer Slide Time: 14:19)



So, this is how your reaction is split up, these are droplets which will have your reaction mix along with your super mix, enzyme, DNA, RNA, primer probe whatever.

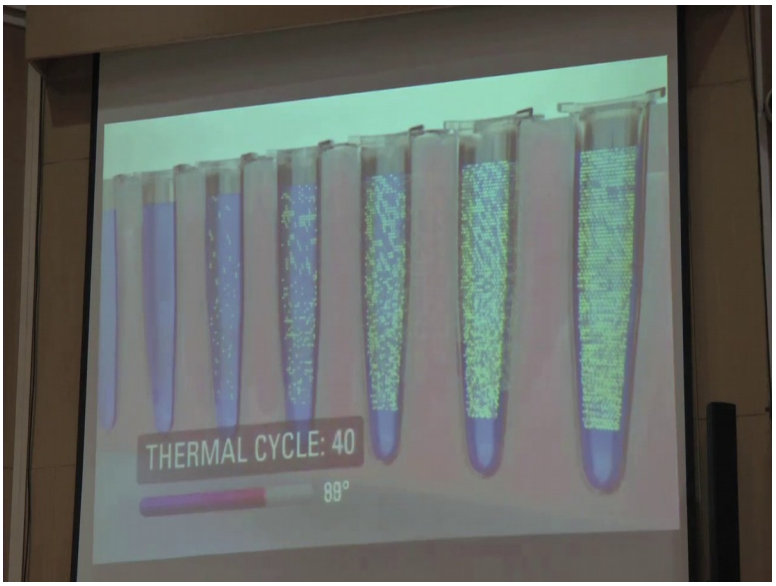
(Refer Slide Time: 14:25)

And as it undergoes a PCR the primers would mind there would be amplification which is happening. If there is an amplification, the fluorophore is emitted, and there is a fluorescence right.

(Refer Slide Time: 14:37)

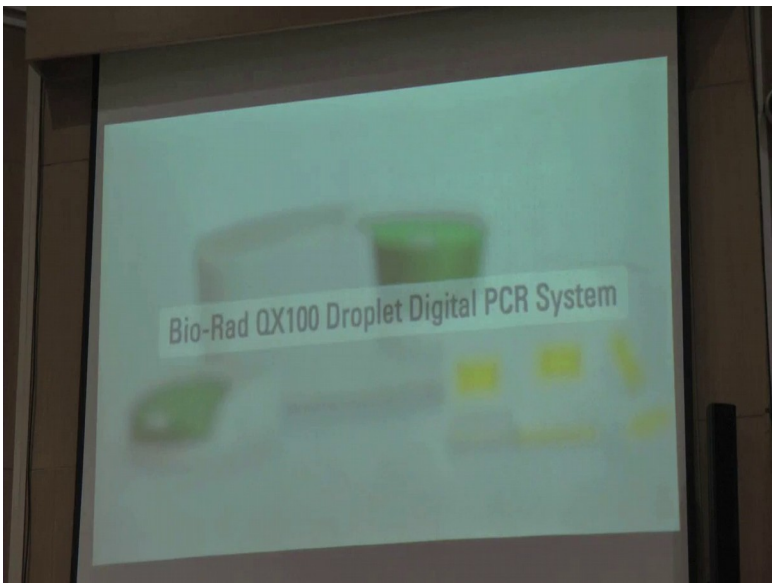
And the once we do not have any amplification would have no fluorescence. And the systems scores each of these droplets individually for a fluorescence versus non fluorescence.

(Refer Slide Time: 14:51)



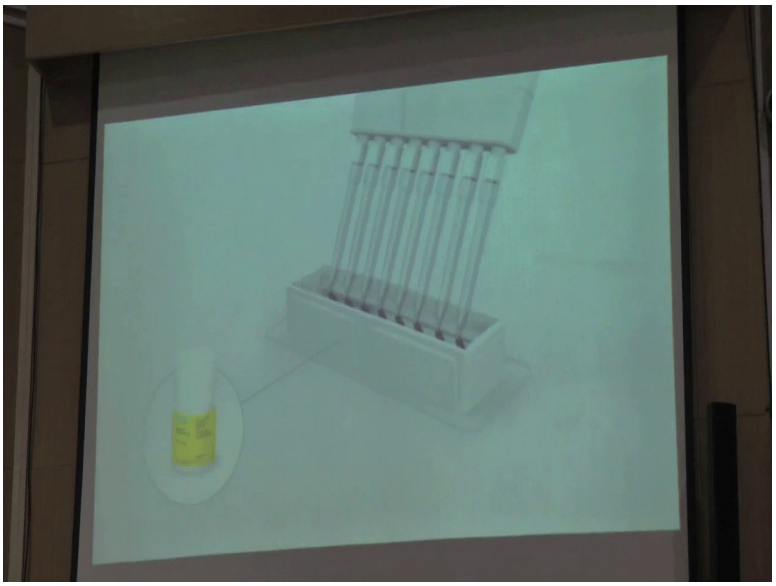
So, this is when it is undergoing your thermal cycling a PCR. Few of the droplets will have amplification the ones green will have your amplification, the others will not.

(Refer Slide Time: 15:11)



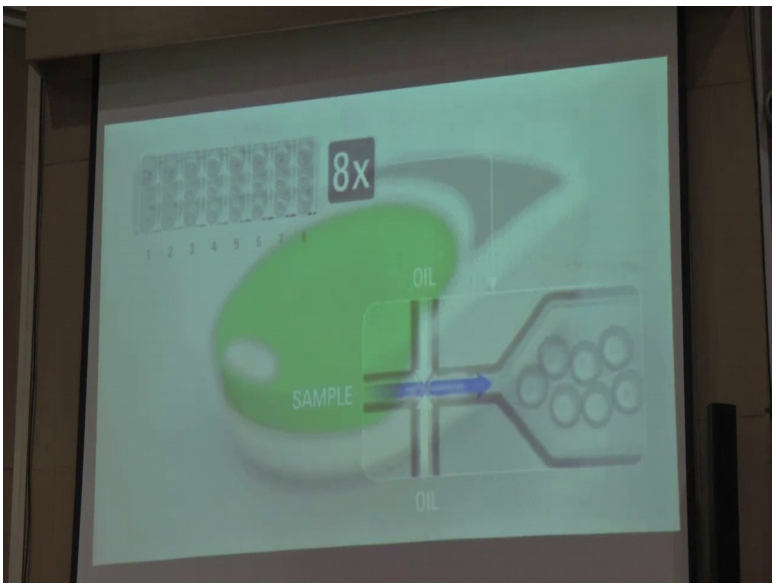
That is how the workflow goes what I was explaining, you have a reaction mix which is prepared.

(Refer Slide Time: 15:23)



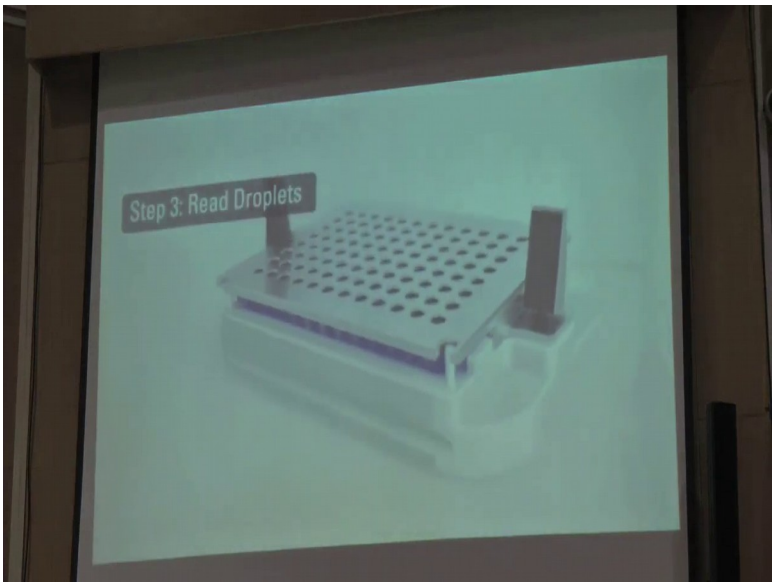
This is the cartridge you add your samples in the middle row. And the oil goes in the bottom row of each cartridge. You add your oil and you will put this cap you know droplet generator which will now make droplets from each of these wells. That is how this is the actual video of it makes your droplets.

(Refer Slide Time: 15:45)



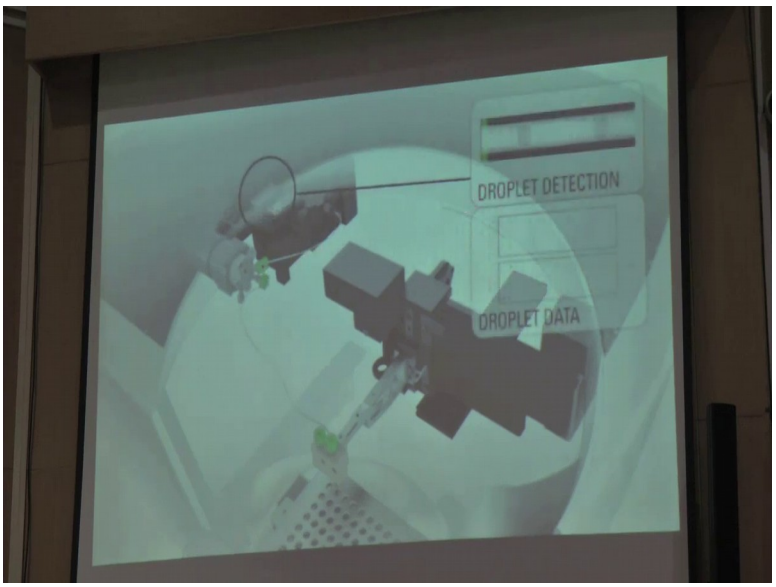
Once they are done with the normal pipette, you take these droplets, and then put it in a normal PCR plate, and put the plate for PCR.

(Refer Slide Time: 16:19)



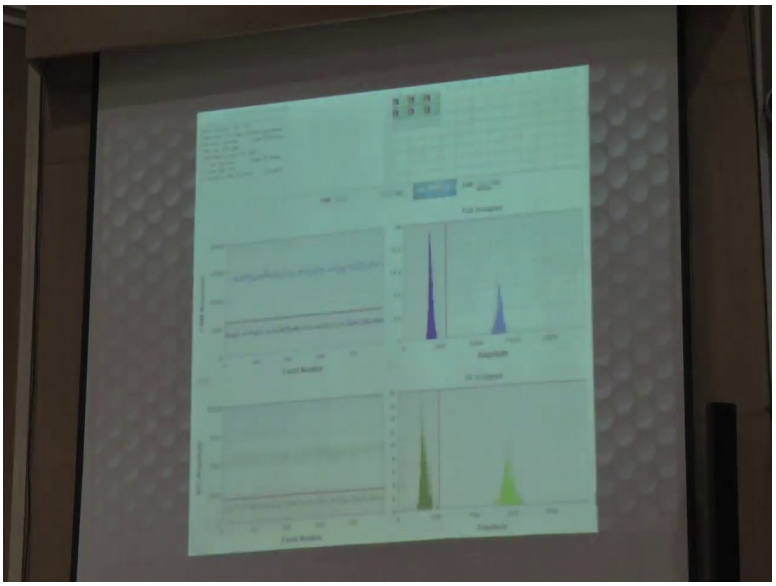
The reaction undergoes PCR and you remove that plate and put it in a reader the droplet reader. Now, each of this wells wherever you have added your droplets.

(Refer Slide Time: 16:27)



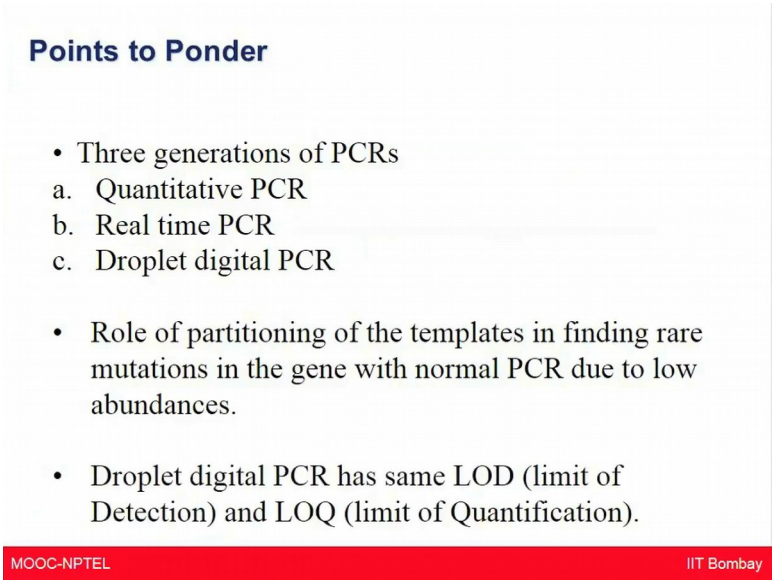
This is the probe which will enter each of these wells. This is how it happens, it takes in all your droplets right which goes here there is a laser system here, two-way laser system which is now going to detect your fluorescence whichever droplets have your amplification, we detect at positive; the ones we do not will be detect at negative, and that is how the data is being quantified.

(Refer Slide Time: 16:51)



So, that is a real time reading how it comes for each of your droplets as it pass through.

(Refer Slide Time: 16:59)



I hope today you have learned how droplet digital PCR or real PCR is 20,000 times more sensitive and provides you very specific data as compared to the real time PCR. Advancement and software provides very accurate concentration. Therefore, you need not to do DNA quantification separately. I hope you also learnt about LOD which is limit of detection, and LOQ which is limit of quantification of this technology. The next lecture will be the continuation of droplet digital PCR by Mr. Abijit.

Thank you.