Interactomics: Basics and Applications
Prof. Sanjeeva Srivastava
Dr. Sanjay Navani
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay

Lecture – 52
The Human Pathology Atlas: A Pathology Atlas of the Human Transcriptome-I

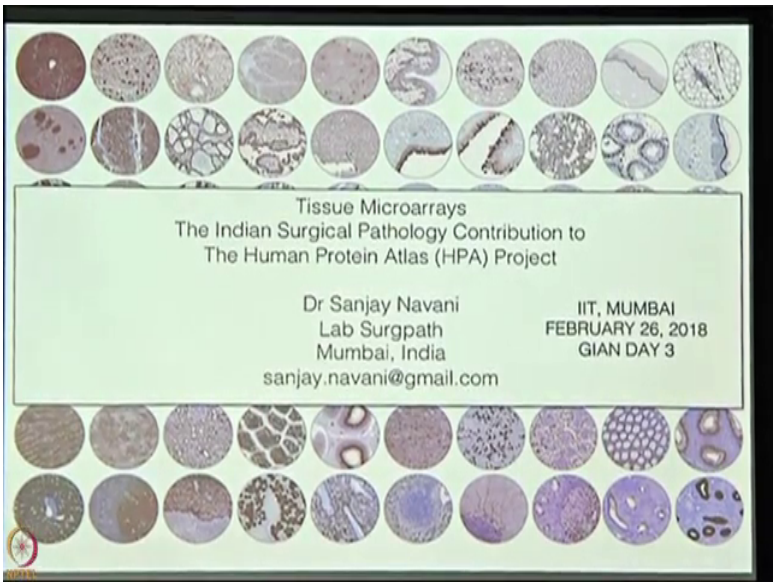
Today we have a distinguished scientist whose gone a talk to us about the latest advancement in the areas of Human Protein Atlas. This lecture will be delivered by Dr. Sanjay Navani. Dr. Navani is a consultant surgical pathologist an immunohistochemist who provides surgical pathology services through his laboratory Lab Surgpath. Lab Surgpath is currently the one of the laboratories in the world which offers immunohistochemistry stained slight services for more than 200 diagnostics and 20000 research laboratories.

Lab Surgpath pathologist have contributed immensely to the human protein atlas H P project by manually annotating approximately 16 millions immunohistory chemistry images over a period of 8 years. In today's lecture; Dr. Navani is going to provide more detail about his contribution on pathology atlas. Which just to give you some understanding that there is a quest to know all the proteins present in human both using mass spectrometry based approach and the protein array based platforms.

Dr. Byathes Woolhen and his group has really contributed immensely to the antibody based, array based platforms where they have tried to look for each protein a specific antibodies and then trying to see where they are localized. These are one of the major tasks mammoth effort conducted by his team where Dr. Navani and his team has also contributed immensely for the pathology project. So, I am sure were all very excited to have Dr. Navani's talk today and let me welcome him for today's lecture.

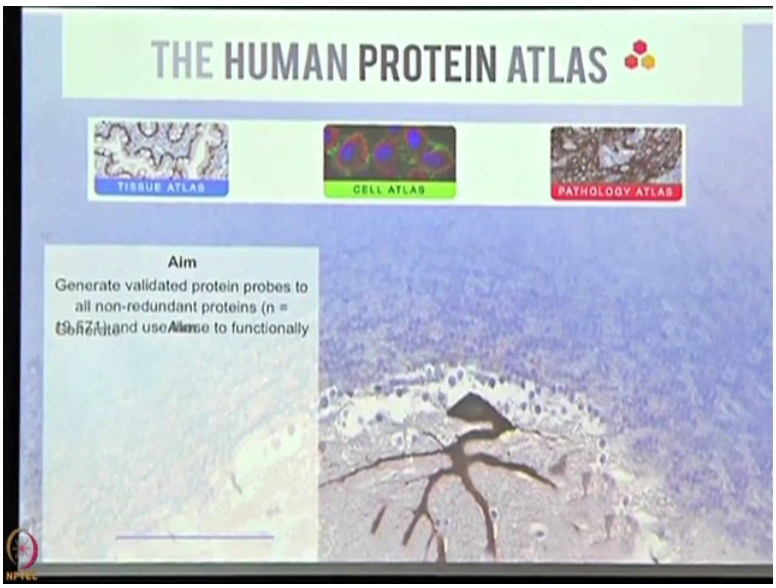
So, I am what is called as surgical pathologist. Formally trained as a physician with my MBBS and then MD in surgical pathology which means that I look at biopsies and I do a lot of immunohistochemistry.

(Refer Slide Time: 02:54)



We are also the pathologist for the protein atlas. As you know all the images that you see on the protein atlas website have been annotated in India. And I hope to tell you a little bit today about what not only what we have done, but what the protein atlas is actually currently doing.

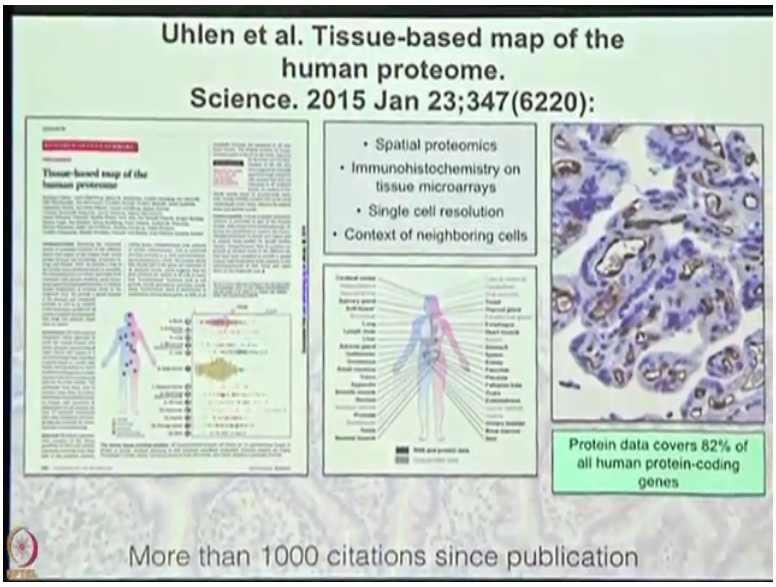
(Refer Slide Time: 03:14)



To give you a brief history about the protein atlas, it started off in 2003. The idea was to generate to generate antibodies and study their presence in normal and cancer tissues and to produce one antibody for every non redundant gene. The protein atlas is composed of 3 main parts. The tissue atlas which is normal tissues which we had a nice discussion about and I will like to expand some more on it. The cell atlas, which is now the centre of many things happening and the pathology atlas which consists of the use of antibodies on pathology tissue.

But mainly when I say pathology I mean that it is a cancer atlas. First publication came out in 2015 that when I say first publication of course, there were many publications before this.

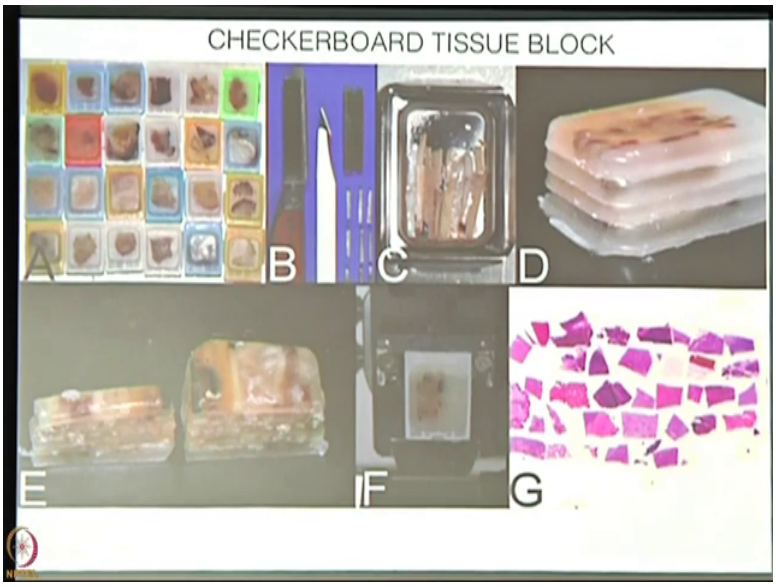
(Refer Slide Time: 04:25)



But that was when we get we put out the full draft of the first human proteome as seen in normal tissues. The main platform for the protein atlas is immunohistochemistry and tissue microarrays and all of these images which are evaluated are available on the website for any researcher student to see.

To give you a bit of a background about surgical pathology and what it is and how we interacted with this project I thought I will just explain what tissue microarrays meant.

(Refer Slide Time: 05:08)



And I put this slide in just to show you how tissue microarrays actually originated. With the advent of immunohistochemistry more reagents became available and these reagents could tell us things about the tissues that we did not normally see under a light microscope. If we applied them in sequence.

So, they told us that this looks like this, but it is actually got an oestrogen receptor and that was valuable information because, we knew where that protein was being expressed at that particular time. The problem with immunohistochemistry amongst other things was also that it was and continues to be a fairly expensive investigation. And therefore, the expenses directly related to the amount of reagent that you use on a slide.

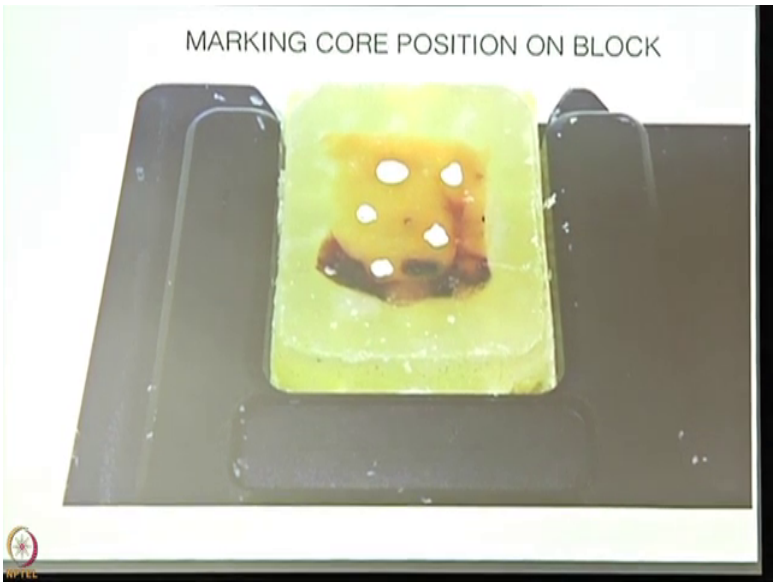
So, if your population for study is 5000, your entire funding could go down only in staining your slides. And therefore, you see now the need why a different technique was required and why pathologist decided to start to try and cram all tissues on to one slide.

Put all the tissues or one slide, put everything on one slide, use the same amount of reagent, save your money get the answers. Now the origin of that idea which finally, today is called tissue microarrays and forms the main platform of the HPA program started with something like this. This was what was called a sausage block. So, people took out interesting bits of tissue that they wanted to study from many blocks from different patients and they put it all into one block and then cut a slide that look like this.

So, one piece of this was lung tissue, one piece of it was colorectal cancer, one piece of it was whatever you like. And with all these tissues on the same slide we put just one antibody and it stains everything and our cost goes down 80 to 90 percent. So, that was the idea. The sausage block came about because people wanted controls for immunohistochemistry as you need with all tests. You need to know whether it is working properly or whether it is not working and therefore, if you have multiple tissues on a block.

Some of which you know are positive and some of which you know are negative and then you have some in there which are also your tests it is looking good just to give you an idea how the tissue microarrays are made.

(Refer Slide Time: 08:28)



This is a usual tissue block, this is the kind of block that I usually use to give out a diagnosis. So, we have cut out a piece from a tumour. We have taken a slice of it have looked at it I have said this is cancer and the reports gone and for diagnostics that is the only way that it must be done. Tissue microarray is a research tool it is not a diagnostic tool.

So, pathologist cannot do that on every case that comes into the lab. But once this has been established as breast cancer for example, we can use to take cores out of this tissue and put it into another block where there will be tissues from other blocks and finally, we will create a tissue microarray block. I hope I am quite clear so far. Now the way that this is done is a slice of this full piece of tissue is cut and viewed under the microscope. The pathologist is asked to mark characteristic diagnostic areas in that slide which contains the whole section.

So, he or she does that, that slide is passed on to the technical person and he is instructed to take cores from those places that the pathologist has identified as classic. Because you must remember that when you see a tissue block of this size, there is tumour, there is stroma, there is normal tissue. So, somebody needs to tell the technician where to put the core.

Student: Is that tissues stabilised to be some sort of pigment or it will the (Refer Time: 10:29).

So, that is a standard process followed in surgical pathology as soon as the samples are taken out of the human body. They the specimens are in sized cut into small thin slices usually which do not exceed 2 to 3 millimetres in thickness and not more than one and a half to 2 centimetres in maximum length and they are placed in a fixative call formalin. Once it is fixed in formalin, the tissue is so to say preserved. We then need to get it ready so that, sections can be cut and viewed under the microscope. A process that is referred to as tissue processing.

And after the tissue is processed, the tissues put into this thing which is actually made of paraffin. So, when you hear the term in the papers you read or the people you talk to and people say FFPE that is Formalin Fixed.

Student: Paraffin (Refer Time: 11:41).

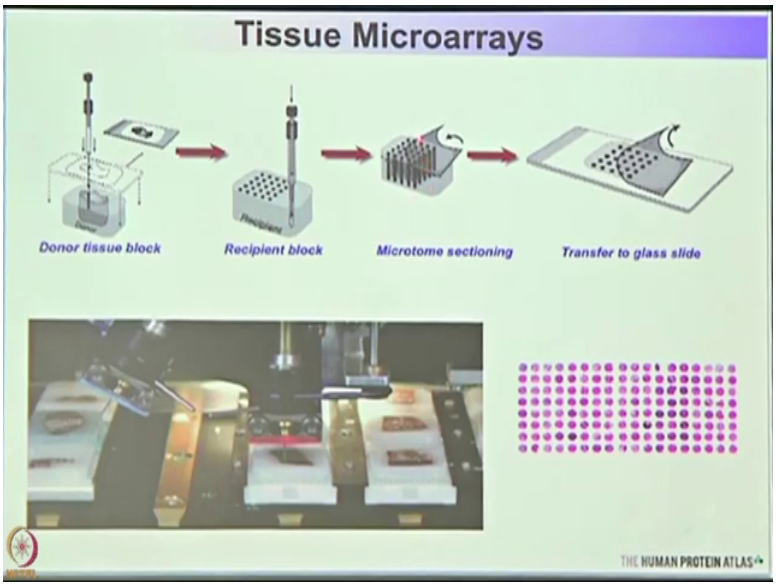
Paraffin Embedded. Amazingly immunohistochemistry has been able to do that. So, when formalin fixes the tissue; it denatures the protein or if I can quote an example that I have used if this is your antigen after formalin fixation it becomes this. If there is no formal and fixation is this. So, if it is this, there are epitopes regions of the protein specific enough that can be recognized to give you a clear staining pattern. But the point that you raised is very relevant because it brings into question the question of the antibody.

So, the antibody has to be good has to be validated we had a nice lecture from Dr. Joshuva about validation. You can produce something from a rabbit, but is does it actually work, that

is the big question. So, we have about a 150 to 200 antibodies that we use in the clinic. So, the antibody that I use for oestrogen and progesterone receptors on the basis of which a woman is going to get some drugs for breast cancer, they have to be very good. But most of the antibodies that I am talking about today are research antibodies whose aim it is to get validated. So, just that you know in the spectrum where the antibodies lie.

So, they are mature antibodies, they are new antibodies, their antibodies with partial data on them, how good is your antibody? That is the big question ok. Let me move a little quickly otherwise we will never get to the end of the lecture.

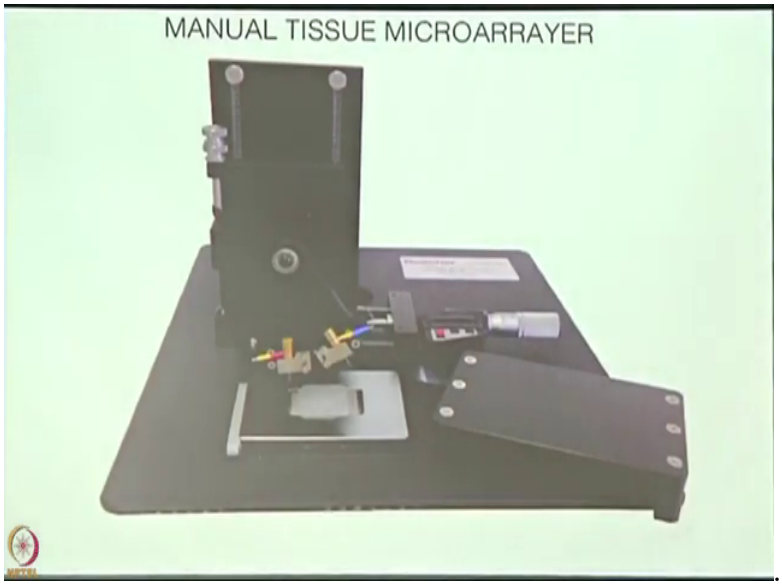
(Refer Slide Time: 13:54)



This just shows you briefly how tissue microarrays are made. So, you take a core from that block that I showed you and you put it into a different block in which there is space for more

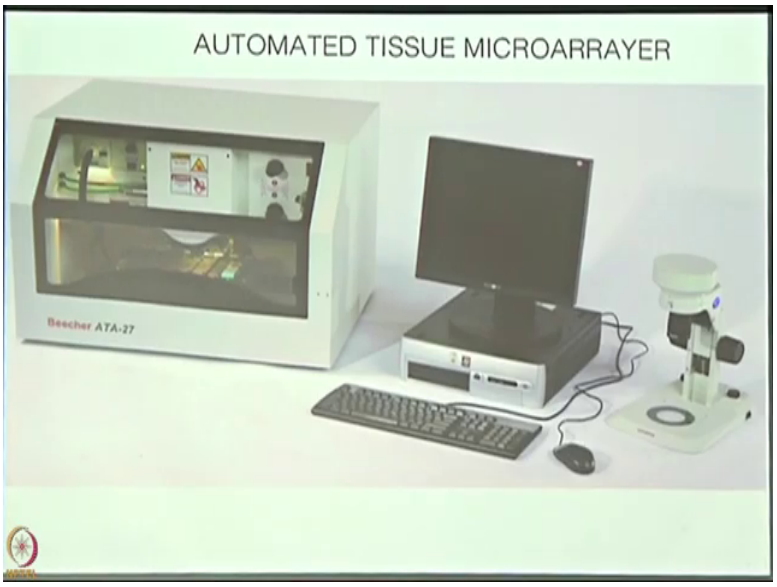
cores. This is then cut as a section put on a slide and then when you stain it this is an ordinary agent is stain that is how it looks.

(Refer Slide Time: 14:20)



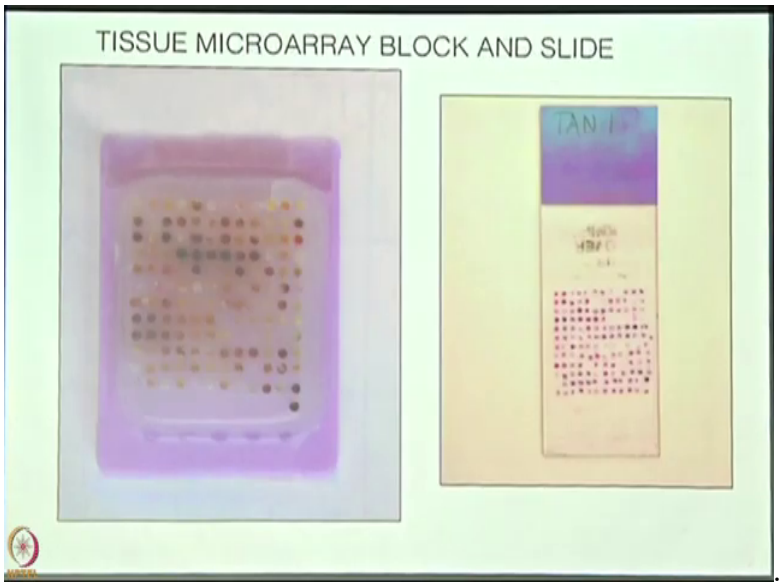
That is the machine that does it is called Tissue Microarrayer and they are many types of those that is a manual tissue microarrayers; the one that we use in our lab it is fairly easy to use good device.

(Refer Slide Time: 14:30)



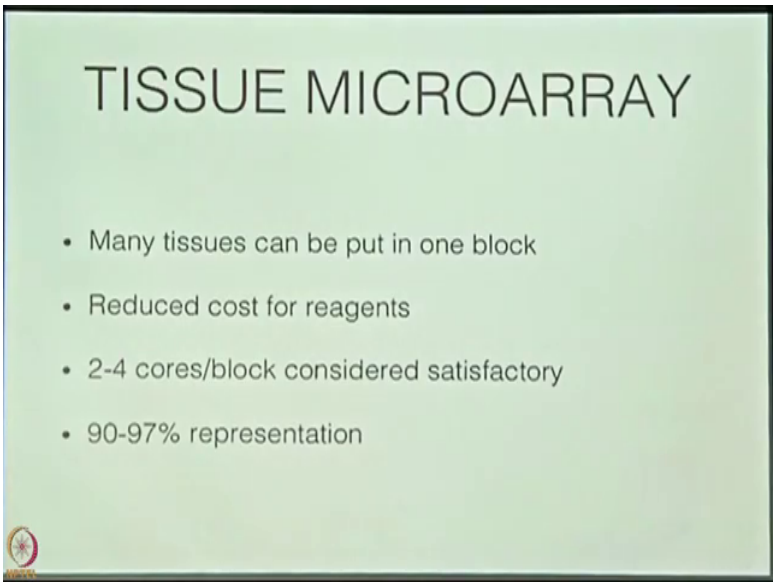
But Swedish colleagues use more sophisticated tissue microarrayers that are automated. Automated microarrayers can make up to 20, 25, 30 microarray blocks a day. But it is an instrumentation which has it is own problems. So, if you do not have that workload it is nothing to that in to get into you are much better off with a manual microarrayer.

(Refer Slide Time: 14:59)



And that is how or the block finally look, nice neat equal sized cores and when you cut the section that is the slide ok. Now this is already been stained with an (Refer Time: 15:13) which is a regular stain, but any kind of stain can be done it any immunostain can be done on it.

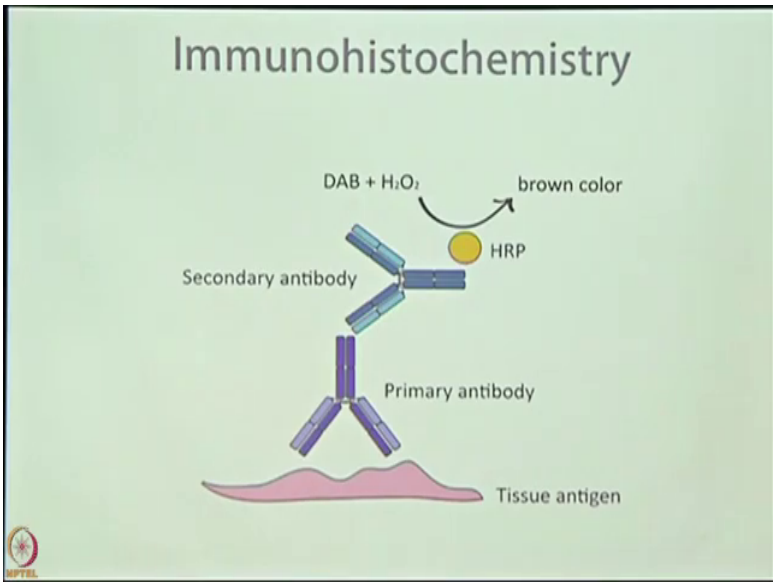
(Refer Slide Time: 15:20)



So, the advantages which I have already outlined many tissues can be put in a block as a reduced cost for agents. How many core should you take to obtain a correct representation. 2 to 4 cores is normally ok. Most people internationally use 2 cores for a cancer. The protein atlas uses 2 cores for a cancer.

But for normal tissues, we use 3. And if you do that then we get a 90 to 97 represent percent representation of the original block. Provided the correct areas have been marked by the pathologist. So, let me tell you a little bit about immunohistochemistry. You have this antigen which is in the tissue.

(Refer Slide Time: 16:13)

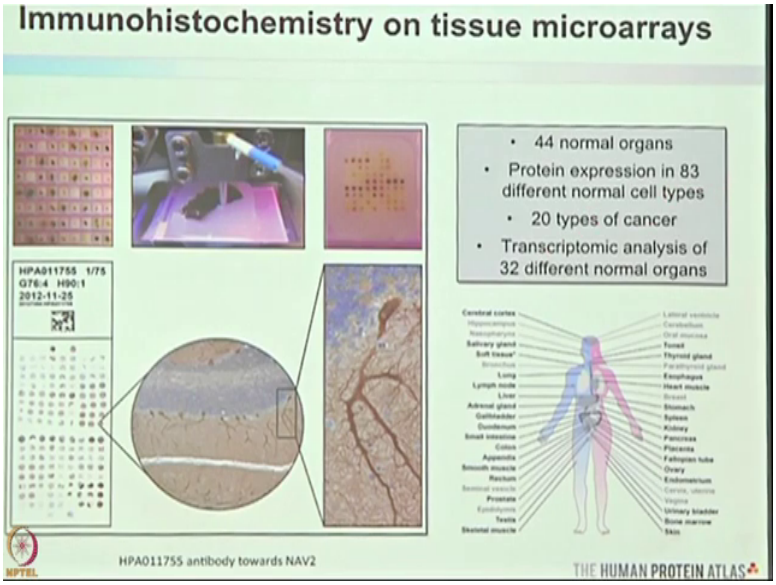


Which we cannot see normally, you apply a primary antibody which goes and binds to the tissue, then you apply a secondary antibody which detects actually the primary antibody. And finally, you add a colouring agent which usually gives you a brown colour. And when you see that you infer that the antigen is actually present.

So, everything depends on your skyscraper of I immunohistochemistry standing up correctly. If any of those products are faulty including the tissue antigen which may not have been preserved properly. If it is not properly fixed in formalin you will get a false negative. If you do not follow the other procedures use very highly sensitive reagents you will get a false positive. So, it is the only thing which we have in which you can see the tissue morphology and also see where that protein is localized.

So, now you understand what is the importance of using tissues to test your antibodies because then you know where those reagents where those proteins actually are. So, that is a brief view of exactly what happens a lot of blocks go and make a TMA, you cut a TMA, you stain it with immunohistochemistry, you look at that stain and that is what it looks like.

(Refer Slide Time: 17:50)

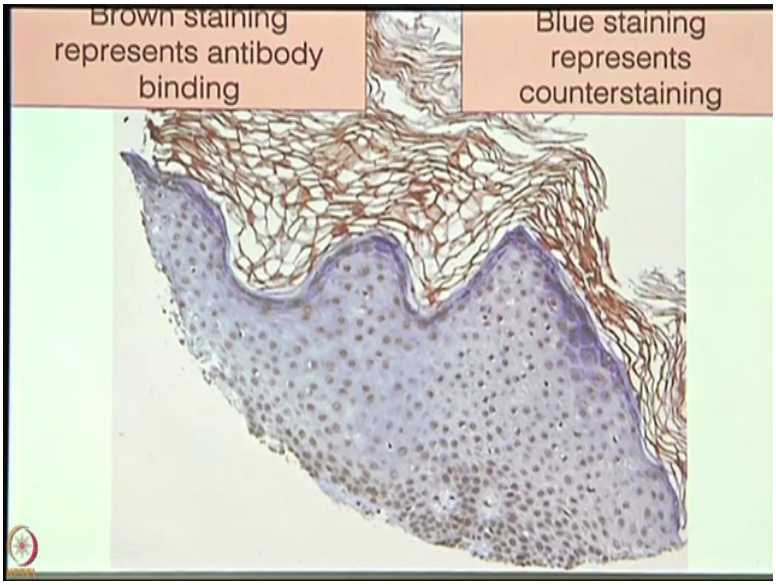


The protein atlas covers 44 normal organs. It studies the protein expression in 83 different normal cell types, 20 types of cancer and it has transcriptomic analysis of 32 different normal organs.

And I was very pertinent question raised about what is considered normal. Normal means in this case not cancer. Most of these tissues have been obtains from Swedish subjects who either underwent an autopsy or who underwent surgery for a non cancerous condition. That is

why in some of these tissues there is a mild inflammation present. But it as close to normal as we can get and that is how it appears finally, that is something what I see.

(Refer Slide Time: 18:52)



So, you see this brown stuff the brown stuff in the nuclear and the blue stuff. The brown stuff is the stain; that means, the antibody is positive. The blue is the counter stain and it is put only because it helps us to appreciate the brown better. If you are not careful with the blue and you will leave the blue on for a little bit longer, it will mask the brown.

So, immunohistochemistry all though if you speak to diagnostic pathology or if this was a diagnostic pathology platform you would hear me (Refer Time: 19:32) about how we can do great things with that. But trust me when I tell you it is on of the most fragile platforms that you can ever encounter. Today you will get positive, tomorrow will get it negative. Your quality control reagents storage it is a got to be up there.

Student: Sir related to that then are their internal controls that you need to have present at the tell you that the staining is just right is that (Refer Time: 20:11).

So, that is.

Student: (Refer Time: 20:13).

A valid point and usually many tissues that we study at least on the diagnostic platform. Will have an internal control and one internal control per batch at least and if we move to external, then at least one external control on the same slide in the batch. So, that is a very valid point Dr. Joshuva is raised. How do you know that everything is working. So, to give you the most common example of breast cancer in which we take a fragment of breast cancer tissue. The periphery still contains normal breast tissue.

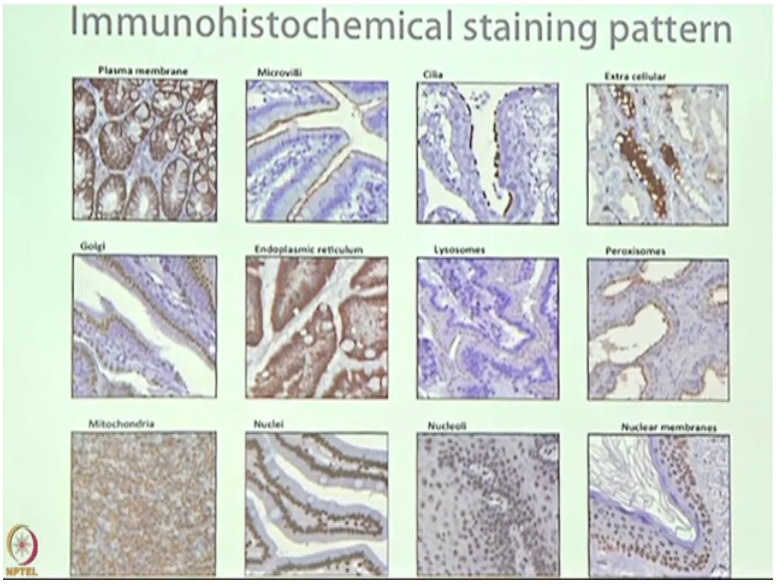
So, when you put the stain there and you say it is negative you look up there and it should be shining. So, you know it is working. Now in this situation you may well ask the other question yeah, but what if it is a false positive. It may not be it is likely not to be because you should have then seen that in the tissue is well not only in that place. The reason I am going into this into a bit of detail is I really want to get across the point about how fragile a platform this is.

So, whenever you are getting any data on immunohistochemistry please read it and thoroughly and look to poke holes in it and please do not accept it at face value. It is most important right. So, the idea after carrying out immunohistochemistry on these tissues with all the antibodies is to print to find out the relative distribution and the abundance.

(Refer Slide Time: 22:00)

Where in the normal tissues is a present and how much of it is there? Because and you must remember that at the start of this exercise approximately 40 to 45 percent of the proteins were unknown. We knew only what 50 to 60 percent of the proteins do where they are even today we do not know 100 percent, but certainly we have covered a lot of ground since then.

(Refer Slide Time: 22:37)



Student: What about the (Refer Time: 22:38) point and it is section was (Refer Time: 22:41).

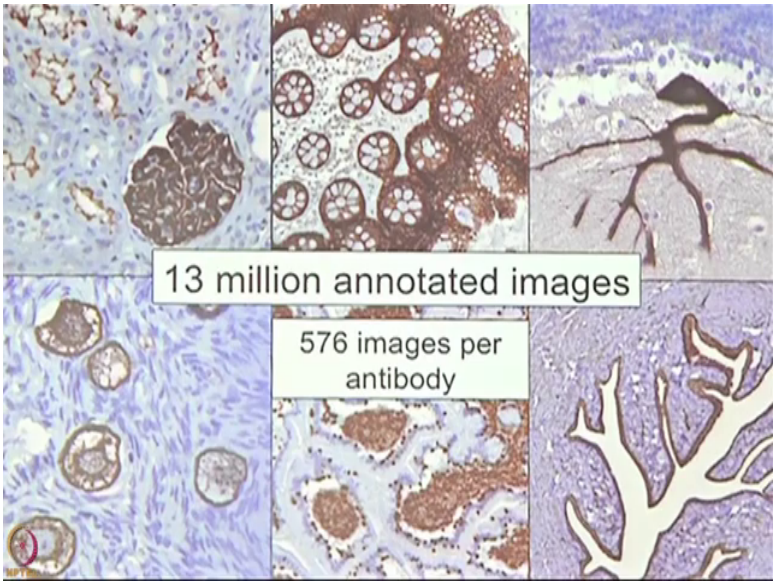
How would that affect? Once the tissue is formalin fixed, the moment that issue comes out from the body, it loses it is normal, nutritional support in the form of blood vessels and there are several studies which have documented that best results are obtained if that issue is obtained within an hour of surgery and placed in a fixative. There are many ways to look at proteins. We are looking at immunohistochemistry in tissues because it gives us the localization.

So, if what she says is correct. It is giving me the localization, but it is not actually correct. How would I check that? You use a different platform and look whether you are getting the same protein in that platform also. I will just come to that in a minute. So, with the antibodies

that were produced as you can see there were some lovely images that we got to see and we were very happy about it.

We were the first people in the world who saw these images came to know where those proteins are located a lot of them did not have any localization patterns predicted for them. And we do not normally see all these proteins in diagnostic pathology. So, it was in many ways a quite a visual treat for us to see and know where proteins were being localized.

(Refer Slide Time: 24:18)

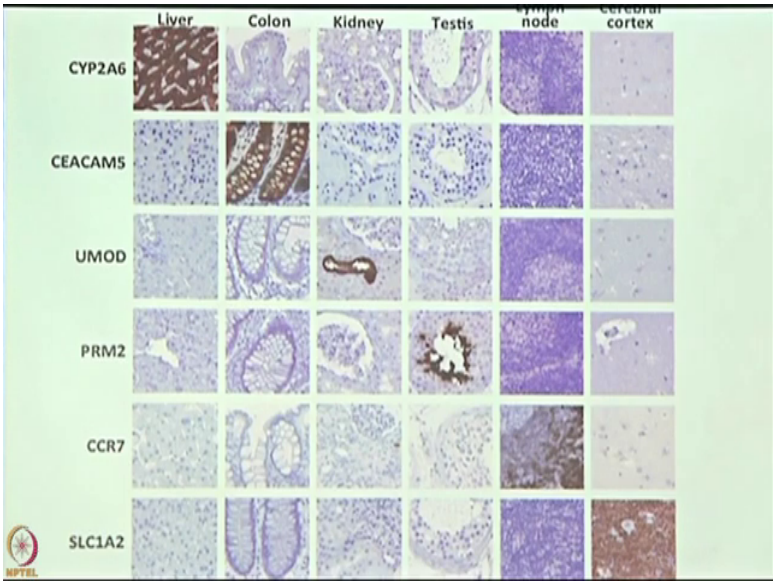


So, that is the job that Indian pathologist did. We looked at 13 million images over a period of 7 to 7 and a half years. It is the largest effort in human history and every image on the atlas has been annotated by an Indian surgical pathologist. There are about 20000 genes as you well know and the aim of the protein atlas project was to produce one antibody for each one

of them. So, if you produce one antibody, we wanted to see the global distribution means; all normal tissues 20 types of cancer.

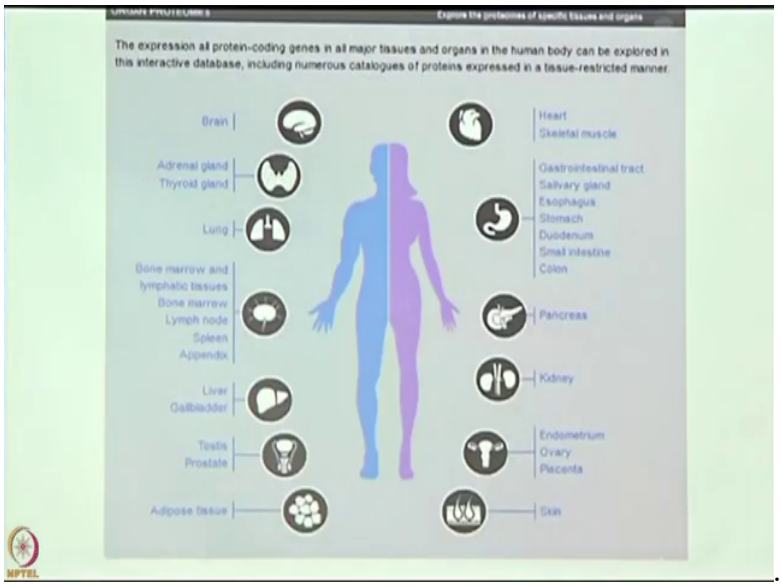
So, that gives rise to very large number of images. So, there were 576 images for each antibody. And that is a nice picture that is put together by one of my Swedish colleagues which shows you the organs on the top the proteins along this side and how each one of them is showing a specific pattern in a different organ.

(Refer Slide Time: 25:20)



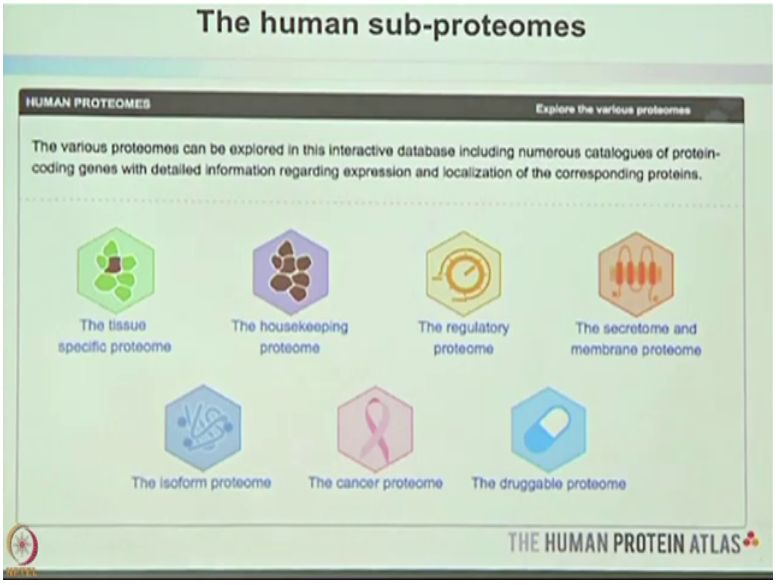
I must point out to you that this is a nice picture to show in a lecture, but I do not think that you should be going out expecting that is how it is clear is going to be. This is just to get the point across that there are some proteins expressed in some tissues that are vastly different from others.

(Refer Slide Time: 26:01)



If you go to the protein atlas, you will find this caricature and if you click any of these links and you can do that you do not need to register it is it is a freely accessible site, anybody can use it. And if you access anything say you a click pancreas it will immediately take you into the genes that we found in the pancreas and you can take it further from there. You can actually also see the images that we annotated we have we get several email from people all over the world, some of them complementing, some of them not so happy. After looking at the images and giving us their feedback.

(Refer Slide Time: 26:51)



We basically classified the normal tissues into 7 different types of proteomes. And this is also available on the website. Everything that I say is available on the website and in far greater detail. So, there was the tissue specific proteome that is specific proteins in specific tissues, there was also the housekeeping proteome which I will tell you a little bit further down the line about how important that is.

The cancer proteome which I am going to just start talking about. And very importantly also the druggable proteome. Most of the targets in the pharma industry are proteins. We manipulate them and then we get the desired result with some side effects. In the pharmaceutical industry; particularly the bio discovery area nobody speaks of using a drug by itself. They always talk about benefit risk ratio because, everybody knows that a drug cannot be only good.

Therefore, how should one look at these targets and is there something that we can do that will make these drugs do their job and at the same time not give a side effects.

Student: Sir I have one question like the tissue specific proteome ok. So, there these are proteins which are expresses as specific proteins, but many a times I have seen that they are dependent upon satisfaction also.

It depends on many things going into the tissue specific proteome, every tissue proteome I do not think this is in the scope of this lecture, but I would very much encourage you to visit the website.

Student: Sir I wanted to ask one more question (Refer Time: 28:54) clinical samples they will give you the lock masses of the tissue.

Yeah.

Student: Ok.

Yeah.

Student: The same kind of disease there can be different different tissues so.

Yeah.

Student: After fixation how you counter or how you reduce the biasness.

That is the job of the surgical pathologist. The surgical pathologist can look at it and tell you where to classify it, that is what a person like me does. Each tissue looks different. So, if you do not tell me what the tissue is in you slip it under my microscope I can tell you what it is and I am glad you asked that question at this time because you might be wondering whether

you are in an abstract art class or in the middle of a science lecture. But I am going to try and make a point about surgical pathology.

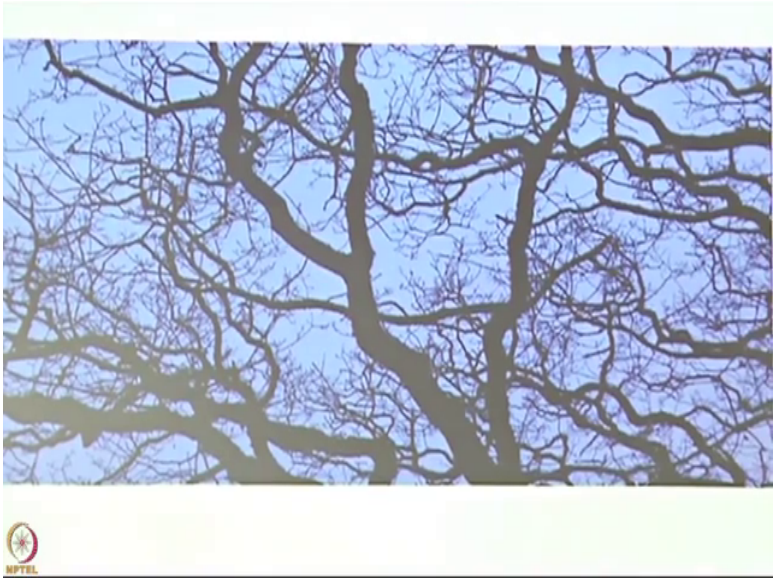
So, when surgical pathologist first see an agent is stained slide under the microscope. The first thing they see is something like this there is a lot of things happening some big, some small, some tiny. There is no central focus it is one picture that you have to take in fully and then you begin to start thinking what is important.

If you look at it for a little bit of time it crystallizes a bit and you begin to think yeah that actually I think this looks like branches of a tree, but it is no there are no leaves in it. And I think it is branches of a tree, but I am not sure and then we look at it some more and then we seek now I am sure.

A tree a definite, but there are no branches on the tree. So, I would like to equate this image with what a pathologist sees on a normal H and E section. When he sees an H and E section and he is come to this point this is short. This is a tree, they are no leaves and this is it. When I say I am sure I mean as much as can be possible maybe some of you did not notice Joshuva laughing about it because he is probably encountered the problem people see the tree and they say this is not a tree.

But I am just trying to build a grade to get my whole concept in. Now that is all that we see if it is so clear 10 out of 10 pathologist should tell you that is a tree and that is a big job in itself to get those 10 jokers to agree is a big thing. So, they will tell you now this is a tree. Now starts and that is something we have been doing for more than a 100 years.

(Refer Slide Time: 31:54)



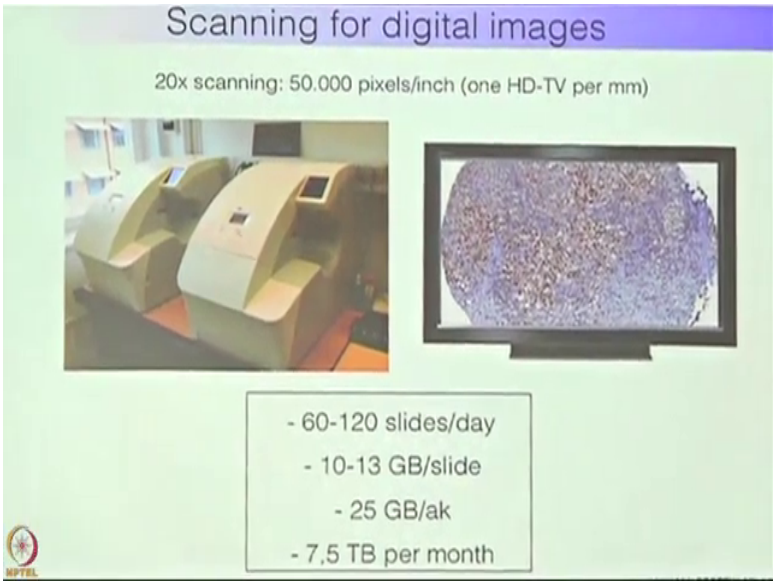
Now, starts what more can you tell me about it. Because, if you go only on this, then all 10 trees should die without water. But they do not only 3 die, 7 do very well. So, if you have 10 invasive carcinomas of the breast which has staged alike which looked the same then, why she dying and she is having a party. There comes in the question of is there something else. So, the first immunos might help you and they are exemplified by the leaves.

So, that looks different that is not that tree, that is going to die this is not going to die, that is what we are trying to do. And I am using breast cancer only as an example, but this is very far from the ultimate holy grail I should say which is we want to see it like this.

There is a tree, there are leaves, they have different colours, this is yellow, this is orange, this is red, this means this, that means that. Well you are in the middle of it right. Now that is what the whole effort is about which protein goes where what is it is role is it going to fall

off, is it going to continue, how important is it. So, now, you get the picture I hope from the surgical pathology point of view right. Let me go on to the next slide when all these immuno stains are done on the tissue microarrays.

(Refer Slide Time: 34:00)

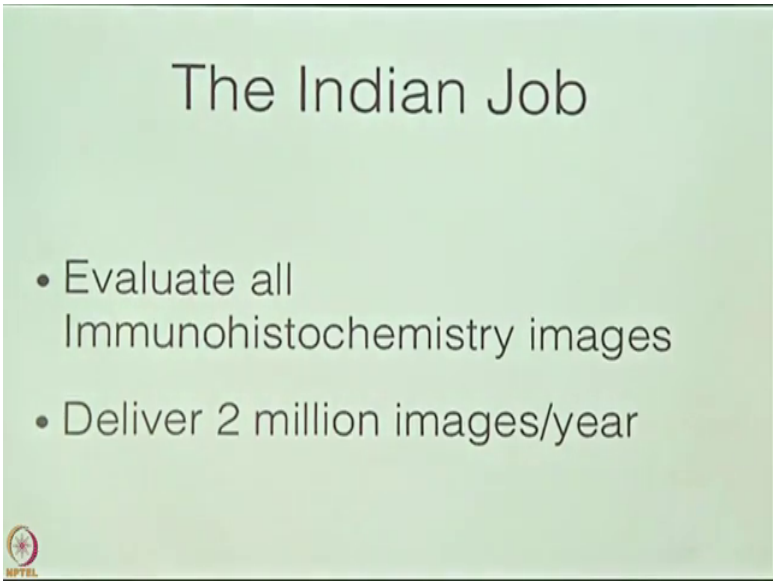


If you give that slide to the pathologist with all those 150, 200 or sometimes even 1000 cores; they are never going to talk to you again. Because, it is a physical challenge to look at that each core under the microscope and then say that that core shows this it is one of the most frustrating exercises in the laboratory. So, comes in the need for digitization. Just scan the whole slide and present it to the pathologist 1 core or 2 cores at a time. And if you add a software to that so, that they can go and click off the boxes then the pathologist is going to invite you for dinner.

So, these are some of the scanners that I use and I think that the ones that we made use of in this project. Internet in India was just coming in a big way the big lines you know we have not had these high speeds more than for more than 12 to 13 years actually, less than 10 years. In the big challenge was how would all these images be transferred to India were the pathologist to see it if you do not have internet that is fast enough.

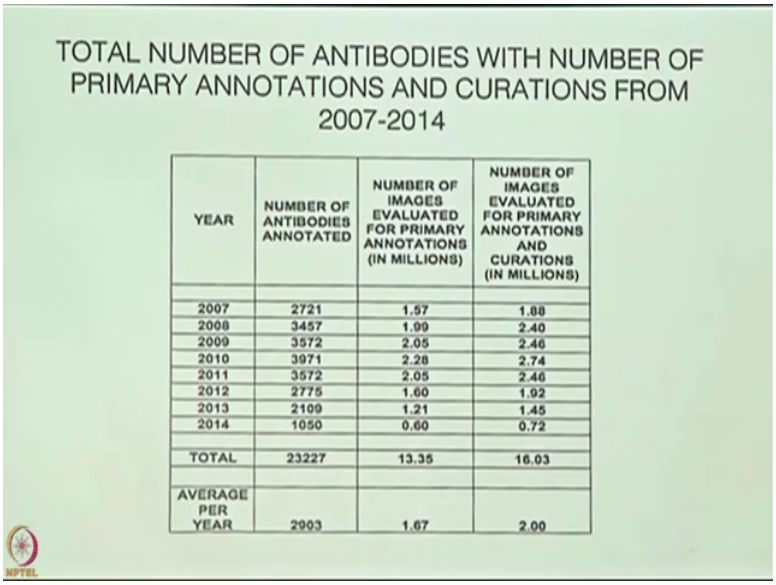
So, we actually downloaded several million images and it was a difficult time for us to work in. Because, when we started the speeds were not that fast and people used to be waiting and you would get only half the core. Then a glitch in the line then another half. Now luckily over the last 5-6 years it is much better.

(Refer Slide Time: 35:42)



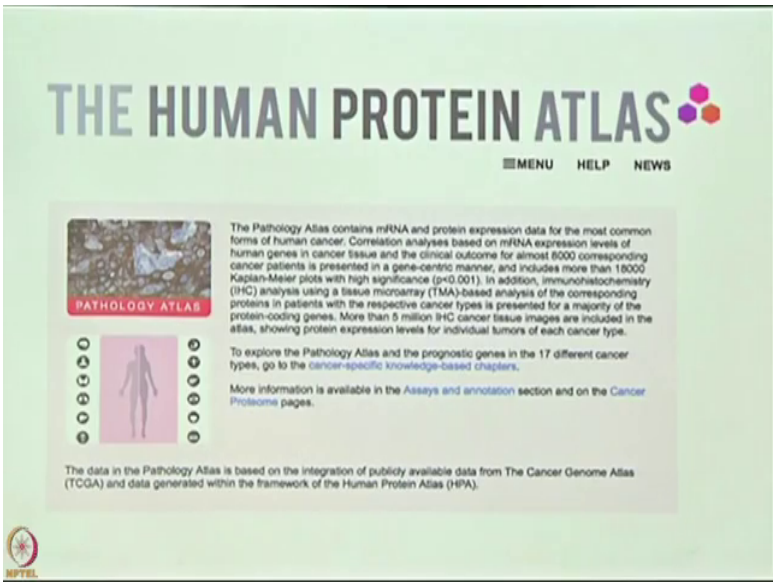
So, that was our job when we took it on. We had to evaluate all the I H C images and we had to deliver 2 million images a year. 2 million images a year is roughly 7000 images a day manually.

(Refer Slide Time: 36:03)



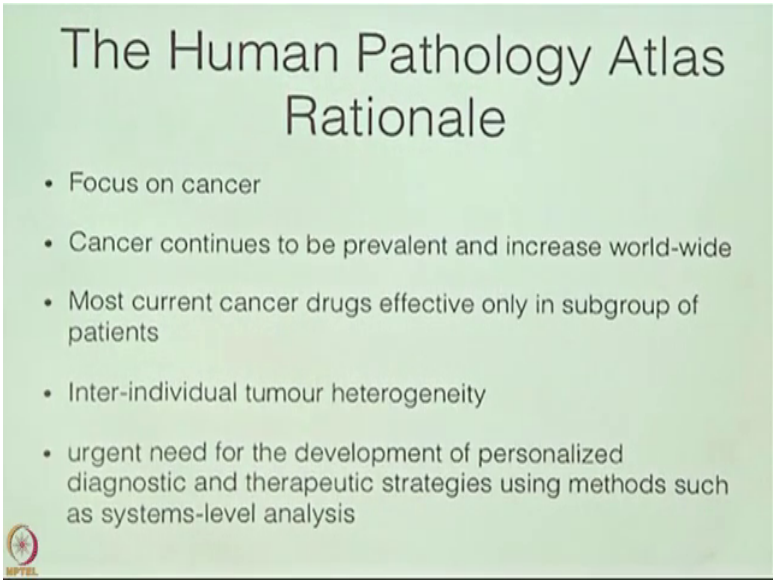
And those are the statistics that we followed from the year 2007 to 2014 that gives you an average of 2 millionaire year. It gives you 2 million because they were original images that were annotated. And then came the big question what about quality control this is from India and we had to work very hard to prove that there was really no difference whether they were annotated in India or in Sweden or in Norway or anywhere else and everybody agreed. The static became statisticians became our best friends because, they proved that.

(Refer Slide Time: 36:49)



Now, let me go on to the second part which is a more recent event in the protein atlas what I just spoke about was normal tissues, but we also have a pathology atlas section that deals with pathology, mainly cancer. And that was published in August 2007. The big proteome of the pathology atlas according to the transcriptome.

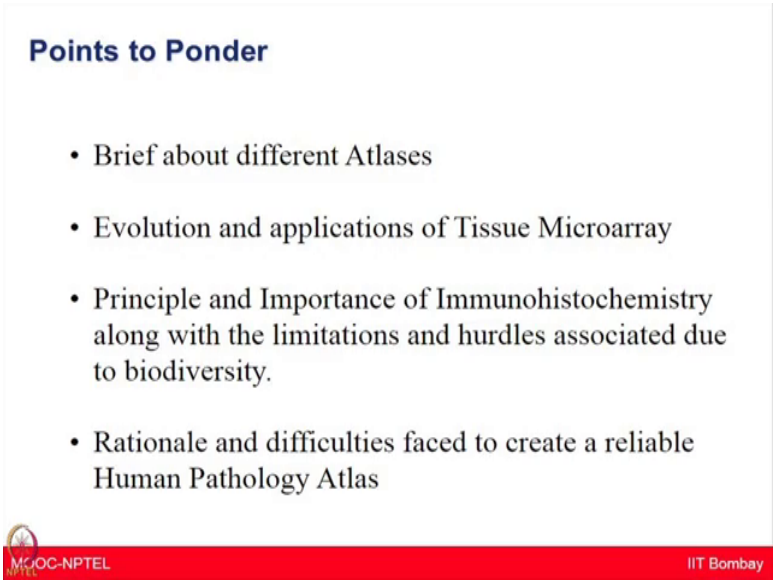
(Refer Slide Time: 37:29)



The focus for the human pathology atlas was on cancer for reasons that most of you are already aware about. The big challenge was that most drugs in cancer are effective only in a small subset of patients.

And therefore, was there any difference between 2 tumours from two different people looking the same under the microscope remember the tree without branches and remember the tree with leaves. And therefore, if there is so much difference, should the treatment protocol get personalized according to what that patient has or should they just be hit with everything. So, what was studied in this paper was a huge effort. It was a genome wide effect of genome rearrangements amplifications and specific cancer driving mutations in all the cancers that we could see.

(Refer Slide Time: 38:35)



In today's lecture, you got a understanding of a big project human protein atlas and also the major contribution by Dr. Navanis group on human pathology atlas. The pathology atlas contains MRNA and protein expression data from 17 different forms of human cancer. Some findings obtained from this project have really contributed immensely to the whole biologists and the entire field. Let us continue this lecture and more discussion about the human protein atlas and human pathology atlas in the next lecture.

Thank you.