Maynard Olson

So, I'm Maynard Olson. I'm Professor Emeritus of Medicine and Genome Sciences at the University of Washington. And the Emeritus part is for real. I actually did retire a few years ago and spend very little time at the University of Washington and am involved in a variety of other activities, but not any longer active in research.

And I grew up in Bethesda, on Edgemoor Lane, yes. I -- my father was an intramural NIH researcher in the -- what was then called the Division of Infectious Diseases. And at that time responsibilities had not sorted out between the CDC and the NIH to the degree that they have now, and most of my father's work was more similar to what goes on now with the CDC than what goes on at the -- in the NIAID. But, in any event, I kind of grew up in this culture and left -- I was born here and left my senior year in high school but -- and -- to the west coast where I've spent a lot of my time since then. But, I went through the public school in Bethesda. Bethesda Elementary, the no longer extent Leland Junior High school and BCC.

I can remember -- I can remember the NIH campus when I was a young child. Still rented some of the land to a neighboring farmer to graze cows on because the initial land gift here was larger than the NIH could put to scientific use at that time, and that was when the NIH, particularly to the north, was surrounded by farm land. So, things have changed.

I was always attracted to the -- to the sort of basic physical sciences and probably would have become a physicist but particularly at Cal Tech, you know, I quickly became aware that many of my peers had more facility with mathematical physics than I did. I'd like to think that I can actually understand these things but it takes me longer than it took them to. And I thought chemistry would be good compromise. It -- it's a basic science. It -- chemists are proud of their field as being what is sometimes called the central science. That whether you're a physicist, or a material scientist or environmental scientist or a biologist you've really got to deal with chemistry. We live in a chemical world, we're chemical entities. So it was a good choice. I stuck with it through graduate school. Got my Ph.D. in Inorganic Chemistry, is really physical inorganic chemistry, studied the mechanism of small molecular reactions, all the while ignoring biology completely. It's impossible to exaggerate my level of ignorance about biology. It is still true to this day that the only biology course of any description that I've ever taken was at BCC in 10th grade. Mr. Butterfield's high school biology course, and I didn't take any biology in college, not even any biochemistry, and less so in graduate school where I was really on the, sort of, chemical physics. My first paper -- one thing that Francis Collins [spelled phonetically], the current NIH director, and I have in common -- a little known fact -- is that for both of us our first serious scientific paper was published in the "Journal of Chemical Physics." It's not too well known, even here on campus, that Francis did his -- he was an MD-Ph.D., but the Ph.D. part of it was in Quantum Mechanics. And so, just a long tradition of chemists migrating into biology, and Francis migrated by his root. He knew a lot more biology, of course, when he did his migration because he had an MD. I knew nothing, but that's always been my style. I prosper in fields where I don't know very much going in because some -- something about the way I learn things requires that I start from scratch. I don't -- I don't do well in fields where you have to start with a lot of highly structured pre-existing knowledge. So just for example in chemistry, I was never good at Organic Chemistry. Organic chemistry is a fascinating subject, but the people that are good at it develop a tremendous amount

of highly structured, kind of, pre-existing knowledge and then they attack a new problem from that reference point. I've never been -- never been good at that -- I like to stay closer to things that I've just had to sort of slowly figure out for myself. That's just a characteristic.

At that time at post-docs were actually not much of a feature of chemical education landscape. They existed but they were relatively few of them and they were not an obligatory step, but -- so right out of graduate school I took a -- a real job. A tenure-track job at Dartmouth College in the Chemistry Department. And with the -- still the intent of doing sort of physical inorganic chemistry research and teaching and spent a few years there, but fairly quickly discovered this really was not a good long-term plan. I like the teaching, but sort of needed -- needed a fresh research topic and so molecular biology was an obvious one to look at, it was a very exciting period in molecular biology. The -- the period after the discovery of the double helix and kind of the working out of the genetic code and the mechanism protein synthesis and so forth -- that had stabilized, but there was a big sort of what next kind of question hovering over molecular biology at that time that even an outsider such as myself could -- could grasp and was strongly attracted to. That -- there were, of course, a lot of molecular biologists that were eager to get on with studying the more, what were then, the traditional topics of, you know, basically transcription, translation, control of transcription and so forth in evermore mechanistic detail. That didn't appeal to me. I had actually done a lot of mechanistic studies in chemistry, mechanisms of reactions and that was part of what I wanted to get away from. My view was that these highly reductionist approaches are extremely affective in their early phase, and then fairly quickly get into a phase of diminishing returns. So I was not interested in that asymptotic sort of phase. I wanted to get in on a more ground floor of something and the -- so with a rather vague ideas of this sort in mind I negotiated an early sabbatical from Dartmouth. I did some extra summer teaching and so forth, I really wasn't eligible for a sabbatical but took one anyway. And it's a complicated and sort of personal story as to how I ended up in Seattle, but my sabbatical was at the University of Washington. We will discover, if we keep on this biographical track that, you know, my relationship with the University of Washington is less straight forward than it might seem. People like to give formal introductions of me because I was at the University of Washington. I went, then, to Washington University and then back to the University of Washington and then, of course, was born in Washington, D.C. So, in any event this was my first phase [laughs] at the University of Washington was on sabbatical from Dartmouth, working in the laboratory of Ben Hall who's a highly distinguished molecular biologist who shared a few things with me that made us a really good match. His Ph.D. was in chemistry, biological chemistry, to be sure. He worked on nucleic acids but from a very chemical point of view. Paul Dotey [spelled phonetically], he -- and he likes to do new things, if you look at Ben Hall's CV it's a series of new things. He has an attention span of about five years and then he moves on to something really new. And I was a bit like that, although I like to think my attention span is longer, maybe 10-15 years. But I also like to move on to things that are new. So what was kind of new for us was, you know, somehow -- this is sounds absurdly vague idea, but it's pretty much the way we talked about it at the time, is that they -- molecular genetics of at least simple eukaryotes -- we're working on yeast saccharomyces -- was actually very well-developed by 1974, when I showed up in Ben's lab. Quite an impressive intellectual and technical edifice doing Mendelian analysis in yeast, unlike bacteria where most of the early molecular biology was done. You know, yeast has Mendelian genetics. In fact, its Mendelian genetic size often been described as sort of so good it looks like it was designed by a geneticist. It -- it's an immensely better organism than peas or people to do Mendelian genetics

on for various, kind of technical reasons. And this was all quite well developed. There was no molecular genetics in yeast. A little biochemical -- what we'd call biochemical genetics, sort of protein level. And a few interesting fusions between the Mendelian genetics and the biochemical genetics, but nothing that we would currently really think of as molecular genetics. But 1974 was the year that the first recombinant DNA papers were published and they were much on Gund's [spelled phonetically] mind and actually mine. I read about them in The New York Times. I have a somewhat famous first front page paper in The New York Times about the development of recombinant DNA methods. Cohen [spelled phonetically] and Boyer's [spelled phonetically] work and -- it's mostly famous because an intellectual property lawyer at Stanford read this article and told the University that, gee they should look into patentability of this method. Which, neither Boyer nor Cohen had considered and they were right at the -- kind of the deadline for doing that but managed to get a patent. I read about it and said, wow, this sounds like you're going to be able to get at genes as sort of chemical entities. And it was actually, you know, a lot of nice work in bacteria that had been done kind of getting at the genes, but it was all done biologically by, you know, these other esoteric methods like transduction and, you know, with prophages that excised imperfectly and various things. Just the vocabulary tells you that it was a very biological approach. Cohen and Boyer or I could understand this, you know. You get these molecules and you manipulate them, that's what chemists do, and then put them into cells and -- and there was a generality to it all that greatly appealed to me. And also it fit the criterion that I outlined earlier that nobody knew anything about these methods. They -- Ben Hall and it was -- there was -- was really interested in getting them going in his lab but didn't have them going. The only person in Seattle who had any experience with them was Dan Falko [spelled phonetically] who is -- was, even then and certainly now, sort of distinguished student of the molecular genetics of bacterial pathogens and that was key because, of course, it was bacterial technology and he was much ahead of anyone in the genetics department where I was. And so we learned some things from him about doing recombinant DNA, but -- anyway, there was no knowledge base. And our goals were vague. We wanted to get the Mendelian genetics sort of together with the -- with the recombinant DNA methods and that -- I was good for this -- to take on this project because, you know, the biologists in Ben's lab and the genetics department -- you know, they all had this functional orientation. And so they were always looking for some project that would teach them something functionally about biology. I was ideal because -- although, of course, I understood that function. I understood then and I understand now, that function is really the ultimate goal in biology. I looked at what people were actually doing to learn about function and the methods all seemed to me just so hopelessly underpowered. I was coming from chemistry, where we actually had some tools [laughs]. And, you know, we had NMR and we had very sophisticated spectroscopy. We had -- we had stable isotopes and radio isotopes and so forth. We had a lot of tools and we could actually learn something about the functions of molecules. Looking around the genetics department, I loved the environment. The people were, you know, they're very smart. Lee Hartwell [spelled phonetically] was right at the -- kind of the peak of his, you know, getting his -- his approach to studying the cell cycle off the ground. You know, these were -- he was another Cal Tech graduate. I hit it off early on with him. He was a young faculty member there. I stayed in touch with him for -- you know -- throughout both of our careers. But they -- I loved the people, I liked the environment. But I look at what they were doing, they had toothpicks and petri plates and they were sort of streaking these colonies out and using velveteen. I don't know, this just wasn't going to get us very -- we weren't going to learn very much about functional biology by these methods. So, I wasn't concerned about getting a project that would get me to study function yet. I wanted

to get these methods kind of, sort of, build -- I wanted to work on methods because I thought that's what -- that's what this field needed was immensely more powerful methods. And recombinant DNA looked like the foothold. So that was kind of my post-doc, was working on that and it went well, actually. And I think a lot of it was our -- our point of view. And also another thing about chemists is that chemical experimentation is a very -- like, because of all these methods and so forth, they're much more sophisticated business. Chemists are not easily discouraged by experimental problems. In fact, chemistry was way past the state even then and is even much farther past it now when you can hope to discover anything interesting in chemistry without developing new methods. You know, the methods had been -- you know, the methods that were very well understood had been applied exhaustively and if you really wanted to learn something new about molecules you needed some new way of studying them. So I was quite comfortable with this whole notion and, you know, I think a key to Ben and my approach is that we were kind of scrappy experimentalists and didn't give up easily. And there were a lot of experimental problems, but the -- it did go -- it did go well. We really combined, I don't know how much detail it's worth going into, but we -- I think we combined Mendelian genetics and molecular genetics in some reasonably new ways and didn't learn anything functionally. I think I left Ben's lab not having learned one new fact about the yeast cells that -- you know, how they do things. But we knew a lot more about the genome and the genes. How they were organized, how -- and more than anything else, kind of how to get at them than we knew five years earlier when I had first come there. So anyway, it went well right from the beginning, so I quit my job at Dartmouth and went -- went back to -- went back to Ben's lab. I took a few -- I had to take a few months to settle my affairs at Dartmouth and then I got back, I basically spent five pretty continuous years there. And it was during that period that I collaborated with Ron Davis. I was actually only in Ron Davis' lab for one week. Ron and I still sort of joke about this that the -- but Ben had arranged this collaboration, I didn't know anything about working Ron's lab, but Ben recognized that he really was at the cutting edge of recombinant DNA techniques then and that expertise didn't exist in Seattle and they -- he arranged anyway. Ron hosted me. He had an extremely talented graduate student, Jon Camron [spelled phonetically], who was -- I worked with, you know, kind of at the lab bench. Jon interestingly later went into clinical medicine -- psychiatry, I think -- and did not pursue a molecular biological research career. But he was -- to this day -- is one of the best experimentalists I've ever seen at the lab bench. And so in a week I learned how to take lambda clone libraries and screen them and brought that back to Seattle and that was one key part of our success, was having some of -- one of Ron's many contributions to molecular biology was that -- you know, he also came from a chemistry department, just as an aside. In fact, the Cal Tech chemistry department. But he also simply believed that these technical problems that were ubiquitous in molecular genetics at that time not only needed more work, but they needed good solutions. I think that's the one thing I learned from Ron -- don't settle, you know, for a halfway solution that barely gets you by your immediate need. If the -- if the problem is a -- you know -- is a fairly fundamental one, technically speaking, solve it well because then that solution will have a whole variety of applications, not just to your project, but that's how the field really advances. The attitude -- this attitude served me extremely well and I really did get this from Ron. The tradition in molecular biology -- I mean, I was hardly -- only person that discovered that it was hard to do these experiments and make them work -- but the tradition was a scraping by tradition. That you get this stuff just to work well enough to do your project, kind of, and I got much more of this attitude from Ron that you should get these methods working well and then when you move

to the next step of complexity get those methods working well. That you can't really build on methods if you're just scraping by.

I mean, I'm a, you know, huge fan of Ron's. I had this early contact with him, but I've admired his work ever since he -- you know, a few years ago I was pleased that I was on the selection committee for the Gruber Prize in genetics, which is kind of a major prize and I was able to play a significant role in persuading them to award this to Ron because I think he's under -- under-appreciated. Yeah, his contributions are so ubiquitous and so manifold that it's hard to know where to start. I remember in the -- in the Gruber -- in the Gruber discussions that Jasper Rine [spelled phonetically], a star post-doc of Ron's wrote a letter of support and I was joking with Jasper about this letter that -- I think that I'm not revealing anything here that's meaningfully confidential, I think all parties here would be happy to know about this story -- that he violated every known rule in writing this letter. For -- we all write a lot of letters like this -- and he violated every known rule because it was about five pages long and it rambled from this to the next, to the next [laughs], to the next. But that's what was needed. I mean, if you go through Ron's contributions, you know, it is many things. And if I -- my best unifying theme is the one I've already used is that he chose -- he was largely methodological, which actually hard to associate Ron's long career with learning anything about how cells work. I'm sure he would come up with some examples, but they're not what stick. It was -- he recognized the difference between kind of a protocol and a method that -- a method is a, you know, some new source of leverage that can be used to address a very wide range of problems. So, the -- I think the early on, he really was the one who got recombinant DNA working on scale. You know, that it wasn't a questions of getting, you know, a few hundred clones and hoping that yours was one of them -- the one you were looking for was one of them. It was a question of you making libraries from complex genomes that had 100 X coverage so that even things that were way under-represented would still be there, and methods of screening these libraries effectively in the, you know, in a day or two, not through some long complicated thing. That was an early phase. But he always understood much better than I did when I went to his lab. It was another thing, I remember discussions when I was there that influenced me. We didn't do these experiments, but how -- he had a much more biological view of genes than I did. I had a chemist's view of genes. You know, these are polymers and -- yeah, they had interesting functional effects, but that was -- I was above my pay grade. Ron understood that genes were, you know, the interest in genes, they were boring polymers. They were -- they were interesting because of their functional effects. And that you had to study these functional effects by getting genes into cells and that was a major theme of -- sort of middles phase of his career, was many different ways of taking genes from anywhere and getting them into cells, usually predominantly yeast cells, and the real strength of his -- of the -- all of the applications of yeast transformation that he developed is that he wanted to have tight genetic control over the genes once they were there. You know, you can do a complemation [sic] -- complementation test just by getting the DNA in there and as long as it gets expressed fairly quickly. Transient assays and various things dominated the literature for a long time. Ron wanted to be able to make a chromosome out of it; he wanted to put it in a chromosome, he wanted to be able to replace the, you know, the gene copy that was there to begin with, exactly, with no other alterations. He wanted to be able to have them that high copy number. Not all at once, of course, these are different goals. And so he gradually developed building on work. There was a lot of good work being done in yeast at that time and centromeres and telomeres replication origins, I think Ron's main contribution was to the replication origin story. But, you know, you need to know all these functional parts of a chromosome to have this

kind of control and that became important later in my career. I think it was, obviously, a direct influence on -- on building YACs because they were yeast artificial chromosomes and they differed, actually, from the kind of thing Ron Davis had been doing much earlier only in that they involved large, large segments of exogenous DNA as opposed to the few thousand base pairs that were the, kind of, 1970s and early '80s recombinant DNA technology. So anyway, he's -- and he's gone on and done still more things. I think there's not an -- any kind of direct lineage between Ron and, you know, say the next generation -- as it's called for some odd reason -- sequencing. But there are many indirect links because he understood that DNA sequencing was too hard and needed to be done much better, and he explored, kind of, many approaches to that problem and trained a lot of people that played a key role in developing next-gen sequencing and is -- I think of him as kind of the Godfather [laughs] of next generation sequencing, he's not the father of it. He's sort of a Godfather. But he had a big influence on me.

Yeah, one of the characteristics of Ron's methods is that they work. And that's not a shock and awe issue because it, you know, the shock and awe is all about somebody got the thing to work once under some, you know, tenuous -- this scraping by, kind of environment, and that's when people pay attention because it's new. You know, Dolly or something is, you know -- was not a -- not a -- not a protocol anyone else could follow and took many years of work to make it possible, even to clone sheep with any regularity, much less mice and so forth. But if you're working within a field, there's a big difference between the scrape by phase and really having a good experimental control over the system you're working with.

Yeah, so that project was kind of a natural outgrowth of what I did in Ben's lab. I didn't work on the project in Ben's lab, but by the end, I was thinking about it. Ben used to affectionately refer to it as my megalomaniac project, but -- and I don't think that was a compliment, actually but I -- you know, we had all the raw materials to integrate the genetic and physical maps, but didn't know how to do it on scale. Basically this project I did on tyrS and tRNA genes in Ben's lab involved integration of the global genetic map because it was -- had been built by then at -- globally with toothpicks and Petri plates. But the -- and we did physical map correlations using this cloning technology, like from Ron's lab, did local correlations of chunks of the physical map with chunks of the genetic map or loci on the genetic map, but it was, you know, quite clear that the next step would be to have a complete physical map of yeast at sort of gene size resolution. Genes and yeast are not very big, and so it, you know, had to have a resolution of, say a few KB, at least, and yeast has a 15 million base pair of genome, and so you're talking about, you know, mapping thousands or perhaps 10,000 sites and restriction maps of those days, you know, they had ten sites on them. Not 10,000. The methods just wouldn't scale, they weren't very reliable, actually. Even those 10 sites because of the methods that they had used -- just good enough, that was his -- so I had begun thinking pretty formally about this problem. Especially the last year I was in Ben's lab and by then I finally, you know, could market myself as a geneticist despite the fact that I still hadn't taken any biology courses. And so I was looking for jobs and found a very good one at Washington University in St. Louis where they were building a brand new genetics department and they -- you know, I don't know, we had a kind of shared view of where genetics needed to go with -- I think the zero baseness was very important. I was not an attractive candidate to the best well-established genetics departments because I -- they could always find somebody better qualified than I was to do any particular thing. But we had this kind of zero-based kind of idea. You know, Bob Waterston was on the search committee that hired me at Wash U. Any event, as I was making that

transition while I was still in Ben's lab I started thinking about this problem of how to -- how you actually would build a physical map across the whole genome. I could picture pretty well how to do correlation with the genetics map, but it wasn't so obvious how to build the physical map and I -- but I had the basic idea by the time I left Ben's lab. One -- the major influence on me during that year was Kim Naysmith [spelled phonetically], who's gone on to have a fairly distinguished career doing functional molecular genetics. Kim wanted to understand how cells do what they do. You know, chromosome synapse and so forth, whatever mechanisms and he has excelled at that. But he was interested in this idea, building a physical map and was really the person I talked to the most about it. And the -- so the method I decided on was generalization of what Kim and I in particular in Ben's lab were already doing and I think anybody who was doing recombinant DNA experiments sort of well, meaning in the Ron Davis style, did what we did. Is that -- because in the Ron Davis style, you didn't just get one clone and then go on and study it, you know, you go 20 or 50. And the way you avoided artifacts was that they had to build -- you had to be able to build a self-consistent physical map locally out of these clones. And if you couldn't do that, then that was a good criterion for throwing out spurious false positives from the screen and so forth. And so you would, you know, you typically would be probing for essentially a point on the chromosome, with a hybridization probe and you would get a bunch of clones that hybridized or -- Kim was starting to do this functionally, so he was selecting for bi-complementation in a yeast recombinant DNA library going into yeast. But it's the same idea, is that he was insisting that some particular point or modest number of base pairs the -- in all the clones. But the libraries -- we -- Ron Davis always -- he made libraries with random end points, so it was part of the Ron Davis dogma, so in Ben's lab we made libraries with random end points. There are a lot of reasons why that's a good idea. So, you would get out a bunch of clones that all had some point in them, but they had random end points, then you'd cut them with restriction enzymes and run them on a gel. And, you know, they would share various numbers of restriction fragments depending on how much they overlapped, they would share the restriction fragments in the overlap. So we had done a lot of that and that seemed, to me, a method that would generalize to the whole genome, just skip the probing step. Make a Ron Davis quality kind of library and just start picking clones. And the numbers were not that intimidating, at least for yeast. Yeast has a 15 -- roughly 15 million base pair genome. And the -- you could get -- [unintelligible] with -- in lambda clones, which we were using at the time that were 15,000 base pairs. So one X coverage of these genome would just be 1,000 clones. And so you could get that many on one Petri plate. And, you know, [unintelligible] considerations suggested that even if allowing for some non-randomness in the sampling that 10,000 clones or something should be enough. Well, yeah that's a lot looked at one way, but it -- the appealing thing about it was that you're just -- there's a linearity to the experimental work. If you needed twice as many clones, it was only twice as many work -- much work, or somewhat less than that because of economies of scale. But the number of pairwise combinations that you could consider at the data analysis step -- what is the square of the number of clones? And so obviously I planned to do that in a computer. So this seemed like the ideal approach, is that we make the experimental work linear and then make the map building the N-squared process, so if we were going to do 10,000 clones, then we'd do 100 million comparisons and -- and that was not computationally daunting, yet it wasn't entirely trivial [laughs]. The computers of 1980 -- but it wasn't daunting, you know, I knew enough about computational complexities. I'd been using computers when I was at Stanford in the chemistry department and used them quite a bit at Dartmouth where Jon Kimmeny [spelled phonetically] had been a leader in really getting, you know, time shared computing terminals out there where students and faculty could use them. You

know, instead of having a computer at Stanford, I used to walk out through the Eucalyptus trees to some, you know, air conditioned building that was away from the center of things where they had a huge computer and type out punch cards and put them into the machine and leave them for an overnight batch run and so forth. At Dartmouth I got this idea. Sort of -- Kimmeny, you know, Kimmeny was sort of the Ron Davis of computing. I -- you know -- I don't think he -- I don't think there was any fundamental contribution, really, to, you know, to computer science. He was a mathematician, but he understood that computers should be used. They should be integrated into everything. He was a -- sort of a Johnny Appleseed of this and he became president of Dartmouth and quite prominent and Dartmouth really pioneered, they developed the basic language at Dartmouth. Yeah, so it -- you know, it's not a modern language, but the name is important. He wanted a language that, you know, wouldn't have a steep learning curve. He wanted Dartmouth students to learn it and then figure out what to do with it themselves. Not, you know, not have some highly structured assignments, but get it involved in their work. That -- a lot of that rubbed off on me when I was on the faculty there. And so actually I didn't use computers at all in Ben's lab. We didn't have any and I didn't use them at all. But I knew that we could -- we could do -- we could do the N-squared comparisons and I was pretty confident that we could do the N kind of complexity experimental work. So of course I hugely underestimated how hard it would be, but you have to or you don't start projects if you think they're going to be as hard as they turn out to be. The amazing thing is I actually -- I said, I got a grant. It was -- you know -- there was no -- there was no genome kind of -- the term genomics had not yet been invented. There were no -- you know -- there was no place to send these grants. There were no reviewers that were qualified to review them. There weren't any genome grants. I'm quite confident that if you went through the vast archives at the NIH that you would find this was the first genome grant to come in the door. And to the -- that anybody would recognize, this is a genome grant [laughs].

The amazing thing is it was funded. You know, they -- you know, there's the argument, you know well you know, they don't really fund things that are that far off the wall, but it went -- it went to GMS, and was reviewed by the Genetics Study Section, which was the Executive Secretary of, and had a lot of yeast geneticists on it and fly geneticists and model organism oriented section. The -- although it was called the genetic section, I served on it for many years later -- it -- they didn't do mammalian genetics, it was -- this was -- basically they did plant genetics and model organisms. And so they got this grant and, you know, it had a -- I think, you know, it had a -- it had a cerebral tradition. I saw that later. I don't know who the actual reviewers were, but they liked it. And here I was, as assistant professor, had no preliminary data. I did -- I had done some -- I did some computer simulations when I -- right when I first got to St. Louis. There were no computers in our new genetics department, but I found some over in epidemiology kind of unit and found a mainframe that punched cards. Again, I remembered a key point when the deadline for the grant was getting close, I dropped my deck of cards on the shuttle bus and at that time the instructions in a computer program, they were just ordered by having the cards in order [laughs], and I dropped the deck. And had to reconstruct the program by shuffling the cards and [laughs] getting it working by the next morning and so forth. But I had done some computer simulations that showed that I -- that at least on simulated data I would actually be able to build a map. I didn't have any -- didn't have any preliminary data. Anyway, amazingly it was funded. Five year grant to a new assistant professor to do something that was just off the wall, you know. They weren't choosing the best of many grants of this kind, they'd never seen anything like it. And wish me well. So I started to work, and that went on for -- it was a good solid 10 years before we had a

map that kind of still was not perfect but was pretty good and well correlated with the genetics map -- the genetic map and widely useful even in that time to each geneticist. So I had to get the grant renewed once I actually -- the grant -- for another five years. And this just shows how much times have changed, so we'd published, I think, about three papers in the five years. And they were all purely -- they were component methodologies. How do to this part of it and this part it. We had not yet published our first paper showing that the method was going to work. That we could actually put these component technologies together and make them make a map. And it got renewed for another five years. So, I think that happened because by then the project was well known in the yeast community. It was not well known generally, but it was well known in the yeast community and we were from time to time able to help other yeast geneticists with their problems in some part of the genome or another. Most of yeast genetics at that time involved, you know, cloning and analyzing little segments of the genome. And from time to time we could be helpful and people could see -- you know -- I presented data at meetings and they -- you know -- they could see this was -- this was going to work. Needed just a little patience and so that went on. You know, it was not a big project, I did a lot of the work myself. And never had more than -- never had a student or post-doc working on it. That was true, actually, for the whole course of the project. Yeah, it was me and one or two technicians and, in the later stages I would have usually one computer programmer and the -- so that's kind of why it went slowly. But the pace was actually about right. I think this is an important point that's not well understood about these early stages of technology is that if I'd had a lot more money, there's no assurance whatsoever that it might have gone better, and it might have gone worse. Because the reality is we didn't know how to -- we constantly ran into problems and we didn't know how to solve these problems. Brute force never works. You cannot solve problems that you don't know how to solve by brute force. You have to do some trial and error. It's very difficult to parallel wise, you know, you can't have one person trying one thing and one person trying another and some Darwinian kind of a struggle. I'd never seen that work. You know, you need a small group -- very small group -- that figures out how to solve these problems. And it's difficult to rush. This is a generalization about the history of genomics, is that some of the -- hard to explain, you know. It takes some Silicone Valley type. I mean, they're the worst. They're the worst because, you know, they're smart and think that the kind of lessons that they learned from the growth of IT generalized all of life and you try to explain to the, you know, why it was -- why it took this long -- took as long as it did. You know, it was only, you know, from the time that I started working on the yeast map -- that was 1979 -- you know, to when the human genome project had a fairly good sequence of the human genome, what is only a little over 20 years. And the number of people working in the field sort of grew exponentially and funding grew exponentially through that whole period. And eventually the field was mature enough that a lot of parallel trial and error could go on. But anyway, to try to explain, you know, why it -- why it took 20 years to somebody in IT is difficult because -- and the core reason is that they're not dealing with a real world. They're dealing with a -- an idealized world that they create. That's what computer science is -- it's an idealized world. You know, the transistors are on or they're off. You know, a bit is either set or it's not set. And the deductive logic and the combinatoric kind of finite mathematics works and applies. And -- but most of them don't actually understand what was going on with the electrical engineering side of computers. I mean, to get computers reliable enough so that they -- so that software could become the major problem took many decades because a real transistor, of course, is not on or off, you know. It's a complicated device that, you know, has its own ideas about it; they're quality control problems, all these kinds of things. But that's not part, actually of the Silicon Valley legend. The Silicon

Valley legend is that somebody else, you know, figured out how to build these chips that were reliable enough that everything became a software problem. And, well, in recombinant DNA techniques, just in experimental biology, you're not working with materials that are as easily modeled as doped silicone. And so we had, you know, bigger problems everywhere you looked. And they were messy problems and they took a lot of trial and error. And this is a history of the human genome project that has not been written and is worth looking at, is really actually what were, kind of -- even if you write a pure wig history of the human genome project, much less a history that captures the confusion that prevailed at -- along the way, you would quickly identify at least a couple of dozen. And that would be a very sparse list of problems that got solved that proved to be critical, that no one recognized very far in advance of colliding with this kind of problem and where there -- where there was no consensus about what the best way of solving the problem would be and often even the people who solve the problems didn't recognize the really -- the really high bounces [laughs] were the solutions that had unintended consequences and they -- good unintended consequences, that is they affected multiple things. I mean, I just could give a very trivial example, would be that in the -- in four colorful essence sequencing, you know, it still depending on electrophoresis, with single nucleotide resolution. You know, these are polymer molecules, they -- that actually met -- how they migrate actually -- their sequence actually matters, not just their length, whereas, kind of, the idealized model depended on single nucleotide resolution by length, but independence of sequence. And so you got these compressions and un-sequenceable sequences because the molecules adopted odd conformations and so forth and people tried running the gels very hot and putting in a lot of denaturants and so forth. None of it worked very well. A major breakthrough -- wasn't the only one, but a major breakthrough was that for completely independent reasons, they -- the fluorescent labeling kind of went from sort of five prime end labeling to three prime end labeling of these molecules with dye terminators. And the dyes, you know, were so highly decorated -- the nucleotides that were being added at the -- were so heavily decorated with all this organic chemistry that didn't belong there, that it turned out to interfere with the formation of the hairpin structures at the end of -- ends of a lot of these molecules that caused a lot of these compressions and un-sequenceable regions. That was an example of an unintended consequence of a couple of different things going to dye terminators, making these energy transfer dyes, which required a lot of -- a lot more modification of the nucleotides and it, you know, it solved a problem that had lingered for years and years, largely solved a problem that lingered for years and years and there are many other examples. But anyway, our 10 years on the yeast project, I, you know -- I -- it's a -- that would be a micro-history, but I could -- I could defend most of them. Obviously I made -- you know, there are some places you just find I made a mistake. Read the data that we already had wrong and in retrospect it was pretty obvious, I should have done something different. But on the whole I can defend most of that time in this way, you know. We just had to get -- kind of muddle through and -- but always using kind of the Ron Davis rule that, you know, of course initially you muddle through but then you figure out, you know, what -- what's -- what am I basically doing right that's solving this problem and let's really understand the problem now and make this a robust solution because genomics has always been -- experimental genomics has always been a, you know, a very frustrating from a process point of view. You know, it has a zillion steps and there's no -- none of them is strongly a rate limiting. This is process engineer's nightmare, is, you know, many, many steps, you know. Starting with, you know, some blood that's drawn from a, you know, a patient and ending up with a GenBank file. There are all these steps and -- and there really is just not a rate limiting step in there. There's nothing even, you know -- there are 10 or 15 steps that, you know, very slight changes in the way

that you're doing things can shift the -- shift the burden. But even then they're not typically strongly rate limiting and so that means that no one thing you do is going to have a big effect. That's just reality and that's as true today as it was then. It's just the way it is. Now, of course, the -- you know -- the methods are much better now and there're fewer steps. There's still a lot, but they're fewer steps and they work better. But there's not -- you tell me what -- what's the rate limiting step in clinical sequencing? You know, you've got a cancer patient you want data now. You know, you take a blood sample or a tumor biopsy or something and you want a whole genome sequence of this thing. Well, there are a series of steps that you've got to go through and making any one of them go even -- go away completely doesn't actually change the throughput very much because it's just the way it is. And these are messy materials, as I said, this is not dope silicone. We don't have good models for most of these steps, actually, so there's a lot of empirical work that goes into characterizing kind of how you get robust protocols, but that's -- so it was the yeast map, you know. It worked. Took a long time. The -- we had good -- you know, good relations from really the beginning with the only other project, which was John Sulston and Ellen Culsen's [spelled phonetically] [unintelligible] that was -- I learned Bob Waterston was the intermediary because he trained like I guess like John Sulston, but John Sulston had been party of Sydney Brenner's kind of MRC nematode group. I'm not sure what exactly -- what his role was -- how that happened -- but anyway, he was a card carrying member of that small group. Bob Waterston was a near charter member himself having post-doc'd there and he was, you know, he was my neighbor at -- and cheerleader, chief cheerleader through this whole yeast project phase at Wash U. And so he was in close touch with John and was the only person, initially, who realized that I was setting out on this yeast project and they were setting out on this worm project with similar goals and certainly closely related methods. The methods differed in detail, but they were random -- both random clone strategies that involved getting these restriction digest and -- you know -- and complexity picking of clones and N-squared complexity map assembly by -- in a computer, and -- so they -- you know, you can't have better competitors than these. You know they, first of all, were really great people, outstanding scientists, and just, you know, a pleasure to interact with and we traded a lot of information throughout.And importantly we both had this -- similar attitudes toward what we were doing and were both reluctant to publish kind of a landmark paper because it was -- we had too many problems and didn't want to sweep them under the rug. But finally Sydney Brenner, who had been watching all this pretty close up, sort of decreed that it was time for papers and so he instructed John to write one and me to write one and can't say no to Sydney Brenner. You know, obviously it was very generous for a scientist of his stature to take such an interest in this, and he communicated these two papers back to back in the PANS in 1986. And that was -- he was absolutely right. He had a much better sense of the, kind of, politics, if you like, of genomics. He saw that that was kind of the end of the period when, you know, these -- you know, small groups could work year after year, after year, pretty much on their own with essentially no competition and not much interest in what they were doing and that it was, you know, going to become a big field and would have the -- acquired the dynamic of a big field as opposed to peripheral activity within molecular biology and genetics. So we did. We published these papers, and you know, we were both I think essentially at the same stage. They were working on a bigger genome but they, you know, the -- they were -- you know, they adjusted details of their strategy so that the amount of work that they had done was comparable to the amount of work we'd done. The -- they had the same continuity problems that we did and -- but also had the same successes. You know, they were -- you know -- we were both at the stage where we were building very good contigs. They were disappointedly small, typically, you know, the typical contig was

only a couple of tiling links of whatever kind of clones you built them out of. And, now, my computer simulations, once I got the cards in the right order in my deck of cards back in, you know, many years before when I had simulated the problem I got bigger contigs. And there are a lot of reasons for that, but that problem actually continued to plague genomics and is not totally gone today [laughs]. The main reason that the human reference sequence is sort of God's gift to -- or the NHGRI's gift to, kind of, the genomics community is that -- is that it has excellent continuity. It's not perfect, but it's excellent and achieving that level of continuity today -- starting from scratch -- is extremely hard. But having one prototype, it's relatively easy to assemble very similar genomes. And, anyway, we hadn't solved the long range continuity problem yet and that -- so that was '86. It was another few years before either they or we started to produce contigs that were long enough, at least in our model organisms so that the -- you know, a good criterion is how many centimorgans are the contigs. Forget about kb, it's centimorgans. Because the -- we were trying to integrate physical and genetic maps and you know they have to be a few centimorgans or else geneticists are going to be frustrated with them and getting them to a few centimorgans was hard. But we got there.

I think that one would have to ask Sydney, kind of what he knew and how much of his vision was his vision. I mean, he's a visionary, and how much he -- you know, he was obviously several orders of magnitude better connected than I was [laughs] to kind of -- kind of the network of people thinking about these problems but I think his thing about us needing to publish in '86 was a recognition that, you know, there were going to be big efforts and the -- so it was about then and I actually -- you know, in retrospect there was not some moment when I heard about the human genome project. It -- you would probably know the date, I'm not sure the, you know, the famous Cold Spring Harbor symposium, was that '86 or '87? '87, yeah so that would have been the spring of '87 so by then the -- I didn't go to that meeting. I was -- you know, I was not well known. I was reasonably well known within the yeast community but I was not well known outside of it and the -- Sulston was much better known, although not really in human genetics, for example, his -- he was already -- John Sulston won the Nobel Prize for the -- for his work on the nematode cell lineage, which had -- which had preceded all of this physical mapping and that work was really seminal and he was very well known amongst molecular biologists, so he had a high stature of -- in biology that I didn't have. I was just an idiosyncratic yeast geneticist. Or in a way that I would have been, I think, described then -- is the way I thought of myself. So anyway, I wasn't very well connected and I -- but I started to hear, you know, secondhand, third-hand reports. The Department of Energy, Los Alamos, Technology that was going to really, really knock out this problem on a human genome scale and the -- and by the time -- by the time of that Cold Spring Harbor symposium I, you know, I didn't go to the -- Sulston went to the Sinsheimer meeting at Santa Cruz. I just wasn't well known at that time. Wasn't on the, kind of, invitation list for those kinds of events. There weren't very many of those events, but anyway he -- so, I sort of knew that a plan was being pulled together. I remember being at Cold Spring Harbor at a Banbury meeting and I'm virtually sure this was before the symposium and I can't remember the date or the topic. I've been at many Banbury meetings, but the -- but I remember sitting next to Jim Watson [spelled phonetically] at dinner at the Banbury Center and asked him what he knew [laughs] about this human genome project idea. And so that shows the level of my naiveté. He didn't know very much, actually, but the main thing he said was he would support it if they got somebody good to run it [laughs]. This was a long time before he was a candidate for this job, but his point was, I learned, I got to know Jim well after that, I met him a couple of times before then, but I got to

know him well during the, kind of, the genome project and still talk to him pretty regularly. But there's always a point, you know, they say when Jim says something there's always a point. Now, often people don't like it and, you know, he says things that he shouldn't say and so forth. This is who he is. But let's just stick to science policy issues -- there's always a point. And he usually makes it very indirectly and the -- I told an anecdote, this is a total digression, but I told a total anecdote about this, which you can find in a book review that I wrote in BioEssays reviewing the -- a book of tributes to Jim Watson. Something -- something science. What's the title of the book? Easy to find, it was a Cold Spring Harbor publication of just essays of people who had worked with Jim and knew Jim and I wrote a review of it, I didn't write any chapters for it. Wyngaarden has -- Wyn -- Wyngaarden? Jim Winegarden has a -- actually a somewhat important essay in there. For I know, the only place there probably are NIH archives, but it's the only place that I know publicly where he sort of wrote down his thought process in sort of basically grabbing the human genome project for the NIH over a lot of intramural opposition. But in any event, it's an interesting collection of essays. In this book review I tell an anecdote about where the $3 billion budget for the human genome project came from and it's a -- it's a Jim Watson story and it illustrates this point that he makes his points very indirectly. So, his point about if they got somebody good to run it is that he doesn't think anything good ever comes of a bureaucracy and his worry about the human genome project is that there was going to be a big bureaucracy built up and it would be an embarrassment to the whole field because it would flounder around. That -- and Jim actually often starts with these, you know, essentially political points. And only later did I really extract from him his thinking about why it would be scientifically useful [laughs]. I think he took it for granted that it would be scientifically useful if it were to boondoggle [laughs]. And so what he was thinking about that night was how to keep this thing from being a boondoggle. But anyway, I -- you know, so I heard these things I -- and my -- I got swept up in the, kind of that policy world. You know, because the -- I had nothing to do with the launching of the -- what turned into the Alberts committee, the NRC Committee on mapping and sequencing. I don't know the story, you know -- Bob Cook Deegan's book probably the best published source about the kind of the -- kind of how that happened. But you know, one thing the NRC does well is, you know, they have a good staff and they research things. And you know, scientists, they -- but the staff, I think they get the staff, kind of outside volunteer balance well. They do that well at the NRC. The staff is very active and compliments -- this is at their best -- compliments the weakness of people like me, that you know, we don't systematically look at fields. We know people, we hear things, we read papers, but anyway they study the landscape and you know, I think things that they discovered my project in St. Louis as just one of a small number of activities that were really directly relevant to this proposal and so I was invited at one of their very first meetings. I'm not sure if they had met previously to this meeting, but it was the first meeting where they were taking, kind of outside input to talk about my project in St. Louis. And so I came and I talked about it. We had just really gotten -- we were really just getting -- yeah I had the used stuff. It was at a, you know, contig closure phase. And -- but we were, the YACs were working by the time I talked there. It was pretty clear they were working; hadn't done much with them, but they were working. And the -- so Bruce, in particular, was really encouraged me to stay in touch and asked me to write him a letter, which I did. I don't have a copy of this letter, but I -- must exist in the archives somewhere, but I wrote him a letter. It was the first time I remember really writing down what I thought about this whole proposal and what the issues were going to be. And I remember the main -- the main -- it was not a long letter. It wasn't supposed to be a big white paper, it was a letter. And I remember the main advice I gave was not to listen to people who were saying that it was

going to be easy. And that there were two reasons that they shouldn't listen to them, there were quite a few, including Wally Gilbert [spelled phonetically], who was on the committee, who thought this was not going to be very hard. I said you don't want to say this for two reasons. First of all -- and be wrong and sets you up for failure because it is -- it's going to be hard and if you say it's easy going in and it turns out to be hard then that's not good. And the other reason it's just simply politically. You've got to argue for a special effort. That was the key phrase in that NRC report is we need a special effort. It's not going to happen by itself. It's -- needs -- you know, this is going to need careful attention. It's going to need institutional nurturing. It's going to need a bureaucracy, God forbid. And, you know, if you make it sound too easy you're undermining the case that any real special effort's needed. The scientific community does well enough. If things aren't so hard, making out hot coalitions and scraping together enough funding at least to really get the ting seriously off the ground and -- so, I want to make this -- and the good thing about this argument is that it's actually true. It's going to be hard. What's difficult is explaining why it's going to be hard and we've already sort of covered some of that territory. But, anyway, Bruce liked my letter. I don't know whether he shared it with the committee or not, and the -- you know, the next stage of the story is well known is that so Wally got so impatient with the deliberations that he did a kind of -- this story actually is not so well documented, but it's known, but it's not very well documented -- he decided on a sort of what was kind of a pre-solera sort of a move, you know. Let's get a little company together, raise a little capital and go and do this. And took some steps in that direction and so he resigned from the committee. It was an obvious conflict of interest, and the -- I don't know if it was Bruce that decided or you know, whoever, but I got a phone call from John Buress [spelled phonetically] who was the Executive Secretary, asking if I'd take Wally's place on the committee and so I had real apprehension about doing that. I mean, you know, there was Sydney Brenner and Lee Hood and you know, Frank Ruttle [spelled phonetically] and Jim Watson, and you know, Bruce Albertson, and so forth. I'm leaving -- you know, I'm leaving out many famous people. There was nobody on this committee that was, you know, a back bencher of my standing. The only -- really the closest peer I had on the committee was Shirley Tillman [spelled phonetically]. And she -- but she was much better known than I was, the leader kind of lineage and had already done well known work in molecular biology on functional aspects of globe and regulation and so forth. Anyway, she was much better known, but about the same career stage I was at. And -- but Shirley and I worked well together. We were the only people on the committee that had sequenced any DNA to speak of. I've told a joke a number of times that at a -- which is, I think, it was accurate -- you'd probably get different versions of it from Shirley and Dave Botstein [spelled phonetically] and me, but -- and I don't know which one would be right, but what I remember is that at a coffee break one time, Shirley and I were conferring about the gap between the reality of DNA sequencing in 1987 and the billions of base pairs and, you know, we compared notes as to how much sequencing the two of us had done and, you know, she'd done more than I had because she worked on bigger genes. But, you know, it was way up in the many thousands of base pairs. And so we wondered how many people on the committee had sequenced with their own hands, at least one thousand base pairs of DNA and we -- so the only taker was Dave Botstein claimed to have sequenced a thousand base pairs with his own hands. Nobody else even claimed to have. Keep in mind, you know, we have Lee Hood and -- on the committee -- and Wally wasn't there anymore but I'm sure he wouldn't' have claimed [laughs] to have sequenced a thousand base pairs. You know, we had -- kind of -- they didn't do this stuff and Shirley and I had done it and -- not on a large scale, but we had actually done it.

And so we initially disallowed Botstein's claim because it was based on having sequenced the URA3 gene in yeast, he has a paper on that. But the gene's only 1,100 base pairs and there were two other authors on the paper and so we just thought it was unlikely that Dave had actually sequenced 1,000 of the base pairs himself. This was undoubtedly unfair and it was all -- this was all for fun [laughs]. So Shirley and I and, to a more limited extent, Dave Botstein, you know, had a little experience actually in the lab doing these things. And we did tend to be -- the three of us -- tended to be allies on the committee on the side of caution. You know, we were caution in the sense -- don't make this sound too easy, it's not going to be easy. The scale up factors that we're talking about here are too big. We -- you know, the report rather famously, and I think this was my suggestion, adopted, you know, the de facto rule of Fred Sanger [spelled phonetically]. I don't know whether he ever wrote this rule down, but it was commonly discussed amongst people who followed technical side of sequencing. You know, that you wanted to move from project to project with a scale up factor that was big enough that you couldn't just do it the way you had done the previous one, but that wouldn't break your system completely. And he settled on an approximate factor of three. And, you know, he looked -- he went from a PhiX to human mitochondrial DNA to lambda DNA to EB virus, you know, the four successive factors of three. And if you look, every one of those was done by really substantial innovation, but not so much innovation that you really had to start from scratch. Well, if you take, you know, take a base line of a few thousand base pairs, which was the state of the art in 1987 and get to billions of base pairs, that's a lot of factors of three. And -- but, you know, we just hammered at the point that if you're going to line up with anybody's view of how to improve this technology, Sanger's probably the one [laughs] to line up with. You know, we should claim that we know how to scale up DNA sequencing better than Fred Sanger did. And so that did kind of prevail if you read the report, that sort of attitude. And actually if you look at what happened, you know, it's not a bad approximation, you know? At some stage, you know, the technology -- that generation of technology got generic enough that it could be parallelized; the factor of three is -- when there's no protocol to copy, you've got to work out the protocol. And the -- there was a lot of working out of protocols. It's more than protocols, you know, it's more -- it's a little more -- there's strategy. The strategy was pretty set. It's more, say, tactics. And there was still quite a lot of tactical maneuver going on until really the end of the '90s when there was a kind of a massive convergence on a particular tactic and very little difference between the practice here and practice there.

And it became an issue of making best practice within a -- within a pretty tightly confined strategy. Makes sure best practice spreads quickly. And I think they -- that was one thing that NHGRI I think did well, actually. It's easy to complain about the G5 System, I was not a part of it. And there still are hard feeling about many things, but I think it was an effective method of having best practice spread rapidly, and that that was a -- that was probably the highest priority that needed attention at that time.

Well, framing it as a Langer and Botstein question sort of invites a slightly humorous response, [laughs] is that one of Botstein's major contributions was that he kind of created Eric Lander [spelled phonetically] [laughs], but you know, so that was a major contribution. The -- I don't -- I don't have a lot to say about Langer, actually. I'm happy to talk about Botstein. You know, he and I were allies for a very long time. The -- there's an anecdote about Botstein and me, it's a true -- this is a true anecdote, I remember this pretty clearly. I had a -- you know, he was a yeast guy, and a big fan of the work even that I did in Ben Hall's lab, which he followed closely. I remember

him sitting in the front row of a Gordon conference where I presented kind of a key aspect of this yeast tRNA study the more genomic aspect of it, and he was just sort of beside himself with enthusiasm whereas lots of other people, you know, they thought this was kind of clever, but you know, waiting for the functional punch line. And Botstein took it for what it was as opposed to looking for, you know, learning something new about tRNA genes. But in any event, I always like to tell the anecdote about when David and I first met, the -- was at a yeast meeting in Cold Spring Harbor, which was 1977, and I had just sequenced with Howard Goodman [spelled phonetically], and Ben Maxim Gilbert [spelled phonetically] sequenced the first of these mutant yeast tRNA genes. And that was my first sequencing and Howard Goodman's lab had Maxim and Gilbert sequencing kind of up and running and I went down to San Francisco just for a few days and learned how to do it and did some of it there. And we got pretty good results and found our mutant. This was the first sequenced eukaryotic mutation. They -- it -- and I -- anyway, so I presented this at Cold Spring Harbor at a -- at a session that Botstein chaired. And he had his usual enthusiasm for all of this, but I'd never really interacted with him. And so that night at the bar, I drank a little too much, not unknown at Cold Spring Harbor, and somehow or another got into an argument with Ron Davis and Dave Botstein, this is the three of us. And it was about DNA sequencing and really this sort of philosophy of science element of DNA sequencing and which I still remember the topic, it was -- they took the position that DNA sequences, in their nature, had to be determined exactly. And I took the position that, you know, experimental science is fallible, it's inexact. You do the best you can and you put error bars on things. So this, of course, was an interesting discussion because it propagates through [laughs] you know, another 20 years of discussion. And my position could be interpreted as my not caring about quality, but anybody who knows me knows that's not right [laughs]. I like to think that I won this argument long term because of course I was right philosophically and eventually basically Phil Green [spelled phonetically] showed how to put error bars that you could really work with [laughs] on these sequences and that that was the critical step in doing it at scale. Certainly, arguably the most critical step in doing it at scale. So, but at the night, anyway, there we were. You know, I'd had too much to drink, they'd probably had too much, too. And I'm sure I'd had more than they had, otherwise I wouldn't have gotten into this argument. So here I'm -- you know, I'm post-doc, a nobody and you know, these guys, they were kind of the rising stars in yeast genetics and molecular biology. And so after a while this argument drew a crowd, you know, this is like in a movie, you know, there's going to be a bar room fight or something [laughs], it's a circle. People went, said yeah there's a post-doc out there arguing with Ron Davis and Dave Botstein and so we -- I don't know, we argued for an hour or two. You know, we're all big talkers, especially Botstein and me. And so I can't remember how it went out, you know, I was drunk and so -- what I remember is -- so I wake up the next morning hung over and I said, "Oh my God, what did I do? I -- have to find some other career. I made a complete ass out of myself in front of the two most important people to impress in the field that I'm trying to make a living in." And this was well before I was out even looking for a job and so I decided I was going to -- that it would be bad form just not to show up at the meeting at the morning. I had already given my talk, but I was going to just sort of sneak in the back and -- a little after the session started, so I'd be sure I didn't run into one of these guys and -- or anybody else who had seen this event. So this was back in the Blackford Hall days before the Grace Auditorium was built and so I kind of sneak in the back, back there and just kind of see what's going on, and so suddenly somebody slaps me on the shoulder from behind and in an inimitable booming voice David Botstein says, "Hi Maynard. That was fun last night." So, anyway, Botstein and I have been great friends ever since and we were allies in the, you know, on

the Alberts committee and then through a lot of these advisory committees. We tended to agree. We didn't, you know, coordinate our arguments ahead of time, but we tended to agree and then we would team up, and we were pretty good at improvising a kind of a back and forth approach where I'd fill in the weaker parts of his argument and he'd fill in the weaker parts of mine. But in Botstein's -- he's very hard to argue about. Often I knew more about the technical details and so if -- but I could shore up his arguments and he was powerful and formulating, you know, of a conceptual level policy kind of idea. So a lot of it, you know, I think we've already covered what a lot of it concerned is that I remember the discussion on the NRC committee, for example, about some people on the committee wanted to put, you know, the year 2000 as the goal. Have this sequence done by the year 2000 and I remember without -- again, without prior coordination, David and I were the two people who just thought that was not going to work. Now, it was very difficult to sit there in 1987 and, you know, decide that 15 years might be enough but that 13 years was cutting it, and -- but we were about right and when we just sort of, both of us, and coming at it from somewhat different directions, sort of looked at what was going to have to happen and then happen after that, it was hard to picture this thing getting done by the year 2000. But, you know, we thought certainly 2010, this was probably very safe, but the -- so it's an example. And issues of that ilk, you know, recurred over and over and over again. And the other thing that we both, you know, agreed on that was a powerful position was, you know, we both -- we were yeast guys and we both thought model organisms were actually the key to this thing that not just an add-on but the key that the -- we both took a fairly dim view of the human genetics community as a scientific community, and there's still people that resent somewhat my attitude.

So he had some exposure to human genetics that I just didn't have. The genetics department in Seattle where I made this rather abrupt transition into genetics that had no significant human genetics. Stan Gartler [spelled phonetically] was there, but he was not working with families, for example. Was primarily at that stage in his career; he was a very positive influence on me, but I didn't learn any human genetics from him. He was primarily interested in the mechanism of exon activation at that stage in his career. And the -- I just didn't have any exposure to human genetics. I, you know, in more recent times have become close with Arno Motulsky, but I didn't know him at that time. I went to one lecture he gave. The human genetics division had essentially no interaction with this basic model organism genetics department. So that was where I came from. David came from this Human Genetics Department at Michigan, and had had some serious exposure to human genetics and knew a lot more about it than I did. And, of course, played, you know, a historic role in recognizing that of the various things that kind of we had to offer from the early stages of genomics, that RFLPs were the thing that human genetics needed the most, and wrote that -- kind of that 1980, I think, paper. I didn't know enough about human genetics to even think about that -- about, you know, why RFLPs were particularly the thing that human geneticists needed. I sort of saw human genetics as needing a sort of a massive collection of tools. Which was essentially accurate, but Botstein could see that this tool was ready to have a big impact in human genetics. I didn't have that level of understanding, that they didn't have enough genetic markers, basically, to do much of anything. I could just see that they couldn't do much of anything. Now they've learned a lot of interesting things, I don't want to be misquoted on this. I liked reading about, you know, human genetic diseases. And particularly when there had been some biochemical success, like the inborn airs of metabolism, I was quite impressed. Sickle cell anemia, these were, you know, these were very interesting scientific stories and in some cases had already had some medical benefit. But, it all looked so peripheral to me to the core question of, you know,

just, "Why are people all so different?" You know, it's -- population geneticists, I had a little -- Joe Felsenstein [spelled phonetically] was there -- I had a little exposure to population genetics. No much. I had more in St. Louis when -- particularly when Dan Hartl [spelled phonetically] came to that department. But, I wasn't so interested then, I've gotten more interested since. But, I wasn't so interested then in a population geneticists answer to these questions. I wanted molecular answers. I mean I -- there are a lot of heritable differences between any pair of humans, and I wanted to know what molecules were different. And I could see that the tools just weren't there. And so when I tended to look at human genetics, until rather recent times, this has changed and is changing. And young, human geneticists today are a totally different group than my comments apply to. But, I'd say well into the '90s, I just saw human genetics as an overly self-satisfied kind of enclave that defined a certain set of problems which they could make a certain amount of progress on, mostly imparting technology from outside -- almost entirely imparting technology from outside. And then, that's okay. What I didn't like is that they were always too satisfied with the technology. I spent, you know, I can't tell you how long; through the late '80s and '90s arguing with human geneticists. The Hughes was the worst. I was a Hughes Investigator for a while. A total misfit with this organization, and they couldn't take my Hughes Investigatorship from St. Louis back to Seattle in '92. So I left the Hughes then. But, for a few years before that I was a Hughes Investigator and I used to go all their meetings and they had a lot of human geneticists. And the -- I'd argue with these guys and they, you know, they would differ about whether a 10 centimorgan, you know, DNA polymorphism map was good enough, or whether their might be some benefits to taking it to five centimorgans. You know, there just was no vision there about, you know, "How are we going to ever do this? That, yeah, how are we ever going to actually ever going to really understand, you know? What are the molecules that, you know, make people different from one another?" And they had at any given stage, a certain set of kind of -- they were all doing the same thing. Some of them did it better than others, but it was always the same thing. And, anyway, I was frustrated by this community. I just didn't see any vision there. They were not very enthusiastic about the Human Genome Project. The Human Genome Project was a -- so, anyway, Botstein and I agreed that, you know, [unintelligible] organisms were the key. You know, they sometimes presented as through we made the case that it was an important add-on. It was actually the key. That how -- you know, that how we're going to figure out how to do these things. And not just -- it goes beyond the methods; the conceptual framework. What is it that we're actually trying to do? And what would the benefits of that be? These questions needed attention, and they needed refinement, and they needed, you know, somebody's sort of smart, knowledgeable people arguing about them. It wasn't going on in human genetics. And it was going on in yeast and worms and eventually -- you know, flies were slow to the table. They were slow to the table for actually a very simple reason. First, well, you know, a couple reasons. They had some of the insularity of human genetics. You're either a fly guy or you weren't a fly guy. And they -- and they had polytene chromosomes. And so they had, you know, they had cytogenetics at a resolution that other people could only dream of. That's not going to solve the problems I mentioned; it's not going to tell you why one fly's different than another fly out around a garbage heap somewhere, but they could do a lot with them. And so they stuck with chromosome walking way too long because they could walk better than other people because they could -- they could see where they were on this cytogenetic map and so forth. But, you know, they came around. But, they didn't leave at all, they dragged their feet. You know it was really worms and yeast that led. And they -- and then, you know, model organisms like arabidopsis, for example. I mean they understood, you know, this can transform our business and so forth. And it, as I say, it was more than methods,

it was just sort of a conceptual framework, you know, what needs to be done? What's the relative importance of all these different parts? How do they fit together? And, also, just getting experience; the factor of three kind of thing from Sanger. But -- so that was another thing that David and I agreed about. But, you know, I think, you know, David's contribution, he -- his contributions as investigator were modest. He was primarily an intellectual force. Every field needs more intellectual forces; I've never been around a field that has enough, you know, has too much intellectual force. Excessive, extra, intellectual force. I've been around a lot of smart people, but intellectual force is a different matter. And the -- and I think it was that that things I've already mentioned. And, well they're all things I've already mentioned. But, it was this combination of, you know, significant experience with the techniques; much more than many of the other big talkers. And the strong rooting in model organisms and -- and this awareness of really what understanding human genetics at a sort of conceptual level was a powerful combination and his -- with his personality. So Lender [spelled phonetically] that'll have to be another day or somebody else, is a different -- it's a whole different discussion.

You know, so obviously, you know, that so one of the -- one of the primary pleasures of a career such as mine as, you know, occasionally to, you know, interact with young investigators that -- at critical stages of their career and sort of feels as though you did something good. You know actually my biggest regrets and my whole scientific career they're not, you know, various scientific errors that I made. You know there's students I feel I let down, you know, I didn't quite figure out. You know, I'm sure there are basketball coaches that feel this way. You know, they just never out quite what to do. You know, this guy had a lot of talent, but I never quite figured out how to plug him in so he could shine. But, you should always focus on your successes. Eric's one of my successes and yeah. So Eric, yeah, Eric is better known to many of you than to me I've seen him over more years, and so forth. He's immensely better organized guy than I am at every level. So it's kind of interesting that we worked so well together during this formative period because our personalities are actually quite different. I haven't been surprised, either by Eric's level of success, or even the general trajectory that he's taken because I always saw, you know, he's a leader. I not -- I'm actually not a leader. I'm a -- sort of an old fashioned professor. I'm at my best when I have a lot of time to think for myself. I, you know, I talked about my sort of zero-based approach to things. Well, that's not an efficient process. And I just -- you know, I've never -- I've never -- I've been an acting chair of a few things just out of community service. I do have some sense of community service about me, but no desire for leadership. You know, I just never pursued any opportunities in leadership; that's not what I want to do. The -- Eric, you know, is a leader, I could see that. And, you know, he'd be -- you know, I was an odd choice of somebody for him to postdoc with, and I don't -- you know, he's the one who'd have to say why he did that. It was pretty adventuresome. He'd had a very successful -- he's an MD-Ph.D. Had personal reasons to want to stay in St. Louis longer after he'd finished his MD-Ph.D. His Ph.D. studies had been extremely successful; were very well regarded in the medical school there; which was a leader in sort of glycobiology. And, you know, he worked on sulfation of the glycosylated proteins and it was biochemistry. It wasn't really biochemical genetics, it was basically biochemistry. And he'd done very well at it. So he's out at this sort of functional extreme of the spectrum that I've been discussing. And here I -- you know, by then, was a well-known figure at Wash U. This was, I don't know, 1997, something like that; 1987, I mean. And wasn't, you know, so well known, nationally, but I was well-known at Wash U., because more and more things were being built on this, kind of this little start-up project of mine. I mean, you know, there was no genomics or

anything that resembled genomics when I went there in '79. And when I left in '92, it was this kind of huge center of genomic activity, and grew more after that. But if you look at that whole history, you'll see that there was never any leadership of mine [laughs]. I just -- I was never in charge of anything, and it frustrated the, you know, the high command to no end because, you know, they wanted me to, you know, bring in tons of money and organize some big thing. But, I -- it's just not what I wanted to do. But, fortunately there are other people that did want to do it and were highly capable of it. But, anyway, somehow on to that, why Eric decided to do this, I don't really know. It was quite a major shift for him, from just the kind of experimental work he had been doing. I wasn't -- you know, he had an MD and planned, you know, to get -- I don't know if he was certified at that time or board qualified at that time as a laboratory-medicine person, but that was certainly where he was headed. And I wasn't developing genetic tests. There were people at Wash U that were developing genetic tests. There were all sorts of things going on that were much more relevant to what he was doing. But -- so he decided to do that. I -- you know, Eric's not hard to read. Some students are hard to read, but he's not hard to read. You know, he's energetic and, you know, hard-working; extremely competent, hard to stop. And we -- you know, I wanted to get a project going that would sort of do for YACs what we'd painstakingly worked out with Lamden cosmic clones in yeast. And I could see that the parallel methods just weren't going to work. They -- it wasn't going to work. And I'm right about that; no one's ever made them work. And that is we weren't going to be fingerprinting YACs by digesting them with restriction enzymes and doing -- and complexity picking of clones and N squared computing and so forth; it just wasn't going to work. There are a bunch of reasons for that, but it wasn't going to work. So PCR was sort of the new kid on the block, technically. At that time it was brand new. And so that really appealed to Eric because that was -- it's relevance to lab medicine was obvious. And he -- so he, you know, he got that up and running and more than that, I mean it wasn't so hard to get up and running, but got everybody doing it [laughs]. And, you know, he made the PCR transition in our lab, and indeed in the whole department there. And we -- yeah, so we developed this idea that, you know, we could combine the screening of the YAC libraries which we shifted to an almost completely PCR-based method. I never liked hybridization screening of things; I'd done a lot of it and I never liked it. Still don't like it. It resurfaces from time to time, but is not the right way of screening libraries, in my opinion. The -- so, anyway, we pretty quickly realized that you could build quite nice maps if you did -- this is just back to what I was talking about with Ken Haizmyth [spelled phonetically] and Ben's lab. If you go in and screen a very deep library, so we have Ron Davis's kind of conception: Always work with a very deep library. So you've got a lot of independent clones screened somehow or another for one point in the genome, and it's like you have random N clones with all of them contained this one point. And -- so PCR was out point. That was getting down to a, you know, pretty close to a point. And we, you know, we quickly realized that we could actually build very good maps just from the -- just from knowing, getting enough PCR assays across the region. Even if they were randomly spaced, we could order the -- order the PCR assays and the YACs in the same -- all in the same. We can screen and order and so forth, all in the same way. And so that was the idea. I actually, you know, clearly it's a compound idea, and it wasn't some afternoon on the blackboard that we put all this together. But, we were trying to get going with YACs what we had going in yeast with these simple clones. And so he got that up and running very quickly and -- and because he was more interested than I was in collaborating with human geneticists, he was very effective at, you know, not just using any old PCR assay. But, that was a period in human genetics when, you know, laboratories, you know, they had armed guards guarding their PCR assays and they wouldn't publish the sequences of the

primers and they -- because it was the closest to, you know, some positional cloning project or another. But, he was very effective at building collaborations; I just left that to him, completely. I didn't have any interactions with these groups, personally, except just sometimes they would ask me if I was really on board. I always said, "If it's okay with Eric, it's okay with me." But, so, for example, and that's how he built his ties to Francis Collins is that the CFTR search was right in its end game. And the -- so the timing actually was sort of perfect because the Francis and his collaborators sort of found the gene by their methods; in the process they developed huge numbers of PCR assays as they were looking for it. And we were no longer a threat to them. And so they and Francis saw the general interest of doing that kind of mapping, and so he and Eric collaborated on getting all the -- you know, they were the ones that, you know, got all these assays together and we did very deep screening across that whole region and kind of showed we could build these maps. And the -- but, it was a fairly brief period, really, by about -- I can't reconstruct the timeline. But, I doubt that Eric was in any meaningful sense a postdoc of mine for more than two years because, this guy, you know, he was on a fast, upward trajectory. So I remember one day -- so I told him, "Look, I have too much space. With all this space, my lab's getting to big. I don't -- big labs are not for me. I want you to take what was my main lab space, including my office, and I'm going to move down the hall to a smaller space and a smaller office, and start over with some new projects and you should run with this." And so the lab medicine or somebody appointed him, you know, when you're a fast-rising star in a sort of the more clinical side of an academic medical school, jobs are never a problem. And he got some kind of a job there. And he did it first, you know, didn't want to do this and so forth. But, I persuaded him and that it would be good for us both. And so he just went around with all that. I wasn't involved in the, you know, I'm not an author on the chromosome seven mapping paper because I didn't have anything to do with it. He had a nice big contig out there on chromosome seven and he just went to work on mapping the rest of the chromosome. And he worked with David Schlessinger, and that's another complicated story. But, the X chromosome was David Schlessinger's kind of baby and chromosome seven was Eric's. And pretty much finished that in St. Louis, I think. But, then Francis hired him here and he, you know, went on and did things. You know, I think the last thing I'll just stick in -- a lot more could be said about Eric, but the -- as far as his scientific contributions go, obviously he was a key player in that kind of formative stage of genomics. But, of the work that, you know, he really did as an independent investigator here, I do think that his comparative genomics is underappreciated. You know, comparative genomics has become, you know, so ubiquitous, and we're not accustomed to comparing whole genomes. And, you know, it's -- a field is at a much more advanced stage. But, you look at the, you know, he took his intramural center in that direction of doing, you know, multiple species, well-chosen multiple species, fairly long tracks of DNA, you know, across multiple genes, complete sequence and, you know, I think, you know, he showed people that this is the way to go. We need multiple sequence alignment over long regions with well-chosen phylogenetic comparisons; produces an immense amount of information. And I think that was a -- I said underappreciated contribution to just the development of genomics as we now know it.