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 _____

Printed Name John Craig Venter

Date 5/2/12

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.*

Archiving Permissions Form: "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

A short while ago, you participated in an interview with investigators engaged in a research project exploring the history and consequences of the Bermuda Principles for DNA sequence data sharing. We have prepared a transcript of your recorded interview. As indicated in the Informed Consent statement for this project, you now have the opportunity to review this transcript and make deletions and corrections.

Your transcript has been sent to you in either electronic format (via Dropbox.com or e-email communication) or hard copy format (via postal service). Please follow the instructions provided with your transcript when making any changes and when returning the document to us. These instructions are specific to the format in which you received your transcript. If you do not want to make any changes to the transcript, please let us know at the time you return this permission form to us.

In addition to the use of your interview materials in our research, we seek your permission (subject to any restrictions you impose) to place the edited, written transcript of your interview, and any related documents, on the Internet in institutionally affiliated, digital archives.

These archives may include:

- Archives affiliated with the **Institute for Genome Sciences & Policy**, Duke University.
- Archives affiliated with the **Duke University Libraries**.
- Archives affiliated with the **Genentech Center for the History of Molecular Biology and Biotechnology**, a part of the Cold Spring Harbor Laboratory (CSHL) Archives,[1] or
- Archives associated with the **Human Genome Archive** at Georgetown University.[2]

Members of the Duke University community, students, faculty and staff at other institutions, or members of the general public may access these digital archives for purposes unrelated to this research project on the Bermuda Principles. Typical research uses of interview materials include scholarly or other publications, visual presentations (i.e., powerpoint presentations), exhibits, class projects, or websites. However there may be other uses made as well, since the materials will be available to the general public. Investigative reporters and lawyers engaged in or contemplating litigation have, for example, used the Human Genome Archive at Georgetown.

Your permission to post the edited, written transcript of your interview, and any related documents, to a digital archive is completely voluntary. Unless you consent to their wider use, all materials from your interview will be available only to members of the research team affiliated with this project.

The form below provides you with different options for how, when, and with whom your interview materials will be shared. You also have the option, of course, not to share the materials beyond the Duke and Adelaide researchers. In the meantime, all digital materials are maintained in a secure, HIPPA-compliant drive at Duke University; paper materials are stored in a locked cabinet; and steps are being taken (i.e., via layers of electronic password protection of documents) to maintain the security of your materials during exchanges amongst the Bermuda research team and between researchers and interview subjects.

---

[1] The Genentech Center at Cold Spring Harbor Laboratories was established in 2006 with a gift of $2.5 million from Genentech, commemorating the 30th anniversary of the company's founding. The mission of the Genentech Center is to identify, acquire, preserve, promote, and provide centralized access to the original papers, correspondence, and research materials of the individuals and institutions that were crucial to the development of molecular biology and biotechnology.

[2] The Human Genome Archive at Georgetown University was established in 1988 under a grant from the National Science Foundation, and was long associated with the National Reference Center for Bioethics Literature and other international resources supported by the National Library of Medicine and other components of the National Institutes of Health.

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: _____ Date of interview: _____

Current institutional affiliation: _____

Street Address: _____

Phone: _____ Email address: _____

Interviewer Information.

Full name(s): _____

Affiliations(s): _____

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

(A) [X] 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: _____ Date: 05/28/13

BCD: … you were at the first two meetings, the ones in '96 and '97, but not at the one in '98. We talked to Mark a bit about the one in '98. And we're also trying to triangulate how much of the meeting in 1996 you were at, and in particular whether you were at that last session when Sulston and Waterston were up at the white board writing out what became known as the Principles. So why don't we start there. And actually, if you could just give us a two-minute summary of what you were doing at the time, why you were there, et cetera.

KM: Well really quickly for the tape…I'm so sorry, [CVenter]…just to confirm, we have the informed consent signed and we've done that and that's on record. So as long as you don't have any questions about that, then we're good to continue.

CVenter: Well in fact for most of this I've already gone I think on record with *A Life Decoded*. So I refer you to that for anything that I've forgotten since then. So I'm trying to remember…'96, I think we still had vague hopes from some of the…I think Mark Adams had one of the smaller genome center grants and we had been successful with developing whole genome shotgun sequencing and it was described in the *Science* paper in '95 that this could work at the whole human genome level. So we were still I think very engaged in trying to be part of the public program and to change minds and try to apply some new techniques and some new approaches. So I probably went to the meetings as an optimist then. I don't remember that session particularly. I know…I remember mostly stuff driven by Michael Morgan there, which may be not that meeting, so you can tell me if it was or wasn't.

RA: It might well have been. If you look at the agenda for the meeting a lot of what people remember, a lot of it was discussing what the various centers were doing and how jobs were going to be divvied up so there wasn't duplication, all that sort of thing. And in a sense the session at the end about the Bermuda Principles…or what came to be known as Bermuda Principles…was a pretty small overall portension to the meeting as it was. But I think our understanding so far is that at least that bit of it would not probably have been driven by Michael as such. A lot of the planning and a lot of the preliminary work was Michael really driving the organizational endeavors.

CVenter: My overall just remembrance of the meetings is being grossly disappointed in them to being really upset by them. My guess is it was sort of the steamroller going in one direction. I can remember …

RA: What did you go there expecting? Sorry to jump the script a little bit, but what did you go there thinking was going to happen, given that you were disappointed?

CVenter: I mean this was 20 years ago so…

RA: I know.

CVenter: …[Laughter] I barely remember going. I remember I liked Bermuda and it kind of destroyed Bermuda for me. It gave Bermuda a bad taste after…it's a great sailing destination, so Bermuda holds a special place for sailors. So to have to fly there and deal with bad politics really ruined Bermuda for me for a long time. I guess I was going in part for information. We were still trying to be players. Optimistically I was always hoping for a more open discussion about approaches versus the kings of the different fiefdoms dividing up the territory of the genome and dictating how it was going to be done. I don't know if that's what I thought at the time or if it's a retrospective analysis. I remember the…it depends on which part of the Bermuda Principles…there was the one error in 10,000 base pair one which I think has carried over in a very bad way. A lot of the groups did not live up to that, reported before…Landers' lab clearly had the best data that exceeded the one in 10,000. I don't think any of the other labs lived up to that standard. But I thought it was a bad standard at the time and it was a low standard. And that's what we're still trying to wrestle with out of getting sequencing where it can be used for diagnostics. But I think the most disturbing thing to come out of it is the notion of…I know this was driven by Michael Morgan because whether he did it at the meeting or at the bar, I don't remember…drove this thing of dumping data every night. Did that happen at that first meeting?

BCD: Yeah.

KM: Yeah, it did.

CVenter: Yeah. I was so dismayed by that notion. And he told me personally, and I just can't remember…we were still having drinks together at the bar in those stages, so…that the Wellcome Trust was forcing it because they didn't trust their own scientists to release data in a timely fashion. But I've objected to it, and I still object today even though so many people associated with it claim it as their badge of honor. It's just anti-intellectual and anti-scientific in my view. And I thought so at the time and was frustrated by it. [BCD], you'll remember because this goes back to early debates I had directly with Jim Watson. All he cared about was getting the sequence, and my argument was, if the people generating the sequence aren't part of the analysis then it's just a crude technical program and people who are wanting to get the human genome sequence to understand the human genome are going to walk away from it. And it was…he knows…a debate from the very first stages of the project. And Jim's answer was always, well we'll leave the analysis for future generations. In part that's clearly what happened. But I think both groups in the end in 2000 made valiant attempts to do an analysis. But still today all these different projects, it's a double rule. You're forced to dump data nightly, depending on who's funding things, but people in the community really aren't allowed to use that data and analyze it and publish papers on it. So I think it was a terrible policy then and it's become even worse policy today because it's the height of anti-intellectualism that's driven a lot of genomic research. It's just data generation and dumping versus analysis and understanding.

BCD: So, [CVenter], if you can, take yourself back to that time and what do you think was driving it? What was your analysis of what was driving this move to prepublication data release?

CVenter: Well Morgan told me directly that the experience is that scientists would hang onto the data for years and analyze it and reanalyze it and never release it. So they tried all kinds of models…there was a six-month rule, you might remember.

KM: The NIH rule.

CVenter: Yes. But it was a fuzzy rule because it was never clear when the six months started. So the six-month rule became very fungible and nobody ever enforced it. There were a few bad examples of people that hung onto their data for years and years until, like most rules in society, you come up with the most absurd thing. For example, everybody has to take off their shoes for going through airport screening because one person had a bomb in his shoe, right? So it taints everybody forever because there was one guy with a bomb in his shoe. So because there were a few people that hung onto their data for a long time, and nobody could find an intelligent way to work on the six-month rule, whether it was Michael Morgan individually deciding that, or the Wellcome Trust Board, or I don't know how their internal things worked, but it came out as him totally driving it, of anything the Wellcome Trust funded people had to release their data immediately because they couldn't come up with a more intelligent plan. And also I remember it was very clear, it was like the tail wagging the dog, that even though the Wellcome Trust put in about a billion dollars, basically they were dictating everything and NIH just sort of followed along with it. I think it was an historic mistake and one that sort of changed science ever since in genomics.

RA: And when you say that the rules of the game were that people then weren't actually able to analyze, or they weren't supposed to in some sense analyze each other's data, is that just because they weren't then geared up to working on that particular kind of analysis? Or was it there was really some sort of unwritten code that you just don't use the data as it's generated?

CVenter: Oh, no, these are new rules…these are new rules that have evolved out of this, so the genome…

RA: Originally.

CVenter: … well originally, it was kind of they didn't think that part through, if I recall. And then people started posting their own rules on their websites. So our data is publically available, but you can't really use it and analyze it and publish a paper on it. So it's not really publically available, right? It was trying to have the best of both worlds of, we're such great guys because we're releasing our data all the time, but we don't want you to use it and analyze it and write a paper about it

before we do. And those are the rules that keep getting battled out in *Science* and *Nature* today. And there were some, in the last few years, some historic examples of major complaints because people analyzed some data that was, quote, put out there every night. I wrote about this in my book. I think with the human genome effort part of the mantra became they were doing…their after the fact justification…they were doing it to block me from patenting the human genome. And the irony of this that everybody knows quite well now is that the pharmaceutical companies took the downloaded data every night, sorted through it and filed patent after patent on it. So it was just…things became mantras, sort of like early church mantras to help contain the masses.

RA: What might have worked better?

CVenter: What would have worked better?

RA: Yeah, what would you have advocated in general, or for your preferred sequencing strategy, to how the sequencing should proceed?

CVenter: Well I've proved what I thought would work better. When I couldn't convince anybody else, I was given the opportunity with money from Applied Biosystems to try my approach.

RA: That's a sequencing strategy. But I meant specifically in terms of data release. You had said the six-month policy was also not very useful.

CVenter: Yeah. So I think again what we did at Celera with the human genome was over a nine-month period we sequenced the genome, collected the data, assembled it, analyzed it and published a paper and released the data with it. If that had taken five years to do, that would have been different, but we did that in the fashion of every other scientific paper that I've ever been involved with and almost any breakthrough in science has been associated with people generating their data, analyzing it, publishing a paper and releasing the data. Crystallographers, protein sequencers, I mean can you imagine Fred Sanger just nightly, even though there was no internet, posting on a wall outside his office the sequence of insulin for other groups to go use and copy? I mean it sounds absurd, right? I mean that sort of…the analysis to me is a key part of science. Data generation is the first stage, then there's the verification, the interpretation. So I think the way science was working worked just fine. I mean maybe it wouldn't have worked for a public works program the way the public Human Genome Project worked out. I think the whole thing was flawed at so many different levels, there's one thing for that, but those rules are carried over as a mantra for everything in genomics.

RA: Well in everything. One of the reasons we got interested areas outside genomics, as you probably know, is that all sorts of other top fields in science are picking up the Bermuda Principles and just not very thoughtfully redeploying them for a

wide range of reasons, and that's part of what we got interested in. But anyway, that's a different issue. What occurs to me though is that it really isn't fulfilling what they take as public faith, particularly because of the huge, large scale public funding. Again I know I'm pushing your memory, but was that part of the dialog as you understood it? That this was different in time because it was sheer scale funding and a public investment?

CVenter: You know, I don't remember there being a lot of dialog. It was totally set up as a *fait accompli* and they were there just to present it. Apparently the power brokers had already agreed upon this, so it wasn't as though the Bermuda meeting was to discuss these things and come up with a plan. The Bermuda meeting was a way to get buy-in from the worker bees what had been decided to be the policy.

BCD: So, [CVenter], all these meetings took place in the TIGR phase, the pre-Celera phase. I mean a partial exception in the '98 meeting because apparently you guys had already…you and Mark had already had some discussions with the folks at ABI before that meeting took place because Mark reports saying he felt like he couldn't really say too much at the meeting because he didn't quite know what the status was. But it was clear that a big move was about to take place, though all three of these meetings took place in the TIGR phase. You're at a non-profit that is doing sequencing. You've demonstrated shotgun sequencing in the *H. flu*. And you're a player now in the shift to focusing on the human genome. What's the process at TIGR for generating sequence? How did you go about sharing it? Did it change how you did this when these rules came down? And then how did those rules interact with your association with HGS and then HGS doing all these deals with SmithKline Beecham and the other folks that were on their dance card? That's a lot of questions, but basically, what's the context of how TIGR is operating and sharing data at that point?

CVenter: So we did it the way we started with…it goes back to the EST. You might recall *Nature* did its first ever special issue with a massive release of data that we had of all the EST sequences based on a pretty comprehensive analysis. So we would collect data, we would analyze it and we would find a way to publish it and release it, all at once. With *Haemophilus*, the first genome, we did science the way that I think is the right way to do it. We generated the data, we assembled it, we corrected all the errors, we interpreted it and I think with the *Haemophilus* paper we set the standard for the entire field. And that was our goal—not just to say, here is a sequence—but we're the first ones on the planet to see the entire genome sequence of a species. If we can't say something intelligent about interpreting that species, this whole field's a fraud. And so we spent as much time analyzing it, then sorting out what the genes say and what they do and what it says about *Haemophilus* as a species. And every paper we've done after that followed that same pattern all the way through with *Drosophila* and human. The one exception was because Mark Adams was funded by NHGRI for the human

BAC sequence his team was generating, so he had no choice but to comply with dumping it, and there was no analysis done at that time.

BCD: So that was basically, that particular project was put on a special track for nightly sharing of data?

CVenter: Well if you want to use the word sharing…it was a requirement of NIH funding that the sequence data be nightly downloaded, okay? So Mark complied because that was a condition of that funding. And as a result it became a mechanical process of, you generate the sequence and you download it, versus you generate the sequence and you analyze it and interpret it and say something meaningful about the sequence. So if you look at *Haemophilus* and then you jump ahead to *Drosophila*, I think it shows that the same model can work on any size genome with the team that came together to analyze and interpret the *Drosophila* sequence. And once that comprehensive analysis was done as a community in a short period of time at Celera, with people coming in from all over to do it, putting it together in a massive analysis, and then releasing all the data all at once, I think it moved science forward. It didn't set anything back, in part because we did it all pretty quickly. I think the whole Cassava thing was done in less than a year from the handshake with Jerry Rubin when we were passing out CDs of the sequence and the paper coming out in *Science*. But I think we proved by our actions how I thought it should have been done.

BCD: So, [CVenter], back to this role of TIGR as a free-standing, non-profit, in your book you talk a lot about the struggles for control of data and what's going to get published when and who's going to control that process and all that. Did the Bermuda Rules provide you with a tool for sharing what data you wanted to get out? Or was this the standard publication framework or…did this become part of one of the norms that was used in your toolbox to free up some data when you wanted to do that?

CVenter: Yeah, I think you know the answer to that. And in the irony…all the irony of this is yes, because Bill Haseltine and the people at SmithKline-Beecham couldn't…we set up these rules that they were enforcing on me, of going in the opposite direction, as I could not release data until after I think it was at least a year. They had these strange definitions of confidentiality. And one of the loopholes that I used was if that same information was available from any other source it's not considered confidential. So the irony is the Merck funding of ESTs by Waterston, just as an economic warfare thing, instead of hurting me actually helped me tremendously because I could argue similar sequences are out there so TIGR is free to go ahead and publish them. So I did use it as a weapon. It's like in politics, the extremes of the left and the extremes of the right—sometimes you can use extremes to fight the other extremes. Neither policy was a good policy.

Jim and HGS would kind of hold up data and it was totally driven by fear. And they would clearly state their fear, which was stunning to me intellectually. They were afraid that when I published the data that somebody else might make a discovery that they didn't make. And I think they sort of verbalized what the fear was of other genome researchers when they didn't want to release the data because if somebody took your data and made a brilliant discovery and you totally missed it, I guess then you look stupid. But the amount of data, the number of ESTs and human gene sequences that were released…and by the way, just as a minor side point…if we had not published all that EST data neither Celera's nor the public human genome sequence would have been very interpretable. ESTs were the key to interpreting and finding the genes. But one group wanted to hold them up forever and one group is dumping them…Merck was paying Waterston to dump them to try and screw up SmithKline-Beecham's economic advantage. So there were a lot of different battles going on simultaneously. I made use of those to enable our scientific publications.

BCD: So, [RA], I actually think we've gotten through most of the questions that were really specific to you, [CVenter]. [KM], do you have anything else?

KM: Nope, you asked the TIGR question.

CVenter: By the way, on the TIGR question, we had massive funding from NIH. Certainly the major funding for TIGR and the [CVenter] Institute has always been from Fauci's Institute, Allergy and Infectious Disease. But massive funding from Heart, Lung and Blood as well. And none of the other NIH institutes adopted the Bermuda Principles. The DOE did not adopt the Bermuda Principles. We had no requirements on any of our funding which was, as you know, we sequenced hundreds of species genomes; other than the NIHGRI funding for the segments of the human genome that Mark Adams was sequencing, there were no requirements to do this. In fact most of the other institutes agreed with our approach because we showed that we could sequence genomes, analyze and get data out in a meaningful fashion. Not one NIH or DOE grant came with any of the same stipulations. So it's not like TIGR was doing anything out of defiance. We were continuing science as we were always doing it and how I've always done it before and since. It was only the NHGRI funding that came with that stipulation.

BCD: Were any of these other projects, where you're sequencing whole bugs basically for NIAID, involved with multiple groups that were trying to pool their sequence information where you're relying on a group of folks who are…where there does have to be an allocation of who's going to do what when?

CVenter: We apportioned out the analysis. I mean we were the major center. I mean before all this was going, you recall, TIGR was the largest sequencing center in the world before the massive scale of Celera and the Human Genome Project. And we could sequence faster, more accurately and cheaper than anyplace around. So

people would bring all their projects and collaborations to us and we apportioned out the analysis. I can't think of an incident other than the NSF project on *Arabidopsis*, where we sequenced I think one of the major chromosomes. I'd have to go back and check whether the *Arabidopsis* project required nightly dumping of data. I think it was a different set of rules and it'd be worth somebody following up on that just to check, it could have been the only …

BCD: Yeah, actually we should.

CVenter: But that was the only other multi-parsed project. And it was very frustrating at the time because it was very clear by then…well things are clear to me before they're clear to a lot of people…but it was clear that we could have just done whole genome shotgun and done it very quickly. But because they had all these different groups that wanted to be involved, again they chopped it up and parsed it out. And every program where they've done that the genome's taken five to ten times longer to get completed than it would have otherwise. I think yeast is the perfect bad example. I mean the yeast could have easily been the first genome completed. And they suffered the fate because they parsed it out and were trying to make it part of the economic development of Europe. As you know, they had to retract their first chromosome and resequence it because it was such a sloppy job. But they could have done…and that's what we argued…we could have just done yeast in not much longer a period than it took us to do *Haemophilus*. But analysis was different and I think the epitome of this was the *Drosophila* Annotation Jamboree that we invented where literally scientists from around the world, including Australia, came up and basically in a genome boot camp analyzed the *Drosophila* genome to the point that it was publication ready by the end of the jamboree.

KM: So one thing we haven't talked about too much directly is patenting. And you mentioned that the Wellcome Trust came out as the leader in the Bermuda Principles and that the NIH was just going along with what Michael Morgan and John Sulston and those folks wanted to do. To what extent do you think the NIH's going along with the Bermuda Principles without much argument amongst the folks in charge was a reaction to the EST patent applications and the *BRCA1* and 2 patenting controversies? Just as someone who is a little more of an outsider…

CVenter: It's hard to know because those things sort of became non-thinking mantras…I guess that's redundant anyway…became mantras for people and justifications of…you still see things written that they had to dump the data every night because the evil guy, [CVenter], was patenting the whole genome. And it was just…it was like really just a mantra that had nothing to do with any kind of reality. And look, [BCD] was there, Jim Watson started this because he felt that ESTs were a real threat to the funding for the human genome because it was sold to Congress as the goal is to get the human genes. And we were getting the human genes out very quickly and cheaply. In other words they undersold the value of getting the whole genome and tried to dumb it down and simplify it for the stupid Congressmen.

And therefore when the ESTs came out as a quick way for doing this, he had to find some way to attack it. Then he attacked it over patenting. Cold Spring Harbor has filed probably ten times more gene patents than TIGR ever has or TIGR plus Celera have. And so it just became…like any political campaign, people catch on to different mantras that they know work for people that can't think for themselves. I think it was much more of an aftereffect justification. When it came under intellectual attack that they're not doing real science, they're just dumping… they're just technicians generating data and dumping it, part of their mantra became…and it was used to rally the troops and rally more funding from Congress …Celera's patenting the whole genome, we have to stop them.

KM: Right, right. When realistically this is such a small subset of data that the Bermuda Principles apply to and in the grand scheme of…as you mentioned…all of the other patents being filed on DNA-related things, this was symbolic more than anything else.

CVenter: Yep, it was used as a rallying point, not a … I don't…and the data shows that their dumping of data sped up patenting of the genome.

KM: Right, right.

CVenter: Primarily by Biotech and pharmaceutical…I mean these…I was exposed to these executives because we had to come up with some kind of a business model to justify me spending $300 million, and the model was to license the data early on to pharmaceutical companies who would pay a fortune for it. And part of their argument was…I mean they would literally laugh at the program because they had whole team assisted download of data into the pharmaceutical company computers every night and go through it, then file patents on it. I mean they would just laugh at the program. And the real irony of what made our data so valuable is what I originally offered, and which historically has become such a horrible thing, was the mouse genome data. When I suggested Francis do the mouse it's because I knew the mouse was what would make the human data valuable, not the other way around. And it's because we had the mouse genome, together with the human genome, that the pharmaceutical companies paid up these tens of millions of dollar fees just to get onto a database.

BCD: [CVenter], on this point of patenting, one of…I'm still years later trying to figure out how this worked…TIGR is generating a lot of sequence, HGS has a right to look at it and then own the intellectual property that it deems valuable and then it can share that with its partners. And HGS is also generating some data. Are there any…can you think of any cases where the sequence data became really the basis for a patent? And what I mean by that is …

CVenter: The basis for a patent, I mean they…HGS filed the patent on the whole *Haemophilus* genome.

BCD: Oh, okay.

CVenter: I don't think it was ever issued, but I don't know. And this is something that was driven far more by SmithKline-Beecham than HGS. HGS was sort of just the broker. They would take the data from TIGR because that's for the $70 million in theory I was being given, that I think we only got $30 million of, and they turned around and sold all the EST data that we were generating and sold it to SmithKline for $150 million, and then went public with it. But they…in fact it was…one of my favorite quotes early on in the debate about the number of genes is, before the human genome was…you know, we were far enough into it and saw the gene density, I published a paper with somebody…I can't even remember who right now…estimating a much lower number of human genes, in the order of 30 to 50,000, and the big debate was because a lot of people, as you recall, remember, wanted there to be 100,000 or so genes and the quote from Bill Haseltine was, he knows there's 350,000 because he's filed patents on all of them. [Laughter]. Which sort of makes a joke out of all of it in the first place because there are only 22,000 genes and you filed 350,000 gene patents, you just have mostly noise that's not meaningful.

BCD: So did any of that work turn into significant products, to your knowledge? Can you trace from TIGR through the various partners to products and services that ended up starting from sequence?

CVenter: I think drugs that they proved, their lupus drug, I think everything they have originated from our ESTs. But that's one out of 20,000, or one out of 350,000, depending on how you're counting. So, yes, that was the absurdity of all of it in the debates, I was in this middle world between the left and the right, the left being you have to…it's anti-business approach of the U.K., you have to dump the data every night to destroy businesses and the right being SmithKline-Beecham and HGS, so you have to hold onto it forever. When the reality was I think everybody knew there was going to be maybe a couple of handfuls of genes that would end up being drugs, right? So because they weren't smart enough to know which ones, they wanted to tie them all up until they knew, which to me is as bad the mindless dumping it every night to, quote, destroy the value. So there's a lot of absurd thinking to find the needle in the haystack you go around buying up all the haystacks. I mean that's…I wouldn't describe that as an intellectual approach.

BCD: Well, [RA], [KM], in a final go, [CVenter], thank you so much for giving us some of your time. And what's going to happen now is we're going to turn this into a transcript. We'll send it your way along with a check sheet that says how can we use it. And the only document that will ever see light of day is after it's gone through that process. And it will be governed by whatever it is that you fill out on the form.

CVenter: What is your goal that you want to use this for?

BCD: Well as [RA] was describing, we were trying to figure out what happened at these meetings because the Bermuda Rules have become such a touchstone for public rhetoric about data sharing. And so we discovered we could not get access to the transcripts that were made of the meetings themselves so we've interviewed a whole bunch of people, so we're getting to the end of that …

CVenter: Why couldn't you get access?

BCD: The Wellcome Trust, not the Wellcome Library, owns them. And they're trying to negotiate a transfer from one to the other. But apparently the rules for what was said there were ambiguous and so they felt like they had to go back to everybody and get permission. And they've got this multi-player problem of getting permission and they can't release it until they've got permission from everybody because it's one master tape or something like that. We don't actually know the details.

RA: But there's some sort of…I mean you probably don't remember…but there's some what seems to be an urban legend that it was actually governed by Chatham House Rule, like a Gordon Conference and whatever else. But when you actually look at whatever was sent out there was no mention of that. So it's a little strange, but the Wellcome Trust seems to feel strongly that because people…at least some people…seem to think it was done by Chatham House Rule, they have to get everyone to release themselves from what was previously in agreement that was bound by Chatham House Rule. But then they also seem not that interested in getting people through…just to redact the stuff that they can't get permission for. So, yeah, a year and a half later and a trip to London to try and negotiate, we haven't made any progress.

CVenter: I don't recall any part of that, but as I said, it was not an open discussion; it was not a planning meeting to all the different parties come there and let's come up with a plan and go out with it. The Wellcome Trust and the NIH leadership came in with a plan and gave it to people.

BCD: But anyway, you asked what are we going to do with it. We're going to make as much of this material public for…because lots of people are interested in the Bermuda Principles, so we're creating a resource so these…the transcripts will be part of a public resource. And then [RA] and [KM] will be taking the lead on writing up some analysis. One will be a history of what happened and why it happened and why does it matter. And [KM] is actually looking at the interactions with various government policies, not just the U.S. and the U.K., but also Germany and Japan and France, where there were some complications in translating these Principles, these Bermuda Principles, into policies.

CVenter: Yes, there were. There was a lot of public blackmail going on, sorting through all that, when Germany and France particularly didn't want to go along with it.

BCD: That's what [KM] is writing about. So that's what we're going to be doing.

CVenter: Well good for you, guys.

BCD: Strictly a scholarly thing.

RA: And some reality to it instead of just lots of talk about what might have happened, because there's not a lot out there that actually has any basis, which you've noticed.

CVenter: Yes. Well you guys are free to use anything from this conversation. I've been getting so many…I hate editing transcripts, so hopefully you don't want an edited transcript; you're just going to show me that for the record.

[Several speaking]

BCD: Yeah, you can do whatever you want and then there's…actually the main thing that we're going to send you is a check sheet, and if you just check off the box that says, yes, you can share this, that's all we need.

CVenter: Great. Good luck.

BCD: Thank you so much, [CVenter].

RA: Thank you so much. It's been great, thank you.

KM: Yeah, thanks so much.

CVenter: Take care, guys. Bye.

RA: Bye now.

END OF RECORDING