Informed consent for: "The ethos and effects of data-sharing rules: Examining the history of the 'Bermuda principles' and their effects on 21st century science"

University of Adelaide

Duke University

Researchers at the University of Adelaide, Australia, and the IGSP Center for Genome Ethics, Law & Policy, Duke University, are engaged in research on the **Bermuda Principles** for sharing DNA sequence data from high-volume sequencing centers. You have been selected for an interview because we believe that the recollections you may have of your experiences with the International Strategy Meetings for Human Genome Sequencing (1996-1998) will be interesting and helpful for our project.

We expect that interviews will last from 30 minutes to much longer, but you may stop your interview at any time. Your participation is strictly voluntary, and you do not have to answer every question asked.

Your interview is being recorded and we may take written notes during the interview. After your interview, we may prepare a typed transcript of the interview. If we prepare a transcript, you will have an opportunity to review it and to make deletions and corrections.

Unless you indicate otherwise, the ***information*** that you provide in this interview will be "on the record"—that is, it can be attributed to you in the various articles and chapters that we plan to write, and thus could become public through these channels. If, however, at some point in the interview you want to provide us with information that might be useful for us to know, but which you do not want to have attributed to you, you should tell us that you wish to go "off the record" and we will stop the recording. We will, however, take notes for our own use. When you are ready to go back "on the record," we will resume recording. Anything you say while "off the record" will not be on the audio recording and therefore will not appear in the transcript.

All ***materials*** from your interview (audio recording; transcript; interviewer's notes) will be available only to members of the research team affiliated with this project, unless you consent to their wider use, as described in the paragraph below. The digital materials will be maintained in a secure, HIPPA-compliant drive at Duke University. The paper materials will be stored in a locked cabinet.

In addition to the scholarly articles and chapters that we plan to write, we also hope to create a resource for other scholars and members of the public. We plan to post some of our research data to online digital archives. While we will use your "on the record" comments to inform and write our articles, we will not post your interview transcript or audio recording online unless you give us permission to do so, in a separate agreement. At the time we send your transcript to you for review, we will also provide a consent form asking your permission to post your interview transcript and/or audio recording online. The form will provide you with different options for how, when, and with whom the materials may be shared. You will, of course, also have the option not to share the materials beyond the Duke and Adelaide researchers.

One risk of this study is that you may voluntarily disclose identifiable information that later could be requested for legal proceedings, or otherwise be used against you. Please take this into consideration when you are speaking. There may be other risks associated with your "on the record" views being made publicly available, such as having your views mischaracterized or misunderstood.

The main benefit of participating in this study is ensuring that your side of the story is properly portrayed in this history of the Bermuda Principles, which have become a model for open and collaborative research in genomics and other fields.

To help us protect the privacy of those parts of your interview that are not public, we have obtained a Certificate of Confidentiality from the U.S. National Institutes of Health. With this Certificate, we investigators cannot be forced to disclose information that may identify you, even by a court subpoena, in any U.S. federal, state, or local civil, criminal, administrative, legislative, or other proceedings. We researchers can use the Certificate to resist any demands for information that would identify you.

The Certificate cannot be used, however, to resist a demand for information from personnel of the United States Government that is used for auditing or evaluation of federally funded projects or for information that must be disclosed in order to meet the requirements of the federal Food and Drug Administration (FDA).

A Certificate of Confidentiality does not prevent you or a member of your family from voluntarily releasing information about yourself or your involvement in this research. If an insurer, employer, or other person or institution obtains your written consent to receive research information, the researchers may not use the Certificate to withhold that information.

Signature R J Durbin

Printed Name RICHARD DURBIN

Date March 6, 2012

If you have read this form in its entirety and agree to the interview and its terms, please sign and date above.

Contact information:

Rachel Ankeny, Ph.D. (University of Adelaide)
rachel.ankeny@adelaide.edu.au
+61-8-8303-5570
Kathryn Maxson, B.S. (Duke University)
kat.maxson@duke.edu
(919) 668-0791
Robert Cook-Deegan, MD (Duke University)
bob.cd@duke.edu
(919) 668-0790

*If you have any questions about your rights as a research subject, you may contact the **Duke University Institutional Review Board** at 919-684-3030 or ors-info@duke.edu.*

PLEASE FILL OUT AND RETURN THIS FORM TO: *Center for Public Genomics, Duke University; c/o Susan Brooks; Center for Genome Ethics, Law, and Policy; 304 Research Drive, Box 90141; Durham, NC, 27708.* ***OR****: You may fax it to us at (U.S.) 1-919-668-0799.*

Interviewee Information. *Please list an address where we can contact you.*

Full name: Richard Durbin Date of interview: 9 March 2012

Current institutional affiliation: Wellcome Trust Sanger Institute

Street Address: Hinxton, Cambridge, CB10 1SA, UK

Phone: 01223 496848 Email address: rd@sanger.ac.uk

Interviewer Information.

Full name(s): Kathryn Maxson

Affiliations(s): Duke University

I, the undersigned, have read the above, and I **AGREE** to release my interview materials, subject to any restrictions listed below:

(A) ☑ I place **no restrictions** on my interview materials.

OR

(B) ☐ My interview materials may be reviewed, used, and quoted by the researchers affiliated with the Center for Public Genomics, Duke University; *and in addition* (check all that apply):

☐ Researchers unaffiliated with the Center for Public Genomics may **read** the interview transcript and any related documents only after obtaining my permission.

☐ Researchers unaffiliated with the Center for Public Genomics may **quote** from the interview only after obtaining my permission.

☐ Researchers unaffiliated with the Center for Public Genomics **DO NOT HAVE** my permission to **read or quote** from the interview.

Posting interview materials to public digital archives: In spite of any restrictions listed above, I give permission for my interview materials to be made publicly available on the Internet by deposit in an institutionally affiliated archive:

☐ 1 year from the date of this form

☐ 5 years from the date of this form

☐ 10 years from the date of this form

☐ 25 years from the date of this form

☐ After my death

☐ Other: _____ (please specify a date or condition)

☐ NEVER: MAY NOT BE DEPOSITED IN A PUBLIC ARCHIVE

Please specify any further restrictions in the space below:

_____

Signature: R. M. Durbin Date: 4 July 2013

Interviewee: Richard Durbin
Date, location, method: 09 March 2012, Durham, NC, by phone
Interviewer: Kathryn Maxson

KM: So we have [RDurbin] here on the line, and [KM] on the interview. And just to review for the tape, we have archived and received your informed consent. Do you have any particular questions regarding that?

RDurbin: No, that's fine with me.

KM: Okay, and just to review quickly, we're making a recording of this interview and then we're going to send it to a transcriptionist on a secure server so no one else will see it. And the recording is going to go right into our HIPAA-compliant drive so it's where personal health information goes with the Duke Medical Center. And once we have that transcript we'll send it to you for review, additions, deletions, anything you want. And we'll send along with that transcript a file that lets you choose what it is we can do with that transcript. And only the transcript that you've edited and made any corrections to, aside from our own research group… And you can check a box that says, you can show this to anyone or only you guys can use it, and pretty much anything in between. So that's what our process is. So you'll be hearing from us very soon after we get that transcript and it should be maybe two weeks or so at the most.

RDurbin: Okay, I'm aware of all that. That's fine.

KM: Wonderful. So we have a rough list of questions here and we follow this very loosely, but the first thing we always start with is: we have you at the Bermuda meetings in 1996 and 1998. There were three meetings and it looks like you were at the first and the last ones but not the one in the middle. Does that seem right?

RDurbin: That sounds correct to me. I think that's right. I've been trying to recall all the details. I believe I went to these twice and I think I remember missing one.

KM: So in your recollection, why were you invited and what was your perception of what was going to be discussed at these meetings before you entered into them?

RDurbin: I was invited because I was working at the Sanger Centre with John Sulston, Jane Rogers and David Bentley on the public human genome project. I was a founding member of the Sanger Centre with John Sulston. The Sanger Centre had been collaborating with St. Louis on sequencing the *C. elegans* genome. The Bermuda meetings took place at a point at which the human genome project was getting itself going; the aim was to coordinate the plan for sequencing the human genome. They were organized as I remember by the Wellcome Trust and NIH, and they involved people who were engaged in or wanting to be engaged in sequencing the human genome. This was essentially putting together the consortium that became the Human Genome Project. Quite a lot of the purpose was partitioning up the genome to work on, to get commitments of people to delivering particular pieces and achieve standards, to establish standards and coordinate the project.

KM: And so did you know in advance that there was going to be discussion of data sharing norms?

RDurbin: So it is interesting - that came out of the process from my point of view, rather than being the primary aim of the meeting. I've actually been looking at my email around the 1996 meeting, and I see there was some communication between John Sulston and myself and LaDeana Hillier about this; John was going to talk about the worm and he also did want to talk about data release. St. Louis and Sanger had already been releasing data as they sequenced the genome of *C. elegans*. The genome sequencing center in St. Louis headed by Bob Waterston and the Sanger Centre headed by John Sulston were strongly coordinated in lots of ways with respect to our approaches to genome sequencing. There was a joint agenda to promote early data release as part of the human genome project on the basis of its success in the *C. elegans* project. So yes, I knew prior to the meeting that this would be brought up.

KM: And how did the process work with *C. elegans* and what elements of the *C. elegans* model were you interested in importing to human sequencing?

RDurbin: In *C. elegans* a physical map was made originally by John Sulston with Alan Coulson. Bob Waterston came on sabbatical in the 1980s to the LMB in Cambridge and introduced the YAC technology that Maynard Olson had developed. Bob started working with John on the worm genome at that point and helped complete the map. And then Bob and John jointly took on sequencing the genome as a natural next step. They were supported by Jim Watson in doing that and by the MRC in England. And then there was a deal done; it was a complicated history which you can probably get elsewhere, with the Wellcome Trust and the MRC to set up the Sanger Centre to complete the worm genome and take on the human genome.

In the worm there was a long tradition of sharing data before publication, not just genomic data. We had something called the *Worm Breeder's Gazette* that came out three times a year, and most people who were in the field would write a page about what they were doing with relevant results that would interest other people and share information. John established early on a way that people could see the physical map as it was being developed and see the current state of it. I got involved in this. Before the Internet was available people could log in by modems and access the worm physical map as it was being built. That was very important in the worm field and there was a lot of good feedback. The same sorts of mechanisms also applied in the yeast field and some other genetics areas, in *Drosophila* to some extent. So when we started sequencing, first we had a physical map to start from. Each chromosome was split into two, roughly, between St. Louis and Cambridge. It was very much an equal partnership. But it was competitive. We both wanted to push forward to be ahead of the other. Friendly competition. We used to visit each other on a regular basis, exchange all

the methods. But still quite competitive. We made data available in the same way as the physical map, and people used the data from an early stage. And so there was a track history of data sharing and no sense of threat. The competition that there was, was in a very supportive structure. Data sharing also helped to coordinate things. There were two functions, really. One is coordination and the second is enabling science to progress and getting the support of the community. So something we recognized early on was that it facilitated people giving us very large grants to sequence these genomes. As part of that deal, this was a grant that was large compared to any normal research grant or research group. Our idea was that we weren't going to try and control all the scientific kind of capital coming out of it, that we were giving that back to the community, the research community. We would get credit for doing the job because it had impact, but we wouldn't try and control or unfairly exploit what happened downstream. So that created a win-win situation that was good for the research community and for us, which we recognized worked well.

When it came to the Bermuda meetings there was not such a culture of sharing information in human genetics. It was a much more competitive. There were many people who sort of wanted to think that they were going to get the glory of sequencing the human genome or bits of it. Or particular genes. There were a lot of independent parallel efforts. The map also had not been constructed in such a strong fashion genome wide. In fact all the way through to the draft sequence the map was being constructed in parallel. There were multiple strategies taken. There were different sorts of human genomic maps. And the scale is different. It's 30 times bigger than the *C. elegans* genome. And so various people put together local bits of map by walking strategies to try and find genes that they wanted to sequence…people wanted their bits sequenced. There were maps for regions and there were some whole genome maps that had different sorts of problems and issues and were incomplete. There was a complicated situation essentially. And so there was a sense that a major requirement was coordination, making sure that people weren't sequencing the same piece.

I remember a complicated situation particularly on the end of the X chromosome, which was known to be a gene-rich area where there were a number of disease genes. There were multiple groups in the world who all wanted to sequence the same part the genome and all had different maps, and it seemed pointless to be replicating and stomping all over each other. And so coordination was important. And when you coordinate it's best to share information. So sharing mapping information and sequence information was important.

A second driver I remember was transparency about standards. There was a lot of discussion about how you achieve a consistent quality. We can come back to that. And a third was about exploitation and control and that it was in the interest of everybody if we were seen not to be controlling, and that's always a difficult issue to persuade people to give up something that actually has value. But then our

position was that by giving up something of value you got the confidence and the support of the people, the research community you were giving it to. That then allowed us to focus on sequencing and not on controlling the research exploitation of the sequence.

KM: To what extent also were the data sharing norms a mechanism of keeping tabs on what different labs had accomplished?

RDurbin: There were representatives of centers there and there were various advisors. It was necessary to get people to commit to doing something and then establishing that they had done it. There is a tendency for people to make big territorial claims and then sort of hide. Making available the data of what you had sequenced allows transparency to see what is actually getting done. That was an important driver for the project, for the openness. But also I shouldn't underplay that from the start I think John had strong views about the exploitation side, the patents/IP type exploitation as well as the scientific research exploitation. There was a strong argument that we should not be taking out IP on the genome sequence and that the public consortium should sign up to that as a principle. And so that was a part of this process. And as other people must have reminded you, I believe Craig Venter was at the first of those meetings and signed up to the principles at the time. Circumstances change. That was when he was in NIH, I think.

KM: Yes, yes. And so patenting was also something that you had an idea was going to be discussed before you came to Bermuda in 1996.

RDurbin: Yes, well, gene patenting was already a live issue because there were patents on cDNAs, and I guess I don't even know, were there already patents on pieces of genomic DNA? Patent applications?

KM: I know that the *BRCA* controversies were heating up for breast cancer.

RDurbin: Had that happened already by '96? Or was that 1995?

KM: Yeah, that was '94 and '95, I think.

RDurbin: '94 and '95, I see. So then we'd already gone through that.

KM: Yes.

RDurbin: So Sanger was involved directly in that, as you're aware probably.

KM: Uh-huh.

RDurbin: Because we had sequenced Mike Stratton's clones. And so with *BRCA* we were directly involved in the patent argument, and we had a strong belief that that was

inappropriate, and so we felt that we should aim to establish as part of the international consortium sequencing, that those would be the guidelines for the international consortium. So John, I would say, showed leadership on that.

KM: Yeah, yeah. So you mentioned Craig. And I'm curious how you remember the tenor, the feeling of the discussion in that meeting or in that session at the end of the 1996 meeting, where John wrote up on the whiteboard. Do you recall this session? How do you remember the atmosphere in the room?

RDurbin: Yes, well I know the slide because I have a photo of it somewhere or another. I have it in my PowerPoint slides, John's handwritten slide, which probably you have a copy of or have seen.

KM: Yes.

RDurbin: And I think, as I remember, the room and where I was sitting even. Let's see, where was Craig? I remember it was sort of a horseshoe and I was on the left side, and I think the whiteboard was at the right side toward the door at the back, even. Or was it a paper flip chart? I can't remember that. Now, the tenor of it all. There was debate about it. A lot of the debate was around the issue of quality and whether we would confuse people by putting out unfinished sequence, and there was this issue of the actual measure, that you shouldn't put out too fragmentary things because that would just be contaminating somehow. That has always been a kind of argument against release of data, that it's not ready yet. I remember trying to persuade people that this wasn't a major problem; that people could sort it out. In some sense my view is that you shouldn't worry about the users…the smart people would know what they're looking at and know how to use it, would learn. And the people who got confused and misled, that was their own fault. You shouldn't protect the dumb people from getting confused, and in the process prevent the smart people from getting useful information. That's always been my view about such things.

KM: So you didn't …

RDurbin: I think I remember saying that or I remember that argument being important. And we ended up agreeing that we weren't going to release raw reads at that time. We were only going to release assembled contigs where the contig was bigger than one or 2kb, I can't remember.

KM: Yeah, it was 1kb. *[KM: It switched to 2KB later on.]*

RDurbin: One kb at the time. I think at some later time when reads were longer we probably went to 2kb for some projects. So it was contigs bigger than 1kb, which were therefore less likely to be rubbish and single rubbish reads would not be included. And so that was a kind of crude, practical quality control. And then the argument

was that was useful enough. So there was quite a discussion of that. The discussion about the patenting, so let's see…so some countries couldn't…this was an issue for the Japanese, actually. They couldn't sign up. In some cases the funder was present, like the Wellcome Trust was present with Sanger. NIH was present. You'll have to remind me, was this after or before Francis took over from Jim?

KM: Francis had taken over by this point.

RDurbin: Yes, I think this is right. I think Francis was in charge. Anyway, so NIH was there. But some countries, the kind of group who was doing the science was present but not the funder.

KM: But not the funder, right.

RDurbin: And I think some people, both the Germans and the Japanese, said that they were supportive but they didn't know if their funder would allow them to give up IP.

KM: Right, right.

RDurbin: So there definitely was pushback from various people. What do I remember about Craig's position on this? I remember Craig as sort of being the…not the outsider, an outsider, kind of a slightly…as he has cast himself repeatedly…on the other side. Coming to the meeting I remember meeting him in the bar and thinking, oh, that's Craig Venter, you know. But we all sat down together and talked about things. But we, Sanger, didn't have anything like the exchange relationship with him that we had with St. Louis, for example. But I explicitly remember him signing up to the agreement. There was a reasonably extended discussion and there was a process of going around the table, making sure that everybody had their say, and I believe everybody agreed to the extent that they felt they were allowed to.

KM: Yeah, but …

RDurbin: That's my belief as to what happened, yes.

KM: But with some hesitations because there were some folks who were concerned about their funding agencies.

RDurbin: I remember that caveat. So was Craig at TIGR or was he…I can't even remember now. It was before Celera, obviously.

KM: Yeah, he was at TIGR.

RDurbin: He was at TIGR. He'd left NIH.

KM: By 1996. Yes.

RDurbin: Yes, he was at TIGR.

KM: Yeah. So I want to move on from Craig. What personally did you think was most important to include in the Bermuda Principles? Was it that the data should be released within 24 hours, or was it more that as long as the data were released eventually that was more important? Or what did you feel was most important for the progress of the project and the other goals of the data release policies?

RDurbin: The principles are important. But I think I said in the beginning my sense at the time was that the meeting was not originally about coming up with these principles. It was about coordinating the sequencing of the human genome. And the data release principles were a component. They turned out to be a very important thing that came out of it. But they were at that time project-specific. Though by the third meeting, the second meeting I went to, we actually discussed extending these principles to other projects—but initially sequencing projects, for example the mouse, which was starting.

So what were the important things? I strongly believed from the worm experience that it worked in terms of success of the project benefit, to have the data available early. So two things, scientific side and the coordination side. I was pretty principled about these things, I think, at that time. I think that I believed it was good, that was the right way to do the science. And the coordination side also I would make all those arguments. I actually believe in that strongly too. I think it was certainly good. I have no regrets about the Bermuda Principles. I think they were a very good solution for the Human Genome project. In contrast, I do think it's reasonable to question what came out of, say, the Toronto meeting which only happened a few years ago. We should be able to ask, for what projects and what data ought there to be a principle of early data release?

KM: Right, yes.

RDurbin: I think that there's a very strong principle that primary data should be available at publication. Clearly results should be available when you publish. For protein structures there had been a big fight about whether you put structure coordinates in or have a nice picture or something. And for DNA sequencing, the DNA sequence in GenBank, that's really, really important. On publication or no later than publication, getting the thing that you are reporting as a data set at that level is absolutely essential. I think that it's part of the scientific requirement, where data generation is a part of the project. But very early primary data, I think that's discussable. Do you make the gel sequence available? Or do you make sequencing reads available, or do you make the assemblies, or exactly what level? There are pragmatic discussions there. Also I think there are discussions about whether it should be done early. There are good reasons for doing it early, but I

don't think that that has to extend to all research that everyone in the world ever does. Certainly in these big, coordinated projects I think it makes sense: where relatively large amounts of funding are being committed in a concentrated way to produce a community resource as a primary output, I think early data release is a good thing. So we've moved on to a philosophy.

And to what extent did I think like that at the time? I saw this as extending the worm principles to the human, and us taking on board the commitment to sequence human, and this was the right way to do it. And us, being the Sanger with Wash U, supported by the Wellcome Trust. Some of this was Michael Morgan in particular, and through him the Wellcome Trust. And in fact, no doubt others in the Trust, and NIH. And I think they were the people who drove it, with people like Maynard Olson and other people in the American context being supportive as well.

[KM: Maynard Olson actually had several reservations. See interview with him from February 2012.]

KM: So now that we have moved on to some more philosophical topics, how would you describe the subsequent implications of the principles for the nature of research generally and for researchers? And I'm coming at this question with the idea that lots of folks in the literature refer to the Bermuda Principles as this touchstone for open science and Mertonian norms of sharing data. Do you think that the Bermuda Principles helped to kick off an era of more open data-intensive science? Or what are your views really on the philosophical implications of these for the practice of science?

RDurbin: So one thing it's important to be clear about, is that this was not kicking off, this isn't the start of openness of scientific data. There were many examples beforehand. For example I have talked about this particularly in the model organism genetics fields, and Bermuda directly came from that. But I think in other areas of science there had also been people who had collected and put together important data sets and made them available to people as they collected them, then published papers at the end of that process. So I don't think there's anything that was new. But they are a kind of a touchstone. I do feel it was very, very important. It was important that the human genome was done that way and I think it's not a coincidence that it helped the human genome get sequenced efficiently. There are other ways it could get done. And basically it did get done in parallel in a different fashion, but I think it's very good for research science that there was an open strategy. And it was useful during the time that people did make use of the data as it was being collected. I think the transparency helped us both coordinate and achieve standards and be honest about what we were doing. And it helped in the emphasis on making available the results and therefore the use of the sequence. And essentially the public sequence from my point of view won, although it was considered in some political sense a draw, because almost

all research is based on the public sequence, and I think that's because of the informatics framework that we put in place through NCBI and Ensembl and EBI and Santa Cruz. All of them played key roles, and other places, in presenting the sequence, making it available. The lack of IP actually was essential there in some sense to getting it fully used, from my point of view. So I think that's all good.

What came from that? Certainly for the other genome sequences, starting with mouse and then rat and so on, a good model had been established, and carrying that forward was an obvious thing to do and was good. And then in other genomic science or large-scale science, I think that there's been a driver. I think it's actually in the funders' interest for the data to be available. There is this problem. If you focus resources into somebody to produce a really important resource and you have some sort of coordination across the world to ensure that it's done without redundancy, then you're giving a sort of monopoly of information to the people presenting it. And I think it's become a really useful and powerful principle that as part of that deal you prevent them from controlling the resulting information as a monopoly. Part of the deal is that they make it available. And once that principle is established, there's no reason not to be making it available as you're collecting it. That enables transparency. Time is of the essence in research. It's very time competitive. Getting there first is what counts. So that all works. I had a thought about the way that human genetics has gone because of these issues about ethics and privacy that create a conflict with totally open data release. And I think that that has become a complicated problem, that that has hindered the field and continues to do so.

KM: During the course of the human genome project did you see any major impediments to the implementation of what became the Bermuda Principles?

RDurbin: Well, I mean the deal was it was meant to be every 24 hours or nightly. That was not adhered to strictly, I would say. So there were clear impediments. It took effort to achieve that, and there were people for whom that wasn't a priority. And there's been a lot of that sense, a lot of kind of lip service, but actually things not appearing or taking a very long time to appear. So it's easy to lapse. There've been a couple of famous occasions where people were scooped. Some in a very blunt way where somebody sequenced something, was completing it and had released the data early. Somebody else just took their data and basically published it. That happened, as far as I'm aware, once to Lee Rowen who worked for Lee Hood on part of the T-cell receptor region in humans, or maybe mouse, not sure. And it happened also with Bart Barrell who was sequencing yeast at Sanger, who was always more suspicious of this approach. And somebody took all a bacterium, I can't remember, some microbial sequence and published it. So you are exposing yourself to that risk. And it also happens in less dramatic ways that you are exposing yourself to being scooped. So people are wary. And it's quite hard to get, it's just repeatedly hard. There are people who have released data early and been good citizens, and then somehow slide and don't for one reason or

another, or get talked out of it. I think that has happened. So it is not a continual …it's not like once people see the light and start doing it that they do it forever for everything. I think that we have to face up to that not being the case, really.

KM: Right.

RDurbin: So I think overall there has been a continuing trend. Also when there is a new data type you have to persuade a whole new set of people. And people always have arguments for why they don't want to do it and why their data are different. And in my view typically those are somewhat spurious.

KM: And just kind of along those lines really quickly, it does sound like you were at the Toronto meeting. Were you also at the Fort Lauderdale meeting in 2003?

RDurbin: Yeah, I was at that.

KM: And what was the discussion about there, and also in Toronto?

RDurbin: The Fort Lauderdale meeting was explicitly aiming to extend the Bermuda Principles from the human genome project and from specific other projects to a broader set of projects, and it was aiming to sort of codify the Bermuda Principles in a general form, for more general data types. There was this idea of a tripartite responsibility that came up and was codified in some sense. So yes, the Bermuda Principle was about that particular consortium. It wasn't trying to preach to the world about the way that science should be done. It was about what should be done with the human genome project. Fort Lauderdale was about actually large-scale sequencing, I believe, primarily, maybe looking toward other disciplines. Toronto was a sort of follow-on, trying to draw in other communities and extend out further the Fort Lauderdale principles. One thing that people misinterpret about the Fort Lauderdale principles, they think that…so there's this tripartite responsibility, the publishers, the data creators and the users…and they think that the deal is that the users are not allowed to scoop or to publish, to scoop the data generators. But actually there is no contract in place. It was very explicit indeed, that the data producers are not allowed to constrain the users, to prevent them from publishing. It may be bad form, it may be inappropriate, but the users are definitely allowed to use the data. I think Fort Lauderdale was trying to establish what the ethical structure should be in some sense. And also what the rules should be. The funders are also involved here, so it was trying to confirm the model under which the funders could require the data producers to make data available early and what the principles of that would be. So the Fort Lauderdale principles, for example, have been used by the Wellcome Trust and referred to by them in their grants to others and us as being the basis under which they're expecting, and I think maybe it's a general term. And I don't know if…NIH also I suspect refers to them. That was the aim there. There was argument. It's interesting, at Toronto, I guess…are you going to talk to Sean Eddy?

KM: Well, we hadn't planned on it. But one of my last questions was, is there anyone else you think we should talk to who maybe wasn't in Bermuda or that we wouldn't know to talk to?

RDurbin: Right, okay, well we'll get to that. So Sean Eddy, I think, has a really good and clear view of these things. It's a personal view, but I've been influenced by it. I think Toronto had more wishful thinking about how great prepublication data release all was and everybody should sign up to being open all the time. And I think it drifted away from realism a bit. And Sean's view was that actually the norm in some sense should be open data release at the time of publication and that prepublication data release should be exceptional and we should think through carefully what the criteria are for that. I think Toronto actually was trying to make the argument that all data should be released on generation or soon after. And I think Sean took a contrary position about that or a minority position, and I have sympathy for him, and I share that to some extent. But Fort Lauderdale was sort of tighter and was trying to establish a model that then you'd have to agree to enter into for any particular project. Or may be forced by your funder to enter into. But at least it was just trying to establish what the principles were. So I think Fort Lauderdale was quite valuable. Toronto brought in people from lots of different fields, and maybe it was sociologically valuable. It helped open people's eyes…and I don't know. Anyway, okay?

KM: Yes, yes. And just quickly, is there anyone else aside from Sean Eddy that you think that we should interview, that maybe we don't know?

RDurbin: Okay. So who are key people? Key people are John Sulston, Bob Waterston, Michael Morgan, and Francis Collins. I would say Mark Guyer and…who's the woman who's at NIH? Sorry, I've…

KM: Jane Peterson?

RDurbin: Jane Peterson, yeah. Jane Rogers is with us. Okay, so Maynard Olson with respect to Bermuda for sure.

KM: Yeah, so Maynard is actually a visiting professor at Duke this semester so we've interviewed him already and gotten to spend some time with him.

RDurbin: Phil Green, I would say. Is he on your list?

KM: Yes, we've already interviewed him.

RDurbin: LeDeana Hillier was there. I don't know how much she would add. She was kind of a more nuts and bolts person probably, but she's in Seattle, still connected to Bob and Rick Wilson. Of course, Rick Wilson. Okay, then…so Sean Eddy, I don't know if he was in Bermuda.

KM: He wasn't.

RDurbin: No. He had been at Sanger as a post-doc and he went to St. Louis, and he was a separate faculty member in St. Louis. He worked on kind of genomic things. He was later on, on NHGRI council. He was certainly involved in Toronto. I think he was involved in Fort Lauderdale. He's somebody with a perspective on this. But if you want to limit yourself to people who were attendees…I'm sure you've got the list of attendees, so…

KM: Yes, we do. Well this is really helpful. And do you by any chance have any documents or notes or anything that you'd be willing to share with us that you think might be helpful in understanding these meetings and what happened and their implications?

RDurbin: Yeah…so what was the date? It was like the 29th of February or something.

KM: Yeah, the 27th and 28th of February in 1996, and I think it was similar in 1998.

RDurbin: Okay, so I've just…I found an email from John. Let's see…it's from John to me. It's on the 22nd of February. So I think we should do this off the record because I actually think, since it's from John, we should check with John that he's happy.

KM: Yep, and so I'm going to go ahead and turn off the recording because we're at the end of the regular interview. So I'm going to stop this.

END OF RECORDING