

Interview with Dr. John Sulston (S)  
Lina Lu (L) at the Center for Public Genomics, Institute for Genome Sciences & Policy, Duke  
University  
November 16, 2007

Lina Lu: The information I am about to give you and your response will now be recorded. My name is Lina Lu and I am a student at Duke University. I am in a course on the history of genomics that includes an oral history project. One goal is to produce a written transcript of interviews with important figures in genomics. Some of the interviews may be archived or made public through a website. I selected you as the person I would like to interview. The interview should last about 45 minutes. Your participation in this interview is strictly voluntary, and you may withdraw at any time. You do not have to answer every question asked. The information that you provide will be —on the record and may be attributed to you. If you do not want something public, then we will turn the tape off and will ensure it does not appear in the transcript. This interview is being recorded and I will take written notes during the interview. The interviews that are posted publicly will be archived as a history resource. It may be included as part of the —Human Genome Archive established at Georgetown University with funding from the National Science Foundation, and now shared between Georgetown and Duke with funding from NIH and the Department of Energy. If you prefer that the interview be used only for the course and not made public, please indicate this. One risk of this study is that you may disclose information that later could be used in ways you do not like and cannot control, by those accessing the archives. The benefit of participating in this study is ensuring that your side of the story is properly portrayed in the history of genomics. Dr. Sulston, do you agree to the interview?

So, to start things off, could you tell me a little bit about your background, what got you interested in genomics, and how you got to where you are today?

Dr. Sulston: Well, it was natural progression, from working with the nematode, *Caenorhabditis elegans*, in the group here at the Medical Research Council, the lab at Cambridge, the Laboratory of Molecular Biology, and there we were engaged in a sort of expanding movement, starting with Sydney Brenner back in the '60s, to study developmental biology, in this little worm. Now, what was happening, and this is true for all experimental animals of this kind, model organisms during the '70s, is that the genetic side of things was going very well, but we didn't have the ability to isolate genes quickly, to actually get hold of the genes and study them. Sequencing methodology was invented towards the end of the '70s, which was the Sanger method, which allowed one to sequence genes once isolated, but it wasn't practical to sequence a whole genome at that time, so that was out. And I was studying various things, such as cell lineage, but anyway, in the '80s, it suddenly struck me that we had to do something with the genome as a whole, to improve the speed with which one could isolate genes. So, I set to with my colleagues and a great collaboration with Bob Waterston at Washington University in St. Louis, Missouri, which was absolutely crucial. And the course for the next few years is that we put together a nice map of the genome. Again, there was no question of sequencing at this stage. It would be far too long and costly, but what we could do was collect clones covering the whole genome and store them in the freezers and then map them one to another, using restriction enzyme procedure. Now, the reason I'm telling you all this is so you can see how it all flowed in a very natural way from biology to a requirement for a new tool. This initiative was paralleled by people working on

other organisms, the idea of laying out the genome, and then of course, the sequencing methodology caught out and became high speed. We were all able to, during the '90s, to start sequencing our genomes. So it was absolutely a natural progression. The odd thing for me was getting involved with the human genome, in a way, because that was accidental. I had no background in human genetics and didn't really expect to be involved. But simply because we, Bob Waterson and I, were involved in the instruments' methodology, we just found ourselves drawn in to the sequencing operation, which was known as the human genome project. In my case, very crucially, the Wellcome Trust became interested and set up facilities that we needed for the nematode sequencing but also for the instruments also used to sequence the human or part of it. So that was the story and it really shows how everything connects together.

L: One of the books we were for our class was *The Common Thread*, and I was really impressed by how firmly you stood against the patenting of the human genome and how it should be for the public domain. Why do you believe that the genome be part of the public domain?

S: This again is pragmatic and I take it back to once again to the '80s, where we were doing this tiny, small sequencing job with the worm. I say small because it was small compared to the human community. We found, Bob Waterson and I, in a very pragmatic way, that it was essential to have communication at the genomic level. Genomic communication, if you like. What it means, is that people doing biology, who also might be geneticists, because that's the goal, discovering genes, mapping them and producing genetic maps, which we can use as a basis for finding important genes that are important in development. Now, the whole process of mapping the genome only works if people share their information. So once we isolate a gene, they have to be willing to put it on the map and say here's a marker. Then basically, you align progressively the physical map of the clones in the freezer with the genetic map and then next time somebody comes along looking for a neighboring gene, there's a marker to go by, so we know exactly which clones to give them. And that process keeps on getting better and better.

If people don't share, then it doesn't work. I also show a diagram when I'm giving a talk about this. I show a database which is shut up so that people aren't sharing. And the result is that this communication doesn't work. You can have, for example, a private database containing sequence data or map data, but it's secret data. And what happens is that somebody pays to go look at the data, but they can't share the data with anybody else, so they can't contribute to each other's operation. And this actually becomes very serious with something like the human genome where the dataset is so vast and so little understood. And we have very much more to discover in terms of control sequences and there are new things coming out all the time, like micro RNA, for example, a whole group of molecules that people haven't suspected. It would be absolutely impossible to work on these things unless everybody has open free access to the genome because that is the map, if you like, as well as the substance that everybody is working on.

Well, let's shift back to mapping, which was the story, to sequencing DNA, the human genome. The human genome sequence is approximately three billion bases long. Now, if you adopt a privatized model, such that the genome, the information you get, is locked up and it's accessed by only paying subscribers, that's the business model of how to do it, which was proposed and indeed acted upon by Celera, and to some degree, other corporations. Now, if you do that, then you, in order to have a business model, you need to collect subscriptions from people, for access to the data. That's, after all, business, and it's fair enough that you get the

money for doing the work of the sequencing and letting them look at it. However, what they have to agree to do when they take a look at the data is they won't share it. If they started sharing the data and started telling everybody else what the sequence was inside your database, then you wouldn't have a business any more. So, all the subscribers of this facility have to agree to very strict rules to go through, that only the subscribers of the data can look at the data and that they may not pass it on.

Now, there are various ways in which people try to get around this and try to make it more flexible, but the truth of the matter is in the end the dataset of this kind is a very complex one which one is going to be working on for a very long time, because we don't understand it completely. The only way for this to happen is for it to be completely in the public domain.

Now, the reason I cite this particular example, I'm not saying that everything has to be in the public domain. There are particular sources of information which really should be out in the public. In fact, everybody agrees to this. People in industry agree to do this. They say it's fantastic that we have a public database for the human genome and all these other genomes because it's really helpful work. They call it tried and tested, saying it's more efficient if you don't have these things done on a pay-per-view business model, but rather just pay publicly once, get it done quickly and have it open. And so that's what we fought to achieve and did achieve.

It applies to, just to illustrate the point, various other kinds of information. A good example is meteorological information, very important with climate change. It's extremely important that people share weather data and so forth around the world. If you lock it all up into individual databases, which I'm sorry to say, Europeans have a tendency to do, then it does no good. The same thing applies to other sorts of biological data, for example, the structure of the protein. There's a big consortium which is discovering protein structures and putting them in the public domain, because again, people can learn a great deal by preparing and sharing them. All of this does not preclude people from going off and making inventions based off of them. It's just the essential data set should be public.

L: Well, during the Human Genome Project, you were basically seen as the figurehead of the British involvement of the public sequencing project. Could you describe your role in the international genetics community, then and now?

S: I would like to slightly modify your statement there. I mean, although in Britain, I was certainly seen that way, because I was put into the position of giving interviews and so forth, explaining how it worked and explaining the importance of public domain. Of course, in the US, other figures, Francis Collins, Eric Lander, and Bob Watson himself, were all very much the figureheads for the public domain. And so, I think that I should not be seen as the international figurehead although, of course, I'm not shy about saying what I think.

So, with that in mind, the information should not be seen as a sort of isolated standard by my part, it was part of the international consortium. Well, then what happened is that we were all in it together. It really was a consortium. We had consortium meetings in different places from time to time, every few months, to discuss things. We emailed a great deal and we also had conference calls on the phone every week just to get together and discuss policy and see how we were moving along with technology.

And now, I, right or wrongly, chose to stand down, when this was all pretty much in the bag, I stood down from the directorship of the Sanger Institute duty, the Wellcome Trust Sanger

Institute. That was a planned move, really, I've always intended to not to go on being the director of big lab for a long time. It was not my forte, not my choice. What I did do, I actually formally stood down in 2000, which was early on, but I remained the sort of leader for the human sequencing project along with extremely able colleagues here, of course. I was very much just the leader of. And what's happened since then, the publication of it came first and various other events sort of led to me being asked this and that and the other.

I found myself now not doing research science any more, simply because I find so much of my time is taken up with other things which I think are important, and so I describe myself as somebody who's involved in well, if you like, call it science in society, or something of that kind. I'm still very much involved in the importance of openness and the appropriate openness, not just any thing out, but the good things that matter. I'm interested in the publication movement.

In fact, I've just accepted, historically, my sort of proper day job is that I'm vice-chair of the Human Genetics Commission here in the UK, which is very much involved in this interface between technology and the public. We work for the government. We advise the government by writing reports on specific inquiries and we spend a lot of time organizing group meetings and so forth with interest known to the public to talk about human genetics and the development and so forth because it's a topic that people are rightly a little bit wary of. They want to know what's going on, they want to know what it means to them. We also have the National Health Service, which I consider to be a really good health provisions system as it accounts for everybody, as everything is free up to the point of delivery, but along with that cause, we must take care that the delivery is done properly and so we have the genetic commission, we have a genetic services subgroup, which constantly knows what's going on and try to make sure enough resources are going in and so forth. So it was responsible for making the British healthcare work and it's part of making and smoothing the passage of technology into healthcare.

A new venture, which I have not really started yet but I have been accepted into, the appointment was with a friend of mine, John Harris, who is a bioethicist and a philosopher at the University of Manchester, and a new job that I'm taking is to work with him on a new group called the Institute of Science, Ethics, and Innovation, which is just being formally set up. And this is taking on a broader view, not just in bioethics, but in general, the way in which science is translated into application, if you like. And I think that my experience in the human genome is that this is actually quite an important thing to take care of, that one really needs to find ways not to do things inappropriately.

In the health care market, for example, there's an awful lot of stuff being sold inappropriately, drugs being pushed on people who don't need them, there are all sorts of claims out there for things that are really not appropriate, things being overprescribed and so forth. And I think that this is something we need to think about - how we can manage that better.

L: So going back to the whole international thing, do you think that during the Human Genome Project, there could have been more international involvement as a whole? Or do you think that that would have led to an increased amount of competition overall?

S: No, I think it would be good if more countries had been more involved. As it was, we had a pretty good spread. It was true that a pretty large fraction of it ended up being done in the US because of the NIH and the DOE's involvement and commitment, and the UK, number two,

because of the Wellcome Trust involvement. The UK government did not put a great deal of money in. The French government were involved some, the Germans, a little bit. Japan, and later on, China, put in substantial amounts. We did have an around-the-world sort of activity, but it would have worked out great if it had been more evenly distributed. It's true, but we got what was possible from what people were willing to fund.

L: So, in your opinion, do you think that America is naturally a competitive country when it comes to industry? Were the differences between England and the US enough that it caused considerable tension between the two when it came to sequencing the human genome?

S: No, because we were in the public domain, and that isn't different. I mean, for America, Americans are a bit competitive, but the British can compete as well. Are you thinking of sort of an industrial pressure?

L: Yes, industrial pressures as well as invention pressures, going back to the patent issue.

S: Yes, well, back in the early days, the NIH had an internal discussion about whether or not sequence fragments ought to be patented. With the change of head to Harold Varmus, that discussion changed a great deal. It was Harold who knocked on the head the idea of patenting the sequence fragments. But that was tried under the previous NIH head, so he made a great difference there.

Jim Watson was always against gene patenting, mainly that was the reason why he stepped down from the Human Genome Project in the early '90s. I mean, the idea of patenting DNA sequences is very much an extension of the database argument, but perhaps with a slightly different slant to it. Again, you see, if you take a gene, a whole gene for the sake of the argument, not just a fragment, and you have a gene sequence, if you have full, exclusive rights type of patent on it, then what it means is that you can stop anybody else from working on it, at least in commercial terms. Many academic labs are also affected, because they are potentially producing profitable products down the line. So restrictions of this kind affect everybody.

So I think that we can say very roughly that if somebody has an exclusive rights patent of an isolated length of DNA sequence, then it stops other people from doing much at all. An example of this would be in the case of Myriad with BRCA1 and BRCA2, where they got the full portfolio on those genes. This refers to the portfolios in the US. The portfolios have been largely revoked by the European patent offices, which I think is a good thing. Unfortunately, it's not for the reason that the European office doesn't think that people should patent genes, but because they didn't think they found the right sequence. So that knocked them on the head. So they were left with a variety of very specific patents and there was a whole interesting story about that on which I'm sure you are aware of and I won't go into any more of that now.

The patent debate, I think has now moved on, to some extent, and the demand for utility has raised the bar to people getting patents. So it's no longer good enough for people to isolate a bit of DNA sequence and say, Look, I want to get a patent on this, which in the beginning, people thought was fine. But now you have to say what you're going to do with it. Now, just going back to the principle of the thing, it seems to me that if you do allow patents on DNA sequences, if you do lock them up in this way, then actually, it's very much opposed to good research practices and good business practices, because it's essentially a very anti-competitive move.

It's like patenting, and a famous analogy I always use, is with mousetraps and patenting the idea of a mousetrap so that nobody can even try to make a mousetrap, as opposed to patenting a particular design of a mousetrap, which encourages people to invent around you and produce a better design.

So it is with a human gene. A human gene cannot be invented around. That means for the duration of the patent, which may be quite long, other people are inhibited from competition. I think that is very poor practice. Really, the Supreme Court is beginning to recognize these things, that this kind of extreme patenting is not very productive, and I think that bit by bit, people are edging towards more rational decisions.

If you want to have patents on things like that, then, in my opinion, they ought to be non-exclusive rights patents. I think that there is a case, that if you have gone through the trouble to identify genes and isolate it, and that it was non-obvious to do so, which may have been the case in the early field of the gene, then it's perhaps not unreasonable to have some sort of rights to it, although certainly, I wouldn't want to do that. But they should not block other people from doing things. And I think that if the international patent community generally were to move towards remuneration rights patents, then I think that would be a marvelous thing, because what that means is that nobody is inhibited from using, from working a particular discovery. They could be required to pay a certain fee to the originators, but they can't be blocked from it and they can't be asked for unreasonable fees, the sort that Myriad would be charging for the BRCA gene. That's my opinion.

L: Okay, so given that a company, I'm going to name one here, Celera, was a proponent for patenting, with Craig Venter, do you think that a true collaboration between the public Human Genome Project and Celera could have happened?

S: We tried. We tried. I mean, there was a meeting at the very end of the millennium, curiously enough, the very last days of 1999. There was a meeting between the representative of the public representatives and that of Celera trying to diffuse the situation which had got fairly aggravated in the press. It all hinged on whether or not Celera was willing to release the data. And the financial head of Celera, Tony White, turned up on the Celera side and said that there was no way that the data was going to be released. Anyway, he was going to have quite a long time before he was giving the data freely.

This made it impossible for us to follow the model which I was explaining earlier, in connection with the research database. We were not able to have agreement. It wasn't to do with patents, really, it had to do with open release. So, I don't know. The thing is when you go up an offer with Craig Venter, it's very tricky, because he now says in his book, if you look at it, that he never wanted patents, that he never went for patents, and that it was misrepresentation to say that he did. Unfortunately, that's not the case. I mean, the record shows very clearly that he was going after patents. However, that was something he said he was forced to do and he didn't want to. So it was very difficult to know what Craig Venter himself wanted since different bits of records say different bits of things. So I really wouldn't want to cite anything with respect to him which is why I mentioned Tony White. He definitely had a particular view of how Celera's information was to be used.

L: Given that it takes a lot of money to do genomics, and given that Celera was the major challenger to the public sequencing effort, do you believe that the governments of the Human Genome Project provided enough incentive to sequence the human genome quickly?

S: You see, I don't accept that Celera had a quicker or faster method, certainly not at that time. It's true that people have moved on very usefully to more and more sort of shotgunning the genome as the machines improved. But to be honest, the cost had very little to do with the technique, which had more to do with the gradually developing technology of the machine. And the figures that were quoted by Celera about the greater progress and the cost of the public domain operation, are wrong. The press releases indicated that the public release was very expensive and it was not true. The public domain was actually very competitive so I don't think that there was anything lacking there. I mean, obviously, the more resources you have, the faster you can go. As you know, there was the element of a race, because if Celera had managed to get a lot of patents filed on the genome, then that would have been very bad for everybody. So we were anxious to make sure that we had data out there at all times to prevent that from happening, which we did.

L: You mentioned the press's role in the Human Genome Project and how they got the information wrong. What's your opinion of the press's role in the Human Genome Project?

S: It's very interesting. Let me illustrate. I had a conversation at a point, after 2002 with someone from the press. I think at the time I was a bit critical as I was fairly bruised by the sort of things that had gone on and the misrepresentations that the journalists would pick up on and then turn into headlines and I was pretty antsy about the whole thing. And she said to me, "Just think, how would you like it if there was only one media in the state control?" And I thought, "Yeah, you're right." And I kind of knew this inside me, it just focused me about the necessity of having a free media. And of course, if you have a free media, then you know there will be a lot of people writing different stories. Some care more about the story, some care more about the headlines, some care very much about accuracy. They tend to write for different papers or different broadcasts on TV channels and so on. I mean, you know the people you want to read because you trust them more than the others.

But the essential message of all this and the fact that there is going to be a feeding frenzy when some interesting conflict is going on, I think that it is an inescapable part of our free media.

Although I was angry at the time, I now see it as absolutely essential. I also now see the sort of world politicians have to live in, with people trying to get mucky stories about them all the time because they would try to sell the papers, and on the other hand, they are also using the media to communicate with the public. It's sort of inescapable messiness with the whole thing. If we manage to have a media which is both open and fair, then that would be a wonderful thing, but I would much rather it be compatible with the openness.

L: One of the big things that the media jumped on recently was the thing with Jim Watson and his statement.

S: Ah, yes, that. Absolutely tragic, and I think that it was a very negative thing on part of the British media. I had been aware of this three weeks ago and all this happened when I was away. As I understand, he was being Jim as always, and I don't think what he said on this particular

occasion was any more outrageous than the things he said before. He certainly has been sailing close to the wind with his remarks about genetics and behavior. We've all been thinking that it's better not to quite express things in this way. On the other hand, you know, at some point we are going to have to adjust to this, but we are going to have to adjust to them carefully. And what little I've seen of the transcript of his conversation, he was making this very unwise connection between IQ tests and particular races. And I honestly would not go there personally because I don't think we know enough yet. We don't understand yet the social factors are involving in things like IQ tests. We don't understand cultural background. All we can see is that in a homogenous group, things are fairly consistent. But if you take two different groupings, then it is not clear. And things aren't clear between IQ and countries and I think he was really not himself to go off base.

On the other hand, he was discussing a subject which was important and I think that the media jumping on him by creating this food fight was dreadful and the fact that he had to step down was a real shame. He is a politician in a sense. In science, he is very high profile and therefore, he has had to accept the results. So I'm very sad, but I don't condone the journalists who chose to play up to this story.

L: Craig Venter just recently published his genome. What do you think of this and would you have your own genome published and why?

S: Craig can do whatever he likes, it's a free world. I don't see any great value in publishing a single genome. All we know about genomes comes out from comparative studies and the reason why he is able to make statements about specific genetic variants he carries is not because his genome is published but because there's already been a lot of work done on those genes. So in that sense, what's new? It's true that now, what he and others say he is going to do is a sort of interest here at the Sanger Institute, sequencing the genomes of many individuals. That of course, would be valuable.

It's having large cohorts of genomes, not having a single one, which is going to expand our knowledge. Would I have mine done? No, I certainly won't want to at this point, to have that sort of money spent on my genome. I think that that would not be sensible. To be a volunteer, to be one of the many who will be sequenced, sure, that's not a problem. I don't mind being sequenced, but I wouldn't want to put myself forward in that fashion.

There is an issue which I hope we will get over with all this and I agree with Craig when he says that it's good if we develop more openness. I'm very sensitive to the people's feelings about this because I work with the Human Genetics Commission and some feel very much that we should keep all of it under lock and key. And so, one of the things I'm working for the Human Genetics Commission and that I'm following very avidly is the work being done on the genome and the US. You know, the bill that's in the Senate now. I think that it is extremely important that we develop a statute that prevents genetic discrimination against individuals. I think that if we can do that, then we have gone a very long way in defusing any worries we might have over people's knowledge of the variants of the human sequence. However, we do have, I think, for people, to have a degree of respect.

One thing that the current generation is going to have to come to terms with is that everybody will know who their father is. And this is quite new in that people have been able to get away with fathering children and escaping the consequences in the past and that will no longer be the case, clearly. There are also people's concerns about the genetic problems about



genetically influenced diseases, the sort of thing in the family that people are going to have to come to terms with and that families will have to share their genetics and that if you have something damaging in your family, then we should know more about it. I think that so long as we respect people, we obviously must respect people's dignity and not just show genomes around without their permission.

But I think we should get used to the fact that this is something that we do know more about and more and more genomes, probably all of the genomes within 50 years or less, will be sequenced. It's just the natural course of practice. And we shall have to not be too sensitive about this data. Just keep it reasonably private with all the information.

L: Where do you think the future of science is headed right now?

S: Science? Science in general? That's a big question.

L: Do you want to focus it down to just genetics, then?

S: No, no, no, not particularly, I hope that it's going somewhere that no one has ever thought of yet because that's the fun part of science, is that new things happen. So if you focus it down and say, "Well, where is genetics headed?", I think that there is going to be more and more information. I think that the understanding, which is to me, the interesting part, is going to take rather longer. I suppose as far as genetics and genomics are concerned, what I want to see—this is happening to a considerable extent—is that it is going to get incorporated back into biology. We've been in a sort of wave of excitement about genomics and I've told you my personal history about how I just got into mapping because we needed the tools and then I got into sequencing because that was a tool and so forth. But we've done that, we've done quite a lot of that and there's still a lot of routine sequencing to be done.

But the important thing now is to use all of this information and to incorporate it all into biology, which again goes back to the reason why it is so important that it should all be open. You can have the genome, and not just the human genome, but all the genomes anywhere. You don't have to pay, you don't have to do anything, you just do your work and they will become part of the tools on every biologist's bench. So that, as far as I'm concerned, is the important message in genomics, you know, since it's not special.

This is just another chunk of information we know more about.

Where is biology going? Well, obviously, to the extent that we can understand, manipulate, and modify, genetic engineering is obviously going ahead. There's this great excitement in biology of the idea of producing novel genomes. That's all very fine, but the important thing is to understand what you are doing. And I think we have a lot to learn before we can start to put together really useful, functional things such as putting together new genomes which are different than the ones we've got. But to start doing really different things, we must really understand the genes and certainly in that sense, biology has a really exciting future. But the thing is, the future has to be thought out in terms of systems biology because we know very well that the genes are highly interactive with one another and many genes operate in more than one place.

Until we understand a great deal more about the control of genes and how they interact, it should come progressively, but until we do that then we can only search for the possibility of producing novel life forms but that's in the future.

Of course, we need to constantly be ethically looking out. I hope that by involving the public, people will go along with it together, so that they won't worry about mad scientists and so forth. It's being talked about a lot and I'm very pleased with the way that communication is being opened up more and more, you know, people writing. I'd like particularly to see developments like the kind that is bringing younger scientists into contact, talking about their work. I think the more of this we can do, the better, and this will mean that we will have a very exciting and not empty journey ahead of us.

L: Thank you for your time, Dr. Sulston.