My name is Jeffrey Schloss. I'm the director of the Division of Genome Sciences, which is part of the extramural program of NHGRI.

I was born in Chicago, Illinois in April of 1951.

I was -- my folks lived in Hyde Park when I was born, near the University of Chicago. Shortly thereafter, they moved to the south side and when I was four, we moved to the south suburbs, about 186 South, so it's 186 blocks south. It tells you exactly where it was, but town called Flossmoor.

It was actually -- I had -- we had a substitute teacher in fourth grade who brought in -- and I love this because it just -- it's really how these things happen. She brought in a microscope that could project the image onto a screen and she brought in paramecium, and I was just blown away. I was fascinated. It was the first time I'd seen anything -- really probably seen a microscope and seen living organisms under a microscope. And I can't say that I had any impression of school other -- before that, and actually, frankly, not much after that in elementary school, but I really thought that that was cool. And I would say that it just really made me interested in biology, but -- and to the extent that, you know, when you -- even when you were a kid. People ask you what do you want to do when you grow up? That was probably the first time that I had any kind of an idea. Now of course growing up where I did and in the environment I did, I didn't know anything about being a PhD. It turns out that one of my parent's friends was a chemist and was a PhD, but I probably didn't know -- I didn't know that until a good deal later. So of course if I was going to do anything related to biology, that meant you were going to be a doctor, right, because that's what Jewish mothers like to hear too so…

Undergrad at Case Western Reserve in Cleveland and grad school at Carnegie Mellon, which was not known for its biology, but they had a small biology department that they were sort of rebuilding. A lot of people had retired and they were rebuilding it, doing some new hires, so we had a few of the older faculty and a few of the newer -- a few newer faculty, including the person who I had done undergraduate reserves for in Cleveland. He'd actually moved to Carnegie Mellon. In between, he had a sabbatical at Cold Spring Harbor, so that was my first exposure to Cold Spring Harbor. I spent the first six months moving and setting up the lab from Cleveland to Pittsburgh, and the second six months of the year in between undergraduate and graduate school actually working at Cold Spring Harbor lab, which was very cool. Lived on the lab premises. We were -- we were a cell and biology lab. We were a lab that was looking at proteins that are responsible for motility in non-muscle cells and we set up one of the first microscopes that was used for -- scanning electron microscopes that was used for biology. They'd been used for metallurgy for a while and they'd hired this physicist to come in who actually understood the instrument and we set up a microscope down in the -- in the basement of the animal house, which is the one across the grass from the -- from the cafeteria.

I got my PhD formally in 79, in May. Actually was done after the fall term so I moved onto my postdoc in -- at the end of 78, but that was 78, 79. Got my PhD and moved onto a postdoc at Yale.

His name was Joel Rosenbaum.

He's known for his work -- mostly for his work in Chlamydomonas. He's done other things too, but much of -- about half the lab was working in Chlamydomonas and there was a -- there were other things going on with cells and culture, and actually went to do some mammalian cell culture work looking at some of the other fiber systems related to the motility proteins in some of the cells that they had growing there. But for about -- within about a year, I switched to Chlamydomonas and from looking -- from doing a lot of microscopy and some protein work to nucleic acids. That was when I started doing nucleic acids.

I was just interested in the project. We had very little exposure actually to nucleic acids research in the -- in the biology department I was in for my PhD, in part because it was a very small department. It's sort of historically oriented. So the lab I was in was almost entirely a microscopy lab and we were moved in next to a protein biochemist who worked in muscle proteins. So it was really a good opportunity. We sort of translated technologies and ways of thinking about the biological questions back and forth between those two labs. So I spent a lot of my time as an -- as a grad -- as a grad student trying to isolate protein complexes from non-muscle cells to try to get them in something of a native state and relate the state of the protein complexes to what the cells were doing that we could observe microscopically. So sort of a function molecular correlation type of thing, but we had very little nucleic acids where it can effect. We had -- as we were -- as they were hiring some people as -- in the last year or so I was there, people would come in and talk about their research using restriction enzymes. So that was my first introduction to restriction enzymes, which I really hadn't heard about and, you know, in every -- we all thought it was very funny they were talking about BamHI and BglII. This was one thing -- it was a little -- so it was very pragmatic, but it was a little discouraging and it's one thing that I've sort of pushed against in my career at NIH, was that we, at one point, went to some of our professors when I was a grad student and we said we'd really like to start a journal club to explore some of the rest of what's going on in biology research because the department is pretty limited. And they said that's great if you want to do that, but just make sure it doesn't detract from your focus on your research. And in one sense, they were absolutely right and in another sense, they were absolutely [laughs] wrong, particularly in a limited department. But -- so one of the -- that really -- I would say that one of the themes that I've tried to promote at NIH has been multidisciplinary interdisciplinary research.

We were a very accomplished light and electron microscopy lab. We had, I think, quite deep insights into what was going on in some structure function relationship with what you could look at in cells that a lot of the rest of the cell biology and even microscopy communities didn't because the way we used several different kinds of light microscopy with very careful electron microscopy. But it was all microscopy and it was sort of a way of approaching the problem. And then to bring in the protein work and the protein biochemistry and start doing protein functional assays along with the microscopy, that was probably -- that was a sort of first step we wanted to broaden even further with this journal club. And I actually don't think we did it.

'84, I left for University of Kentucky in Lexington, an assistant professor job at the University of Kentucky in Lexington in biology.

Well, first of all, I have to admit that I did have to look at the map and see exactly where was Lexington because I had been -- grown up in Chicago and moved to Cleveland and Pittsburgh and New Haven with a little side trip to Long Island. And so it was, you know, I had five or six interviews and it just seemed like the -- it seemed like a really vital department of -- there was a -- there were, again, a bunch of older faculty, but a good solid core of younger faculty doing molecular work in -- molecular in several kinds of molecules. We had nucleic acids, proteins, lipids, but it seemed like a very good group and seemed like a good job offer. And it wasn't -- well, at that point actually, it was far from family because by that point my family had all moved from Chicago to the west coast. So it wasn't close to family, but it seemed like a good move at the time and it was mostly -- it was largely -- it was largely a good experience. It was only a not a good experience to the extent that I didn't get tenure, but that turned out to be a good thing because then I met Jane Peterson at a cell and biology meeting while I was looking for jobs. And the rest is history. She had a, you know, three by five card up on a bulletin board looking for, you know, program director in genome -- the genome -- the genome center at NIH. I had -- we -- obviously we had postdocs and grad students in the labs that -- where I was and one of the postdocs from the lab in -- actually from the lab in Cleveland, I kept in touch with and, you know, we -- she knew I was looking for jobs. And she said, "You really ought to look at NIH. There's some really interesting jobs at NIH being a program director." And I had initially dismissed that, but then when I met Jane, it seemed like a reasonable opportunity. I came here within a couple of weeks I think after meeting her at the cell and biology meeting and interviewed. And they liked me and I liked them, so I moved here.

The other main parts of the program were sort of Betty was leading the sort of functional genomics types of things and I think it was clear at that point already that Elise Feingold would be coming in. She started formally a little bit after I did, but she'd already -- I think she already had some knowledge of NCHGR because she was in the grants associates program and moving from, I guess, a postdoc position to an extramural position. And so I think it was probably already formulating that, let's say, that she might work with Betty, but I'm not absolutely positive about that. And then we had the informatics and we had ELSI. It was pretty small.

Very vaguely I would say I was aware of the human genome project. It was, you know, being an assistant professor, I was pretty busy [laughs], you know, teaching a fair amount. You know, an assistant professor in a biology department teaches whole courses unlike med school, you know, where you teach a couple of classes here and there very often, not everyplace. But -- so I was teaching a couple of courses and had a couple students in the lab, undergraduates and a couple grad students, and I was -- I was the -- sort of the chair of the graduate recruiting that is -- screening at least. I did a fair amount of service even as an assistant professor, which may have been part of the problem [laughs]. But -- so I was pretty immersed in the -- in local activities. So I was -- as I said, I was fleetingly aware. I'd heard of it, but I mean I -- we -- I can't say that I remember any debates within our department, for example, about, you know, whether the federal funding agencies should be supporting this human genome project nonsense or something like that. We -- our battles were about very old school physiology and mycology versus more molecular -- modern molecular approaches and how the distribution of teaching responsibilities and other kinds of resources should be distributed within the department faculty meetings and so forth.

Well, I mean I didn't have any history with this organization to see things changing. Michael was -- I mean I'm sure he still had his lab. He was around some. I just remember him as very personable, just a real calm, pleasant guy. I think he thought that the project was in good hands with the leadership of Elke and, you know, and Mark, with Jane and Betty. So I don't remember him trying to -- I don't remember him trying to steer things very much. Clearly though, you know, there was a search going on, but I mean certainly he came -- I know he came to some meetings. I can't remember if he -- if it was weekly or something like that, but so I can't say that I remember him as a strong influence on directions we were taking. And again, I mean the grant -- I hadn't any contact with this community before that, right, so I don't know what -- if there was turmoil because of Watson leaving and things like that. I wouldn't have -- I can't say I noticed that very much. The other thing -- so in terms of the portfolio, I did not have a lot of grants. The grants that I was involved in with the centers branch was, you know, like, probably five or six at the beginning. You know, I was -- I was brand new at the whole business, right, so I really had a lot of ropes to learn. My first grant -- so I came -- my first day on the job was a Sunday, as it always is, and there was a pre-council meeting that night. And so I went to the pre-council meeting and I had, you know, had spoken to people a little bit beforehand, and there was some discussion of a -- of a Drosophila physical mapping rant application. I've probably mentioned this to you before, Jerry Rubin's physical mapping grant of Drosophila and that was the first grant that was assigned to me. So that was cool because I got to hear the council discussion of it. Obviously it was brought by Jane and then basically worked that up as a new -- as a new grant and see it through its, you know, through its paces. But I, you know, I worked -- one of the things you had asked me about was -- earlier was what -- whether I worked closely with Jane and I would say yeah, we worked -- we were constantly working together, but that as actually true of the whole group. We were highly interactive and so while I was working most closely with Jane because we were running that centers program, so -- and I really -- that was the first time I really thought much about -- my whole stunt was [laughs] people had specific numbers of markers they were expected to map per month, I think. And so I -- and then I had to learn all of it, you know, stuff about NIH grants from the grant from the NIH side. I have applied for grants. I'd had a grant, but that -- it's kind of different from the NIH side. And we -- yeah, just so I had a lot to learn. And, oh, I was -- yeah, I was saying I had just a few grants so this was a center, so it wasn't as if I had 20 or 30 all at once and therefore had a lot of contact with more typical NIH grantee community of academic investigators working on their individual projects. Instead it was this very sort of highly focused program where people were generally at a more advanced stage in their career. They had a pretty big grant for the -- from the centers program and they were grinding away to build these maps.

I would say she was a good guide and mentor. She gave me plenty of rope [laughs]. I don't think she let me hang myself, but she let me have independence. So I don't know why. I don't know if that meant I was doing a good job or if she was just purposely saying well, you know, go do it. It's the best way to learn. We never tried to come back and figure that out, talk about that, but I found her very easy to work with at that time. But as I say, when -- any time we were formulating anything, anything we were taking for, you know, every application that came through, we all talked about -- we all talked about. We all talked about all the applications that were coming through to NHGRI and we could do that because it wasn't very many program directors and there frankly weren't that many grants. There weren't that many people doing this kind of research at that time. The, you know, we had genome study section which was

formulated with a different perspective than most of CSR review and I think I did realize at the time that that was going on, but that it was quite different. It was focused on the goals of a specific project and so we interacted closely with the -- with Cheryl Corsaro who ran that. We also had our own review office and the -- and as far as I remember, the review officers, SRAs we called them at the time, and the program staff were pretty much together all the time. We almost always ate lunch together because there was plenty of stuff to talk about and, you know, we just met in the conference room. We easily all fit around a small conference table over in sixth floor of 38A and it was a very close group. And so I, you know, yes, I worked with Jane, but I felt that I worked with the entire group.

Well, I mean Mark brought the perspective of someone who -- I was conscious of that, of someone who'd been in industry and so he brought probably additional rich perspectives. And so he hadn't been at NIH for that long, I think, at that time. I don't actually remember him having been away that much so that's interesting. I remember him being much of a presence and he was excellent with crafting language, as he has continued to be. And so you could struggle with, you know, with sentences and then he would always fix them so they just -- they flowed better and they conveyed the meaning better. And when he was -- he would always ask you whether you were -- whether he was changing the meaning from -- away from what you were trying to get at, so he's very good about that. He didn't just say this is the way to do it. I guess that -- the main thing I -- the main impression I had over those -- all those years, it wasn't just that year, was that this was not a top down and you're a grunt because you're new so you're going to have to do this, this, and this thing and then bring it back, and then we'll view it and tell you what you did wrong. It was none of that. We just worked as a team.

Well, so I should say that Francis was one of the center grantees and we had -- as I said, we were, you know, the centers were supposed to be mapping certain numbers of markers and so forth, and we had two grantees who were not seeing it that way [laughs]. And Tom Caskey was one and Francis was the other, and they were more -- they were coming much more from a medical perspective than the other -- than the other projects that were probably more from -- I think more from molecular biologists-- as opposed to physicians. And so it was kind of funny to -- that here was this guy who, if you will, wasn't towing the line and he was the one who was recruited to -- of what the -- of what the program was not only trying to do, and he was recruited to be the director. So that was kind of interesting [laughs] but no, I mean he came in and he seemed like -- very much like a regular guy, but very smart and clearly very compelling. So in that way, I think we -- I didn't really know him before. His wasn't one of the centers that I was directly responsible for. That was Jane's and so I hadn't interacted with him a lot, but I think we could pretty quickly see that this was a guy who was going to be, you know, somebody who could go in front of congress and if he remembered to -- when he traveled here while he was still going back and forth to bring a good pair of dress shoes because he didn't always do that [laughs]. So when he had to go to a meeting, I seem to remember him having to borrow a pair of shoes from people because he hadn't brought a pair of dress shoes. I don't know if anyone has mentioned that [laughs].

So it was around -- so I started her in 92. Francis came in 93 and by -- it's by 96 because I know that I, you know, I helped write the -- with Jane, the RFA to preliminary to the scale up. It was sort of pilot projects toward large scale sequencing. So this was a -- some of the mapping goals

had been met, not all of them by any means, but the sequencing that was going on was sequencing in order to define STSs, sequence tag sites. It wasn't really the idea yet of starting to march through chromosomes, which was the way people originally thought you were going to have to sequence the genome, as you know. And so we wrote that RFA and then sometime -- and I'm pretty sure that we got the grants -- those grants started and I remember having been involved in setting up the first round robin because there were, you know, there were questions very early on as what's the sequence quality and how are we going to figure that out. And all the sequencing was based on bacs at that time so it was very easy to exchange templates between centers. Obviously you wanted to make sure that people weren't working on the same bacs so that we weren't duplicating effort, but -- and the bacs were mapped so you could say okay, we're going to work on this pool of backs and they're going to work on that pool of bacs. But to have the comparison, we set up a round robin of exchange of backs and then people were going to sequence them and they were -- there was supposed to be a comparison of the data. And I -- it was at that -- it was in that era [laughs] that I -- two of the -- there were two people here who had been running the technology development program. It was Bob Strausberg and Carol Dahl, and they -- there, you know, there were always questions about what you -- how much emphasis you should put on different parts of the program. That's normal in any kind of research program, and Francis wanted -- was very much we got to start sequencing. And the logical argument could be made that we don't have the technology to do that yet anywhere near the scale and we need a serious investment in sequencing technology. And so I wasn't part of those -- part of those discussions in any detail, but I -- it's pretty clear that Bob and Carol really wanted to try to push the technologies. They were much more interested in pushing on those technologies because they saw huge opportunities and they were in touch with the -- with the biotechs and the research community that could do that and they saw big opportunities. And Francis wasn't apparently willing to put as much investment as -- in that as they would've liked to see. And again, I'm -- I don't remember that we had very explicit discussions about some of this, but that's my perception of how this went and they moved onto another institute that was -- where they could have a bigger investment in technologies. I don't know if it -- I don't know extent to which it was sequencing technologies because I mean they moved to cancer and they -- cancer has supported all sorts of molecular technologies over the years. But we certainly still had the sequencing technology programs here and so it was -- they left and we needed to fill that slot. And basically Elke and Jane and Mark offered that to me. I could continue working with the centers or I could, you know, take on that different portfolio and that was really interesting and attractive to me. I liked working with Jane, but this was a chance to be a little bit more independent and so I just moved. I switched my portfolio and that was when we hired Adam Felsenfeld. So he actually implemented that round robin as I remember this.

So I mean, you know, this is always kind of rolling evolution. There was -- maybe I was already in the business of these slab gel sequencers. There were -- with the fluorescence detection. The throughput and just the number of samples you could put on those was very limited. The, you know, sample -- the lanes were wavy [laughs]. There were all kinds of problems with them because they were still individual people pouring slab gels. So the -- I inherited an RFA that was on electrophoretic sequencing to improve that basic methodology and I just look back at some of the things we issued shortly thereafter. So this is in 97, was this low cost, high accuracy DNA sequencing technologies RFA, in parallel with a more generic just genomic technologies program announcement where that was -- that program announcement, all those applications,

were just out to the community and just, you know, send us your interesting genomic methods, including, as you read the new -- this 97 RFA, including additional improvements on those electrophoretic methods. But what we were looking for here is what else is out there? What are their -- are there some other ways to sequence DNA other than the chemistries and the separation methods? And so I looked to see what some of the things that we funded under this and, you know, there was still -- we didn't want to do slab gels. We frankly didn't want to do capillaries because that was already coming onto the scene, capillary electrophoresis, but -- so these were trying to integrate other kinds of detection methods, electrochemical detection methods, into the capillaries or figuring out way -- or using mass spectrometry. So you'd rid of the gels altogether or integrating sample prep because these were all still very separate processes. You had, you know, you had to grow up your bacteria [laughs] and pick your colonies and -- with toothpicks generally. At that time, it was starting to move toward robotic pickers and, you know, arm robots to pick and place arm robots, like were used in automotive industry or to assemble integrated circuit chips, to assemble sequencing reactions, you know, using little pipettors or things. But this was -- so we had an application -- couple of applications there to try to integrate the sample prep with the separation step and with a different way to detect the -- to detect the separated ants. So, you know, these were the kinds of things that were being funded at that time. There was nanopore. We were -- we had -- that wasn't the first time we'd funded nanopore, but there was a nanopore application and another idea using exonucleases, where somehow you would get the exonuclease to chop off a nucleotide at a time and the products of that would flow out onto a surface, get attached to a surface, and then be detected in -- with some chemical means on that surface. And then you'd move your -- presumably your microscope, whatever, along the surface and read off what basis. So this was -- we were -- we were definitely trying to push it that way to different ways to sequence already in 97. So trying to address this question of sequencing with Sanger reactions and electrophoresis, even capillary electrophoresis, was still expensive and relatively low throughput. Obviously the advent a couple of years later of the higher throughput capillary -- of the commercial capillary systems where you could actually do 96 at a time and not worry about lanes waving into each other and so forth, and do automatic -- automated loading as opposed to the contraptions that people were building, the pneumatic contraptions people were building in order to load 48 or 36 or 48 or 96 whatever samples onto a slab gel that -- where the slab gel was less than a millimeter in between two glass plates. I mean that was -- that was all very hand labor so.

So I mean -- and this was one of the issues, is that, you know, it's how long are you willing -- if you want to do something really different, how long are you willing to invest in it? And presumably will come to that sometime later in this discussion [laughs] -- in the next hour or so, but if you want to change the paradigm, sometimes, you know, the methods were completely different. It was no longer molecular biology. Everything else was really based on molecular biology and on -- I mean if you were -- sure, capillary separations were, you know, that was known technology. It just wasn't -- hadn't been used for single nucleotide resolution DNA separations, but mass spec was well established methods. It just hadn't been used in this application. There was electrochemical detection of other kinds of molecules, but the question was could you bring some of those things together and apply them to this particular problem where there was clear -- if you could do it, you could conceivably have an impact. And then there was the question could you -- could you do it in the academic setting and then could that be engineered to a commercial product that would be robust and cost effective?

So I mean I -- the biotech sector, I would say, is -- sometimes innovates to the level of actually a completely new idea, but there aren't that many completely new ideas, you know. They may -- they may put together two ideas that are reasonably well established, but nobody's ever thought of putting them together. That would be a really important thing for biotech to do, but some of this, you know, these -- some of these methods, if it's never been tried for that particular kind of molecule, those are things that if you can incentivize them, you can get academics to think about. It's not to say that biotech wasn't thinking about it too, but -- and we supporting things in the [unintelligible] on our program too. Now this happened to be R01s and R21s and things like -- and those kind of grant mechanisms. And at that time, I think this was not -- probably not calling for small business grants. But -- because back in those days, we could combine -- yeah, this was R21s, R01s, and program projects. We could combine mechanisms. We can't do that anymore so the -- we can do it, just separate program announcements, but there are a lot of innovative people in academia too who are more into this kind of methods developed in different communities. It's not generally the biology community. So my initial grantees were mostly analytical chemists. You get a few people who were physicists or bioengineers, but in those days this was more from the analytical chemistry community that knew these methods for other purposes and could start developing them. It's a little hard for me to know. You know, I've heard stories from different investigators about who did exactly what, but I mean, it's pretty clear that the capillary electrophoresis that ABI was commercialized was in partnership with Hitachi, and I don't know who did that original research leading to that, whether that was Japanese investigators or American investigators, but I know there were investigators in the United States who were working on that same problem. And this is another just general principle is, you know, we've said in some of our program announcements, our RFAs and so forth that we want to fund whoever is the best place to do this kind of research. And so NHGRI has funded a lot more work sort of R01 research in company settings than most institutes do because for the tasks we thought were important to get done the commercial setting was actually a really good one, but academia can also innovate in those areas. And I feel what we did with the academic grants -- there was no real possibility -- there was certainly no real possibility of direct commercialization because this was stuff going on in academic labs. Now some of them may have thought they could have put the plans up on their website, which is what happened with some of THE, you know, spotting for microarrays in the early days, but, you know, these methods were going to have to get commercialized. We're going to have to get, you know, the licenses were going to have to get -- the license was going to have to get issued from the universities to companies, BUT that's fine. What we were trying to do is sort of feed early pipeline. NHGRI really generally can't afford to fund the commercialization of these instruments -- of these systems, I should say. But anything that we could do in academia, where there's a lot of expertise, to add knowledge about the components and whether that was some of the dyes. There was a lot of dye research going on. Part of it was done by Jing Zhu, who, you know, had gone on to Columbia and done a lot more of that in Rich Matthews' lab at Berkley. You know, they developed some of the original energy transfer dyes. It was done in academia, and I can't tell you whether ABI was working on that in parallel. You know, this kind of information gets around, but I do believe, firmly believe that information that came out of the academic labs helped the companies. So there were some experiments maybe they didn't have to do because they saw the results published from academia, and they could say, "Okay, we've learned from that, and now we can go do our own thing." So maybe the exact path to their commercial product wasn't funded by

NIH, but what was learned from the academic labs. And that's true, for example, of these dyes, a lot of the polymers that are used in the capillary systems, certain software, probably not a lot about the fluidics because that was a different kind of engineering. So I'm completely convinced that it's appropriate for NIH to fund that kind of research to feed that pipeline of technologies. The other thing that it absolutely does is it educates the students, who then go get hired. You know, those students were getting snapped up, the students from the labs that were getting funded to do these projects were getting snapped up by those companies because even if they weren't going to bring their specific project and, you know, their dye that they had worked on, they knew how to do this kind of research.

Yes, it is. Yes, absolutely. So that was -- what was somewhat different about that was that it was a bigger lab, and he had sort of a different approach to how he was going to do this. But he really hired on a wide variety of expertise, and it was an environment unlike many others where you could do this kind of research where you really brought together biologists who understood -- biologists, biochemists, surface chemists, you know, electrical engineers, mechanical engineers. Because if you were trying to develop really novel devices, you couldn't do it piece meal. You had to have all of that together to be able to tackle some of the problems if it was going to have any chance of getting to an end point. That's not to say he did all of these projects perfectly, but the basic approach of having all that expertise in one place and where the people could switch around to work on different projects, and you had an environment where people were willing to try stuff and fail and then they would still be a way to support them to do another project that took advantage of their expertise as well as what they learned from their failures on the other projects. So you had some long term support, and a lot of innovations came out of that group. Often by somebody going there because they knew that the expertise existed there doesn't mean that all the ideas necessarily originated with Ron Davis' group, though many did. And there's been a lot of biotech spin out from Ron's lab because of the environment and the kinds of projects and ways of thinking about what's needed to advance biology and medicine by groups like that.

Oh, that's so loaded. I mean, I think we had a general feeling that it was, you know, it was pretty annoying. It wasn't like he was thinking of something brand new. He was just going to do it faster, better, cheaper, which has clear impact, but, you know, it wasn't like "Oh, I'm going to sequence human genome." That was already going on. There were two main innovations, one of which I think originated with his group, which was the idea that you didn’t have to start with maps by using paradigms, and then not only an idea but a demonstration of the value you could get from doing the molecular biology and following it up with the right kind of software. And he hired really good people to do both of those. And then his affiliation with ABI to really jumpstart the use of the high thermal capillary array sequencers. And so the impact -- those were all really good insights, important insights into what it might take to do this better, faster at least, and reasonable quality, and probably, you know, so there were people within the public Human Genome Project who were sort of frustrated, I think, about the degree of emphasis on quality that if we could sacrifice some of that we could go faster. But there were very strong voices, strong and well respected voices. about the imperative to get right, at least to a certain level. And so, I mean, I think the most important thing he did was light a fire under the public effort. It did -- this is another place with Francis was quite effective in selling this to get more money for the public effort as well as the way he was able to galvanize international collaborators to try to

improve efficiency and reduce redundancy and, you know, buy into the public data release, because this was a very different thing than what Craig Venter was going to do, you know, and so it brought more money, brought perhaps, call it, better collaboration cooperation, and say, okay, we can accelerate this even with pretty high standards. There was a big push to improve the throughput of mapping using basically old fashioned methods but automation. I mean, the automation was a big deal in terms of the applying automation to biology experiments. I don't know if it had ever been done before, but it was certainly done for the Human Genome Project. So it allowed better control and it allowed probably use of smaller volumes and, you know, better tracking of what you had done so you could see where errors might be, you could fine tune and take variability out, and so forth. So there was automation throughout. I don't think Craig particularly contributed to that part, but that was, you know, that was an ongoing thing in the centers that we were supporting and I'm sure at Sanger and so forth. So, I mean, the lighting of a fire was probably good. The role that Ari Patrios played was a very interesting one to try to bring the two -- to try to keep the whole thing from exploding to the extent possible, to bring the two together, at least for the wrap up. Craig is an extremely smart guy. His system was so -- is so different, more of the entrepreneur and startup type. We have some of those in the public sector, too. They manage to come across a little bit differently, though. That doesn't mean that they're any less entrepreneurial and have any less vision, but they operate differently, and, may I say, piss people off but not as much as Craig does. He seems to thrive on it in my opinion, where I don't think that would surprise anybody, whereas I think some other people will do it, thought they may not thrive on it as much, pissing people off. Well, he also was unabashed about reinventing facts. So, I mean, there were things that, you know, that we knew inside of NIH that he would explain them very differently. I'm not going to go into details, but it's probably old enough that we could, and it's probably out there, but I'm not going to be the one -- I can just say there were -- I know of at least one very firm example where he explained things in a very different way from what the facts were. And so I wouldn't be surprised that he did that in other circumstances. But he's not the other only who does that, and I'm not only referring to current politics. So I'm talking about genomic researchers, so. But, I mean, I have to give him credit. He's innovated in a lot of different areas of biology. I think he's often over claimed in what he's achieved, but he's very innovative, and he's thinking always a few steps ahead of what could be important, and I think that's great. So.

Well, I mean, the Genome Project was -- although it wasn't emphasized at the beginning, it was still about variation. It was about human variation, right? And so, you know, the only variation we got was in the back overlaps, when the BACs happened to be from different human libraries. And so, you know, you could -- I'm sure there were people who thought, "Well, we can just get that with microarrays." I'm certainly wrong about this, but it's always seemed to me that if you could get all the kinds of information you wanted from one assay it would be better. Now that could be too high a bar, and I completely appreciate that. So, you know, people would sequence and they would run C and B chips and they would run, you know, single nucleotide variant chips, and, you know, you could do all these different chips, and each of those things is -- in aggregate could be much cheaper than if you just decided you had to do very, very high quality sequencing of everything. But if you could get a single technology to give you all of those answers at once then, first of all, you know, just the sample handling, you could do it one, you know, do it eight times or whatever the number of technologies you have to apply the data analysis should be simpler. And if you ever want to make this routine like for medical use you

probably don't want to be running all that different number of assays just for the sequence. It's not that you're not going to run, you know, blood chemistries for all the other things for patients, but you don't want to be running, I don't think, you don't want to be running a whole battery of tests just to get the DNA sequence right. So that I think was in the back of my mind from a very early day. We certainly knew, actually, in this '97 RFA we talk about, you know, there are going to be lots of other things that need to be sequenced for which we're not going to have all the resources like the mapped BAC libraries. And so there need to be other ways to do sequencing that can be cost effective that can be applied very widely across agriculture and microbes. I don't think we were talking micro biomes at that time, but certainly microbes, agricultural important organisms and model organisms for biology research. So, I mean, that actually -- I was surprised when I look back at this today in preparation for talking with you about some of the statements that were in here. It was all there. This was nothing new when we got to the planning for after the Human Genome Project and had that huge series of workshops, as you know one of which was sequencing and re-sequencing the biome, and that was how it was framed. There was a lot of other sequencing to do including human variation sequencing, and we got to have ways to do this more cost effectively and accurately and efficiently and quickly, you know. So there were all the ways to make it hard: cheaper, faster, more accurate.

Right. Well, I mean, the quality standard were originally set in part for STSs, right? Because if you didn't have an accuracy so that you could design a primer where you needed it, a pair of primers where you needed it to span a region, and that was a, you know, it's an applicon. In this case it was for a mapping STS, but you could do the same thing if you had a region where you were looking for variants that were related to a disease. If you couldn’t design -- confidently design a primer than your accuracy was too low, right? So it had to be -- accuracy had to be pretty high. But I mean, I think there were costs of the -- I'm going to come back to some things to talk about. There were costs of the acceleration that was prompted by the Celera-- by the Venter and TIGR and Celera, which was that I think we could have gotten a better quality sequence the first time around. I might have taken a little longer and it might have been more expensive, but we've done a lot of backtracking I think because of the race. Now you could say that the opportunity cost and just get it done once and then let the whole world work on those ways to improve the quality, and this is a philosophical argument, and -- you could argue it either way, but I think we would have gotten a better quality sequence. We didn't get the small INDELs, and we did a bad job on the structural variants, on the larger scale structural variants and copy number of variants because of the way the sequencing went to shorter reads, and this was a big discussion as we can get a lot more data faster if we sacrifice read lengths, if we tune the way we use the capillary sequencers to more runs per day, but that means we use shorter read length. But that meant we lost a lot of kinds of data that I think we could have gotten earlier. And so now we've had to backtrack and we're still struggling because with the next gen technologies got way worse in terms of being able to collect these kinds of data. And so it's fantastic to be able to sequence a genome for approaching a thousand dollars, but you're not getting the information you need. You can do a bunch of molecular biology -- we're jumping way ahead, obviously. You can do a bunch of really carefully thought out and beautifully implemented molecular biology using 10x technology now and some of the others like that dovetail, but you're still -- it's surrogate. But we don't have the real thing yet. We're approaching the real thing with some of the single molecule methods, which much, much longer reads. It remains to be seen whether those will ever be able to be implemented at the through put of cost

that we really do need, which is, of course I would say need more research. But at least we've seen the vision of what you can get with even low quality -- relatively low quality single reads. So people are complaining now: "Oh, well, you can do back bio." Yeah, but that cost 20,000 or 25,000 dollars to sequence a genome. Well, it depends on what you're comparing it to. Are you comparing it to, you know, a 1500 dollar sequence on Illumina, which still is pretty much without some of those add-ons to do the longer range assemblies or are you comparing it to 10 million dollars in the Human Genome Project? Fifty million dollars, you know, for high quality draft of Human Genome Project sequence? It's come down unbelievably even if it's 20,000 dollars. But, again, my soapbox. But the other thing to note there's been a much bigger investment in PacBio commercialization both by PacBio and now by Roche, and it's come a long way, but it's still not where it really needs to be for the ways people would like to be able to use that sequence, but it is single molecule. And there's information you can get from that that you can't get from the ensemble measurements, including some of the -- directly getting -- better assemblies, but directly getting the methylation. And that's not all worked yet, but it's, again, many steps that you don't have to do. You don't get methylation for 1500 dollars. You have to do bisulfate and stuff like that. So I think vision is still worth pursuing. That doesn't mean you stop and don't do the research using the methods we have today. It doesn't mean you stop developing those other methods. But that kind of refinement, I think, can actually go on in biotech more so. Again, many of the ideas, the concepts behind dovetail in 10x came out of academia, right? Some of them were funded in our organized programs and some of them were just funded by R01s that the professors applied for, and that's good. That's why we try to have both opportunities. But a lot of that can be done with ideas that emerged -- the original idea as well as the commercialization emerged in biotechs. But if you want to develop a completely brand new way to approach the problem I still think that we need both, but I think there has to be a source of funds for academics to do that.

It can't all be on a three year timeline. It can't be all on a three year timeline. It also can't be in an environment where you can define from the beginning how it's going to be commercially successful, which I think biotechs and companies really do have to worry more about. So part of the challenge of rolling out these technologies is each time you have a new commercial technology the bar keeps getting higher. So the bar for PacBio got higher over the time PacBio was being developed because of the other next gens that were coming out, and the fantastic engineering that got done in these companies. I mean, I'm not knocking it at all. These brilliant people working really hard, and they've done a fantastic job. But this was a different vision, but the bar kept getting higher, and then similarly for nanopore. If you were going to have a biotech come up with the idea of nanopore sequencing today and what was going to be their path to achieve the scientific results and the commercialization success to displace other technologies I don't think that's going to fly in biotech. And so you need other ways to do it. I don't know if it can be -- there are a lot of contests coming up now and that's one way to fund things is that whoever wins gets the money to take it the next step. Somehow they have to get the money to develop that initial prototype, but I think -- I've used the term here, it's an ecosystem, and you need lots of members of the ecosystem and you need evolution, and its fine. Some things will die off and some things that could have been really good will die off, but it's because they didn't fit the new niche.

One of the things that I want to mention so that either you were going to get to it or not, but it's the question of so, you know, there was 454 and SOLiD and Illumina and what did we fund? Did we fund any of that? Did we matter? Did NHGRI's investment matter? Well, there was some development of sequencing by synthesis. For years before any of it got commercialized, and it was just sort of going along sort of slowly and people were struggling with it in a number of the other projects that we were funding. And so, you know, we were aware of the concept. It wasn't really taking off. It ended up being hard because the chemistry is quite difficult. Actually, 454 was the first one to get there could avoid a lot of the chemistry because it was pyrosequencing, but that had its own limitations as we saw by sort of the demise of that platform, but it really was the first one out there, and they had a market, and they contributed a lot, particularly to micro biome science. People could do with that platform that was much too expensive with other or impossible with other platforms. But we did fund -- well, clearly as you know we funded PacBio from pretty much the moment they were a spin out from the Cornell lab as a small business grant up in New York, and then they moved to California. We continued funding them for a while, and then they came and got some large grants in our bigger programs. But as we were formulating what would become commonly known as the 100,000 dollar genome and the thousand dollar genome, we were already being approached by some of these entities, and we funded a program project grant to 454. And, you know, we'd been funding their predecessor Kerogen for awhile before that. So there were all these other companies that we were funded before, Anthropometrics and Kerogen, and yeah, in the earlier days of the Genome Project, even before I was here. So early '90s, maybe late '80s, but certainly early '90s. And this was all, again, it was just part of this bubbling pot, I think is perhaps a way to say it, is that we were stimulating that there was an industry there, which is the other important thing is that there's going to be a market for some of this stuff, which is important if you're going to have any kind of venture investment. So Kerogen was the predecessor to 454, and that was all Jonathon. So we funded a program project grant, and what that was they already had their first platform was going to be rolled out. This was before it was rolled out, and actually they were -- I think at the time they were a little bit nervous about even submitting the application because people didn't know a lot of the details about how that system was going to work. But what we funded was the potential for scale up because they couldn't sequence genomes on that original platform, which their business plan was that earlier platform. And I'm sure -- I know because I've been told over and over again by some of the people working with them that they were able to try any number of variables in their engineering because of the grant, and even if those particular variables, those paths didn't work out for them, it helped them to zero in on the ones that ultimately did work for improving that platform. So those are the kinds of things that NHGRI supported is it accelerated what they -- it accelerated their path toward their next platform in large part by giving them more flexibility to try things and fail, which is okay. If we accelerated the development of the platform and also stimulated the field. We stimulated competition in the field. I'm completely convinced of that. We definitely funded work in George Church's lab, which is where the ligation -- the oligo ligation method of sequencing was originally developed and then picked up a spin off from [unintelligible], Agencourt. We funded Agencourt, and we gave them additional money to -- it was a supplement to accelerate that development, and that became the SOLiD technology when they were bought up by ABI. So we certainly funded early work on that. It's impossible to know -- any time these things are bought up by a bigger company they can bring -- they always bring a reengineering team, and they may reengineer that platform, by that I include the instruments, the software, and the chemistry that go in into it, they

may redo everything, but that doesn't mean that the investment to show that it was even possible to do this thing wasn't worthwhile. And we had much less direct funding of the Illumina technologies. We funded Illumina early on when they were a bee hybridization company. In fact, very early on when people didn't believe that they could decode the bees with oligo hybridization. So helped that get off the ground. Subsequently we funded a bunch of grants, mostly through SBIR, to help them with their development of their [unintelligible] synthesis. We didn't fund much of their development of the sequencing technology, though we funded a project that spun out of Ron Davis' lab. Again, I think you know a lot of this, but I'm saying it so we get it on this history video. A product that started in Ron Davis' lab that was being run in part by Mustfa Nagi [spelled phonetically] to develop a sort of a mega scale pyrosequencer. And this was one of those situations where Ron felt he didn't have enough money in his NIH grant and couldn't put together money from other grants, so he spun it out as a Vantone [spelled phonetically], as a small company. And Vantone applied for a grant from us, and before we were able to process the grant all the way through they were bought by Illumina, and that became the MiSeq. Not directly, but a lot of the fundamental technologies, some of the detection technologies I think and some of the engineering, because MiSeq isn't pyrosequencing, right, but a lot of the technology concepts contributed to the development of the MiSeq. And I know that because that was in their press releases. So I'm not over claiming. I'm always very careful about over claiming. The other thing is we funded any number of companies to develop sequencing technologies that never really developed sequencing technologies, and we knew that this would be the case. Sequencing is a really high bar. It's really hard to do, but there's a lot of other nucleic acid analysis that you need to be able to do for mapping, or not for genome sequencing but for forensics, for sample prep. And I have at least one or two companies in mind for each of those. And actually some of these are on my slide setup when I gave my council talk. These were initially funded as sequencing technologies. They weren’t quite up to snuff to do sequencing, but they ended up getting spun out to other applications and commercialized. So, again, I think the funding, if there's a question of should the federal government, which we seem to get frequently, should the federal government be investing in this kind of technology development, I think it's a no brainer, and I think we have a lot to show for it.

I tried to avoid calling it to the 100,000-dollar genome and the thousand-dollar genome. But I failed because it just got too cumbersome. You know, we were talking about two orders of magnitude and four orders of magnitude because it all depended on what your starting point was, both for what your dollar starting point was, but also or what kind of sequencing you meant. And so, you know, it's different to do sequencing when genome versus re-sequencing. And we said whatever you’re doing we're looking for a two order of magnitude cost reduction in five years, and then a four order cost reduction. And it did become more convenient, and people would -- it was a faster way to communicate to say 100,000-dollar or the thousand dollar genome. But there's nothing magical about the thousand-dollar genome, right? It's just a nice round number. I learned, actually, at the sequencing program planning workshop, I may have told you this, in 2014 of an earlier enunciation of it, it's not the right word, expression of the concept of the thousand-dollar genome happened at a workshop that Bob Waterston [spelled phonetically] organized at UC Santa Cruz. So I don't know -- I didn't find exactly what the date was of this, and people who were there who told me about this were Bob Waterston and Dave Housler [spelled phonetically], and they actually used that phrase at that workshop, and apparently it's in some sort of report from that workshop. I mean, I first heard it I think in our

planning for the first bookend -- in our planning for the first bookended meeting. Bookend meetings being the two large planning meetings that early on either side of series of maybe 12 or 14, 15 topical workshops, one of which was sequencing and re-sequencing the micro biome. But so that was like 2001, and I don't remember if I heard it in the planning or it was at the workshop. I know I heard it at the workshop. We had some breakout groups for technology development, and it was being used. Thousand dollar genome was being used there. I don't know who first brought it up there. You know, different people have different stories of this. Dave Schwartz [spelled phonetically], who has been a grantee of ours for a long time, who was developing a mapping -- who developed a mapping technology, said that he think he threw that number out just to something completely ridiculous, just saying, you know, somebody said, "What should it cost?" And he said, "Oh, you know, a thousand dollars," so everybody could laugh.

The way -- one of the ways I heard it expressed, and expressed it to press at the time when we started the program was that this is a cost that's reasonable in a medical setting. There are imaging tests that cost a thousand dollars that get done on people all the time, and so you could envision using sequencing as a practical medical test. Though it's not a cheap one, it's not a screening test. But you know, I think anybody who talks about it today would say that it's a great, and important milestone to achieve, but it's not where we ultimately need to be, and again it comes back to something we talked about later is it's only as good as the quality that you get for that cost. So I would love to see $50 sequencing, but it actually includes long range -- all the long range phasing information that you'd like to see. As well as methylation, and maybe a few other things. You know, whether you could actually incorporate chromosome-folding -- chromatin-folding into that, I don't know. I think we could always continue striving for that. Because we're not only going to need to sequence people once. That was -- it was sold that way early on, but it's just not true for any number of reasons. Particularly if you're looking at epigenetics because you -- epigenetics is probably going to be a very effective readout of environmental effects. Perhaps a better readout than having a lot of environmental sensors that people will wear because that's actually so what's the outcome of all the environmental exposures. What's it done to your genome? And if you need that information you're going to have to sequence people over, and over, and over again. So, you know, even a thousand -- calling it comparable to an imaging test you could do for $1,000 isn't satisfying. So, I don't know what people actually were thinking at the time. They thought it was a very bold goal, and if you could even remotely approach it, it would revolutionize the industry. Which of course it's done, but without the bold goal as, you know, now you got your archive of Francis Croke quotes about, you know, the architects. You know, make no modest plans. That's not the word that was used in it. So I think it was a great idea. I think it was a wonderful idea. We didn't originate it here, and in each year I -- it was, you know, it came from the community which is great because it's actually better for these things to come from the community, but I think it was a good one for us to grab onto. So I just wanted to make sure we got that in.

So we were a very small institute in the -- and this is mid-nineties when, with our roughly eight program directors, and I had formed this bio-engineering consortium. This was largely, I think, due to pressure from the radiology community, and also somewhat from the bio-engineering community that felt it wasn't getting -- they were having a hard time getting grants from NIH, getting support from NIH in part because NIH was so focused on a hypothesis-driven research,

and these people were trying to develop devices and instrumentation that were critical for medical care, could be critical for medical care, but they couldn't tap into the pot of money. this growing pot of money. So in part, at least in part, perhaps wholly in response to that pressure the NIH director's office -- and this was during Varmus' directorship formed the bio-engineering consortium. And it was led by the deeder [spelled phonetically] the Deputy Director for extramural research who at the time was Wendy Baldwin. And they needed -- they asked for a representative from each of the institutes for this engineering and that stuff, and well, you know, I was a cell-biologist by training, I am a cell-biologist by training, but I was the closest thing we had because I was leading the technology development program at the time. Shortly after I -- I think that BCON was formed in '96 or '97 shortly after I took on that role. But I was the closest thing we had so I started attending those meetings, and I learned so much from colleagues whose -- some of whose portfolios really were about bio-engineering, most of them were not. Most of them were people like me who had some part of their portfolio that could be related to developing new methods. But through the ways of thinking at -- we developed some program announcements, since you have to have a product if you have a Transonite panel, right, and so we developed some program announcements that allowed people to explicitly come in with design-driven grant applications, as opposed to hypothesis-driven grant applications, and we stressed several of them, bringing together a biologist with a physician with an engineer. So it couldn't be just some whacky -- I'm being purposely a little bit provocative, not facetious -- whacky engineering ideas. There had to be some solid connection to a biologist or clinician who would want it, right, and so the development of the engineering idea would go hand-in-hand with the use of it. And so that was some of where some of the ideas that led to segs [spelled phonetically] came from, was that interdisciplinary. Also the idea is that -- as we had with the centers program, you had some small, you know, RO1 grants but you also had opportunities for people to put together teams that you couldn't sustain or support on an RO1 grant. Also the idea actually that you should be able to design a project that was not going to go on forever because if we were going to do larger investments in order to support these larger, more complicated teams, you didn't want to have to fund that for the next 30 years, 40 years for a person's whole career. Say "Design a project where you're going to have an outcome. " And tell us where you're going to get to in five years, and tell us where you're going to get to in 10 years because that's it. That's the limit, and probably some fairly heavy component of engineering in there with the biology. So all these things were formulating at kind of the same time. It was through BCON that I got assigned to be one of the NIH representatives to the National Nanotechnology Initiative. This was percolating largely out of NSF and DOE and probably I think DOD. Maybe a little bit of MIST at that time, I'm not sure how important MIST was in the very earliest days that there was this new sort of science emerging that was very hard to get enough funding for, and I think people just thought that this was just some -- this was a really good hook: nanotechnology, and the idea of control of matter at the single atomic level which was, again, at that time something of a dream. There had been some very rudimentary demonstrations that you could laboriously, with an atomic force microscope, move molecules into a pattern that spelled out IBM, or something like that. Then there was some more philosophical, but physics-based books coming out at the time, and famous discussions about the room at the bottom, and so forth, and there was enough to build on there to build ideas. So they formed this thing, and they needed some representatives form NIH, and some of my -- it turns out I didn't know what was going on in the background -- some of my colleagues said at BCON, they thought I would be a good participant in that. I don't know why, but they did. And I wasn't -- it was so -- there was much more

experienced bio-engineer from Drexel -- who had been at Drexel, and he actually went back to Drexel a few years later -- who was working at one of the institutes, and I'm sorry I don't remember which one. Dove Jeran [spelled phonetically] was his name. I can't believe I remembered that. And he was sort of the lead on that, but I was part of that team, and when he left I became more of the lead. So it was all serendipity. The serendipity of my getting to take on the technology development portfolio here, though that. Getting involved at BCON, through that. Getting involved in nano, and us thinking about what we're going to do after the human genome project, and you know the idea from the segs came from Maynard, as you know, you need to come back to original research ideas. Something that's not top-down organized. So we formulated sort of an envelope for it, but the actual research that people would do under segs was completely up to them, not up to us. We, you know, we built an envelope they had to work with and of the two five-year slugs of money, and interdisciplinary, and things like that. Because all of the successes of the human genome project up until that time completely depended on biologists, and biochemists, and engineers, and informaticians, and, you know, if you were going to make any of these fancy systems work, you know, you had to have surface chemists because you were trying to work with very small volumes, and so you had to have some needle or something that would be able to collect the sample off the bottom, so your surface in the tube had to be just right, so the sample would go down to the bottom without a centrifuge. You know, it -- so that was a lesson that I was actually bringing in part to things I did in BCON, and then subsequently to segs from the human genome project. So they all -- it was really interesting time because they all fed off of each other.