James Gusella & Marcy MacDonald

Personal Details

Name	James Gusella
Dates	Born 1952
Place of Birth	Ottawa, Canada
Main work places	Boston USA
Principal field of work	Human molecular genetics

Personal Details

Name	Marcy MacDonald
Dates	Born 1951
Place of Birth	Curtain Lake, Ontario, Canada
Main work places	Boston, USA
Principal field of work	Human molecular genetics

Interview

Recorded interview made	Yes
Interviewer	Peter Harper
Date of Interview	13/09/2011
Edited transcript available	See Below

INTERVIEW WITH JAMES GUSELLA AND MARCY MACDONALD, 13th SEPTEMBER 2011

PH = Interviewer (Peter Harper)

JG = James Gusella

MM = Marcy MacDonald

PH It's Tuesday, September 13th 2011, and I'm talking with Drs. James Gusella and Marcy MacDonald from Boston at the International Huntington's Meeting in Melbourne about the discovery of the HD [Huntington's disease] gene and a few other more or less related things. Jim, can I start at the beginning and ask you where were you born and when?

JG I was born in Ottawa, Canada in 1952.

PH And did you stay in Ottawa or did you have part of your upbringing in other parts of Canada?

JG I grew up in Ottawa; did all my elementary and high school there; went to the University of Ottawa, so I got my Bachelors Degree of Science from the University of Ottawa. And that's when I left to go to the University of Toronto for a master's degree at the Ontario Cancer Institute. On the day that I arrived at the Ontario Cancer Institute, I had already chosen I was going to do my degree with David Housman, and the way it worked, you did a Masters degree first and then you moved into a PhD programme. And on the day that I arrived David announced that he was moving to Boston, and so I had a choice of whether to choose someone else to work with, or to go to Boston with him. And I made the decision, for a whole variety of reasons related to my future wife and the Ontario Provincial Government and their inability to pay the universities appropriate amounts of money for their medical schools; these things all conspired to cause me to go to Boston with David, but finish my masters degree from Toronto at a distance.

PH So what was the initial theme of your work with David Housman?

JG Well, initially with David I was working on this cell line called Friend erythroleukaemia cells, which basically are a cell line transformed by Friend virus which was interesting to study because you induce it to undergo erythroid differentiation, so in theory you could study the steps involved in triggering erythroid differentiation. And when I moved with him to Boston and it was clear that I was going to be switching from a masters degree to a PhD programme, it was also the same time that recombinant DNA technology was just being evolved, and it was obvious to me that applying recombinant DNA technology to biology was the future. So I had set out to clone the globin gene, since erythroid differentiation ultimately gave expression of globin as a product, but I was denied the ability to do so because the Cambridge City Council in Massachusetts ordained that one could not clone mouse or human DNA within the city limits. And as a result my PhD thesis ended up not being the cloning of a specific gene, because I wasn't allowed to for several years; (for a couple of years I actually worked on trying to clone soy bean DNA, mainly because we were allowed to). But in the end the city council relented, after all the genes I was interested in were already cloned, and so my thesis ended up being the development of a technique for cloning human DNA out of somatic cell hybrids so that you could clone DNA based on the chromosome that it came from rather than the gene that it coded for.

PH Can I ask you: at this time, in the Housman lab, was there any talk or discussion or interest in Huntington's in terms of being a disease where new techniques might allow gene isolation?

JG That kind of discussion just began at the very end of my thesis, because what I was working on was this technique for cloning DNA based on where it came from in the genome; and having established that technique it was pretty clear that you wanted to know the location of something interesting in

order to apply it. And so we were just beginning to talk about the idea of mapping a disease gene to a location in order to have something to apply this technology to, when I was making my plans for doing postdoctoral work and going off to Caltech, and working on immunoglobulins, one of those genes that I had wanted to clone earlier but hadn't been allowed to.

PH And what year are we in now?

JG So we're in late 78, early 79, where essentially I was finishing up my thesis in… let me think about this and get it right. I think my official graduation was in 1980 which means that we were in 79 because I was planning to go to Lee Hood's lab for a post doc, and I visited Caltech in June of 79 and that was what made my decision not to go there for a post doc. [laughs] And so I came back to Boston in the summer of 79 and discussed with David the fact that I had grave doubts about going to Caltech, and my wife didn't really want to go to California, and that was when the opportunity to apply the technology to Huntington's Disease came up and evolved over the next 6 -8 months to the point where I decided to stay and start a lab at Mass General.

PH So in a moment I'm going to turn to Marcy and trace her first steps, but before I do that: did you come from a family with a background in science or medicine, or anything of that kind?

JG No background in science or medicine or anything. My father was a commercial artist, graphic designer; and my mother was essentially a homemaker.

PH And were there any particular factors you can pinpoint that made you especially keen to go into science?

JG I had a great childhood in which I was free to explore and I always loved to explore nature, essentially. And it ended up, I think, creating a curiosity for doing and learning things. And at the same time I had a sort of inherent appreciation for mathematics without a real ability to apply it. And I think the combination of those two things leads to science. [laughs]

PH Marcy, can I ask you now: where and when were you born?

MM I was born in Curtain Lake, Ontario in 1951, in Canada.

PH And, as with Jim, were there any particular factors in your background, either family or other, that led you into science?

MM Not really. My family was, there was no prior scientist in the family; architects and engineers. My father was a mining engineer; my mother was a secretary. But I was very lucky: they both fostered, much as Jim said, an independence and a curiosity, and although we lived in places that didn't have libraries, they were both voracious readers and the house was always covered in books. And they really fostered that reading curiosity and that interest in nature; in biology. And I just always had it. It's always something I wanted to do.

PH Whereabouts were you living during most of your childhood?

MM We moved around northern Ontario and Quebec following the stock prices of the various hard rock mining commodities: gold, silver, copper, uranium. So I think we moved about every 3 years to small communities, and we stayed the longest when I was in high school. We moved to follow into a copper mining community north of Sudbury, Ontario, and when the uranium mining went belly up; and stayed there right through high school, which was unusual, because we moved quite a lot.

PH And where did you go to university?

MM I went to university at the University of Ottawa, and chose to go there because at the time, in Ontario high school, there were 5 years of high school and I didn't want to do the Grade 13 in the small town high school. And my family supported me in looking for a university that would accept somebody who hadn't finished high school but could get in with entrance exams. And the University of Ottawa had a

programme, and I applied and wrote the entrance exams and went to the, what was called the pre-university year at the University of Ottawa; and stayed there to do my Bachelors in biology.

PH And what followed on from that?

MM From there, actually, I met Jim in biology there at the University of Ottawa, and I think it was probably in 3rd or 4th year or something that, I remember having a really detailed discussion with Jim about Jacob and Monod [laughs] and the concept of gene regulation. And of all the things that I heard, I had thought sort of vaguely about biology in various ways but was really interested in, I think, what became molecular biology very much later. And the notion of being able to understand at molecule level what might regulate things and set the path of biology, really fascinated me. And we used to have, Jim and I used to have really good discussions, not necessarily around anything except final exams, but the whole idea that this was something that was brand new, just captured me. And I decided that I wanted to go on to do a degree in something that would allow that kind of development, so did a PhD masters, as Jim said, and a PhD eventually at the University of Toronto. I originally was going to stay in Ottawa and then fortunately managed to get a government studentship that could sponsor me to live in Toronto and applied to the University of Toronto and got into the programme there at the Ontario Cancer Institute; essentially a programme in quantitative biology. It was perfect; just what I wanted.

PH So at what stage did your two careers converge? Was that really the Huntington project that brought the two of you working together?

MM I think we've collaborated really since, for a long time, I think. The idea about HD grew later. I think at university I remember a discussion with Jim about "what level of molecule do you want to work on: RNA, DNA or protein?" And you know,at the time it all just seemed so completely brand new; it was really hard to imagine how you could work on protein, RNA or DNA. So choosing to stay and work on a problem which I eventually did in Toronto. When I arrived for a master's programme I started looking at, with Jim Till, in haematopoiesis; a project that was oriented around a new faculty member, Alan Bernstein, who had been a post doctoral fellow with Jim Till, who was off in London doing his own postdoctoral training. So when Alan returned I became Alan's first graduate student, and worked on haematopoiesis, not involving stem cells as Jim Till's laboratory worked on, but rather Friend erythroleukaemia cells,much as Jim said, and was under the wing of David Housman for the period when Alan hadn't quite finished his post doc and had his faculty position but wasn't actually physically there. So David sort of took me under his wing for a little while -

JG About 6 months.

MM About 6 months; something like that. Got interested in the Friend cell project, which is exactly what Alan wanted, and ended up doing a thesis based, really coming out the notion that because there were rumours that an onc gene had been discovered, that this must be the basis for much of cancer, and therefore much of my thesis was set about trying to find an onc gene in a virus, the Friend leukaemia virus, two versions of it actually, which ended up not having an oncogene. [laughs] But nobody knew that for decades and decades. But that's all right; I got to learn molecular biology. So Huntington's Disease, I didn't get involved in till 1985 because I was doing a post doctoral fellowship with Richard Flavell, first in Amsterdam and then in London; and wanted to return to Boston for personal reasons, and met up with Jim who was travelling through, I think it was Amsterdam.

JG It was. I came to Amsterdam for a muscular dystrophy meeting.

MM Yeah, and got talking, and Jim said, "I have this project: Huntington's disease. Need someone with molecular experience. Come work on it for a year." And that was that.

PH So, coming back to the beginnings then of the Huntington's project: how did it begin?

JG So I, as I said, was supposed to go do a post doc at Caltech. I went out to visit Caltech and the weather was smoggy and hot, and you couldn't keep your eyes open because it was like acid was in your eyes because of the smog, and Lee's lab was clearly one that was both very busy but also very single minded. He made it 100% clear to me I would work on one thing and one thing only, which is just not the way that David worked, and it wasn't the way I worked in David's lab. And I couldn't conceive of the idea of not having multiple things going on at the same time. And the same time my wife found out she, because she trained in Canada, California was the only state in the country where she'd have to write special exams in order to do her medical work, so just nothing fit with respect of going to California. And I came back and this opportunity arose through a chance circumstance of David Housman interacting with Joe Martin, essentially as a result of a scenario in which Alex Rich, who is a famous biologist at MIT, had a son who was injured in a sporting event, and rather than going to the emergency room you go to the head of the neurology department if you're Alex Rich and you're on the board of the hospital. And when they were there Jo Martin was thinking of putting in a response to the Government's call for Huntington's research centres, and asked Alex if he could think of a molecular approach to Huntington's because Jo wanted his centre to be one that was taking advantage of the new molecular biology at the time.

And Alex directed Jo to David, and then David said, "Oh and by the way I have a graduate student who is about to graduate who would be perfect for this." So I ended up making the decision not to go to California but to skip a postdoc and directly set up a lab at Mass General. I got a fellowship, a 3 month fellowship from the Hereditary Disease Foundation, that was meant to cover 3 months worth of salary while we waited to hear if the NIH grant came through. If the NIH grant didn't come through, I was then out looking for another job. But fortunately part of the grant came through; enough to get the project going. And when I say part of it, there were components, there was a programme project, and there were components to the programme project and my component was to do linkage analysis in Huntington's families. And there was another component that was run by Ted Bird that was supposed to supply the Huntington's families. And the Ted Bird component was not funded, but mine was. So I was left with the mandate to do linkage in Huntington's families with an approach that we had proposed of using DNA markers as the polymorphisms for linkage rather than the traditional expressed markers, but without any source of families to do it with.

And at the time Nancy Wexler was a bureaucrat at NINDS; she was the person within their office who was responsible for the project, and so she recognised this deficiency and immediately put me in touch with Mike Conneally who had been funded by the NIH to set up a roster of Huntington's Disease families. And he supplied me with one large American family with a promise to collect more, and that got me going.

PH Right. Now, thinking of the other starting point of the project, the RFLPs: what did you use to begin with?

JG Yeah, okay, so we knew that we wanted to use DNA markers based on the variation that should happen in the DNA between people. And at that time the variation in DNA between people had really only been characterised at the globin locus, and it was done in the context of thalassemia and sickle cell anaemia. So there was a lot of scepticism that there was going to be sufficient DNA variation in the genome to actually do the project; and so people thought it was going to take a very long time, if it was possible at all. And that's why I was quite worried the grant was going to get funded because there was a lot of criticism at the site visit of the fact that we didn't already have all these DNA markers, and what if they didn't exist? So we needed a technology by which to find them, and the technology that was then available for looking at DNA, allowed for sequencing but it was very laborious. You couldn't really apply it to large numbers of people, but you could look at DNA by a blotting technique that had been developed by, Ed Southern, known as Southern blots, where if you cut DNA with one of the restriction enzymes that were originally the basis for recombinant DNA

technology because they cut in a sequence specific manner, and ran a gel electrophoresis to splay the fragments out by size, you could then transfer them to a membrane; and if you had an individual cloned fragment, you could hybridize, make it radioactive, hybridize where the two DNA strands, the one radioactive probe, the other DNA strand on the filter would reanneal and you could put an x-ray film over it and see a band.

And since the enzymes were site specific, if there was a base sequence variation at one of the sites, the enzyme wouldn't cut and you'd get a change in the fragment sizes. So that became the basis of a restriction fragment length polymorphism or RFLP that we needed to look for. But to look for them, you had to had the cloned probes to do the Southern blotting, and human DNA, as I discovered during my thesis when I was trying to figure out how to get DNA from hybrids that distinguished the human DNA from the rodent DNA, has these interspersed repeat sequences in it, and those interspersed repeat sequences get in the way of the Southern blotting because they see lots and lots and lots of fragments. So I started out on the project knowing that I needed to generate probes, and I knew the probes needed to be single copy probes, that didn't have any repeat. But at the time there wasn't any knowledge of what the distribution of the repeat sequences was in the human genome. So the starting point for the whole project was to make a set of libraries that had different size fragments in them, from a few hundred base pairs to 15,000 base pairs; and to hybridize with human DNA and eliminate any clones that hybridized, because those are the ones with repeats; and then take the single copy ones and just do Southern blots against random human DNAs cut with a whole, I think we started with about a dozen restriction enzymes, just looking for variation.

And so we ended up finding, as one might expect, that there are not very many fragments of large size that don't have interspersed repeats, but the ones that you do find are very good for finding polymorphisms because they cover a lot of distance in the genome. And so the first set of probes that we decided to run after finding polymorphisms (we found polymorphisms for maybe a dozen probes) really fell into two classes: they were these large sequences from a Lambda library that had no repeat sequences. There was actually a library that had originally been made by Tom Maniatis, and those were, I think, G3, G6, G8, G9 because the technician that had helped in picking those was Ginger Weeks. And the other thing we used was, because we knew it would cover large distances, was coding sequences from cDNA clones, since they would hit sites wherever the exons were. And we had a bunch of those like parathyroid hormone and insulin and things like that, and we found a bunch of polymorphisms; we got to the point of having, I think we had about 15 polymorphisms from about 12 probes, or something like that; and started running blots.

MM It's funny to think that there weren't very many cloned cDNAs .

JG There were very few cloned cDNAs.

MM My post doc project was globin splicing, given that I came from the OCI team, and hematopoietic stem cell group, and Alan and viruses with erythroleukemia and so on. So the project was splicing but we mostly were running in competition with Tom Maniatis to clone cDNAs for globin and then use them to work out the splicing patterns using S1 nuclease in RNA protection experiments. And so that was one of the very few loci that had any kind of cloned expressed, you know, copies; and people thought that it really couldn't be done, and it was really very difficult to do Maxam Gilbert sequencing, as Peter, as you'll recall, to try and demonstrate the sequence of the things that you did manage to clone. Funny.

PH So at what point was it that you had a sufficient number of polymorphisms to be able to turn to the family material and then use that?

JG We started running probes in the fall of 82, and we ran them against an American family that Mike Conneally had supplied, that [sigh] I can't recall exactly how many affecteds there were in it, I think there were something like a dozen affecteds in that family. And we ran, basically we ran a large batch

of probes, large at that time, although eventually we got much larger, and that was a set of about 12 probes but they saw multiple sites, so there were maybe 15 or 16 polymorphisms. And we ran that through in the fall of 82, and this was myself and Rudy Tansi, who was my technician at the time. We did it together, got the autoradiograms off, looked at them, you know, looked at the bands, gave allele names to the different bands, put them on the pedigrees, didn't really see anything; it didn't look particularly exciting, and dutifully put the data down on paper and then sent it off to Mike Conneally to calculate LOD scores. And if I remember correctly I think that that was just before Christmas that we had sent them off, and I'm not completely certain of it but that time frame seems right. And I think that we got, in January we got the data back that had been run in his lab for LOD scores by Peggy Wallace, it might have been Jonathan at that time, but maybe it was Peggy. Can't remember which one it was, but probably Peggy.

Anyway, they ran LOD scores and one of the probes gave a LOD score of 1.65, as I recall. And that came as a big surprise because when you just looked at the pedigree and the way the marker was being inherited, you couldn't see that there was that much support. Because 1.65 was not near statistical significance but it was definitely deviating from what you would have expected in a small set of probes. And so it puzzled me that you couldn't see it. I actually sat with Rudy for a long time working over this pedigree and trying to figure out where the LOD score was coming from because, you know, I was a molecular biologist at that point; I didn't know anything about calculating LOD scores or genetics or anything. I knew some of the theory of it, but not much. I was depending on Mike Conneally for that. And working through where that LOD score came from was very informative because it kind of gave you the understanding of what was happening when a LOD score was being calculated, that the computation was taking into account the probable genotypes of people that it didn't know, because you didn't have them.

And in a Huntington's family, you often don't have people because your affecteds parents are often dead by the time you study them. And so it was clearly calculating based on frequencies that it knew for the polymorphism but of course at that point we didn't know what the frequency of the polymorphism was in the population; it was calculating it only from the family. And it became clear that knowing what the polymorphism frequency was in the population was going to be really important for getting accurate results, and so we set out to start getting those data. But we also looked at and said, you know, "What are we going to do to go beyond this LOD score? How are we going to test is it real or is it not?" In the intervening period while we were funded and then got hold of Mike's family, we'd started to generate probes and learn how to do high volume Southern blots so that they worked all the time, and which membranes would work the best, because you know, in the beginning it was nitrocellulose and then these nylon membranes came out and then did you want positively charged nylon or… there were all kinds of variants. We ended up actually using, we were successful because we used a nylon membrane that was made not for biology, it was made for soft drink testing; testing bacteria in soft drinks. So we got it directly from the manufacturer -

MM From Millipore?

JG - no, no, no; it wasn't from Millipore. It was from this company in Connecticut that Rudy found, and we bought an enormous role of the stuff.

MM Oh, I remember it actually.

JG Do you remember that? I forget the name of the company, but it was an enormous role of the stuff.

MM It was; it was a fire hazard.

JG Soon as it ran out, blots didn't work anymore. After we did get the linkage, for about a year and half, blots were terrible because we didn't have this membrane anymore. But anyway, so during this period of time Nancy continued with her efforts of getting us more families. And one of the things that, you

kind of need to understand the background of where the centres came from and where the roster came from, because the third component came from the same place: and that was that Marjory Guthrie had lobbied the NIH and Congress, essentially, to do more in Huntington's disease because essentially nothing was known and I think she'd kind of looked around the country and realised no one was doing any research on Huntington's disease. And so she managed to be, as a very good lobbyist, successful in getting a congressional commission set up to investigate, "Why isn't there more being done in Huntington's disease, and make recommendations." And Nancy had, prior to her becoming the project officer at NIH, been assigned as the secretary to that commission.

There were 3 major recommendations that came out of that commission: one was to set up Huntington's disease research centres that they, NIH, would fund; originally there were supposed to be either 5 or 10, I can't remember. In the end they funded 2, one of which was funded at reduced levels, and ours was fortunately the one that was funded at full levels because, as you go through the process, they realise that they have less money than they think that do, and they want to fulfil the mandate but you can't do everything you'd like. The second recommendation was a roster of Huntington's disease families willing to participate in research, and that was awarded as a contract to Mike Conneally in Indiana. And the third recommendation surrounded this report that had come out many years before about a high concentration of Huntington's disease in Venezuela, and this was as a result of a physician named Americo Negretti who was, as all physicians in Venezuela, assigned to do service after graduating from a medical programme, and arrived in this town and I think it was on a Sunday morning as it was described to me, that it seemed a little weird that everybody seemed drunk, and it turned out of course that everybody wasn't drunk, they all had Huntington's disease. It was a very high concentration.

So the discussions around this congressional commission and scientific advisors that had brought together to discuss this were aware of the work of Brown and Goldstein, that had used the rare individuals that had 2 copies of the defect in familial hypercholesterolemia to really work out the role of cholesterol in heart disease. They recognised that finding individuals who were homozygous for the mutation might be extremely valuable, and the best place to find such individuals would be in a very heavily, an area that had a lot of Huntington's disease individuals, and that this Venezuelan population looked very promising. So the third recommendation was a sort of further exploration of this Venezuelan population. And for a while they continued to have scientific meetings where a bunch of scientists would come together and discuss the issue, but nobody actually did anything: they just talked about it. And so once our centre was funded, Nancy, I don't know if she stepped beyond her bounds as a project officer at NIH, but she, I think, had some frustration with the fact that nothing was actually happening on that third recommendation and she got on a plane and went to Venezuela and visited the villagers, and the net result was that she recognised this was an ideal source of families for the linkage study, and made arrangements to continue to visit and collect blood samples and have them come to Boston for analysis.

And what we did was… the other thing I did when I started my lab was I realised that if we were going to get deeply into linkage we were going to have to have a way to have a permanent source of DNA from the families, because we weren't going to be able to go and get blood samples all the time, so we developed a method to rapidly make lymphoblastoid cell lines at high efficiency. So once Nancy connected with the Venezuelan population and the ability to collect bloods, we arranged for them to come back to Boston and for us to make transformed lines on them all. At the time that we were analysing the American family, we had been banking lymphoblast lines from the Venezuelans but we hadn't actually prepared DNA from them; we just banked the lymphoblast lines because we didn't want to risk any of the blood that came on making DNA; we wanted to make sure we had enough to make the lymphoblast lines. So when we got this LOD score of 1.65… I think it was 1.65… it was in that range anyway; it's in the original paper. [laughs] We then said, "Well obviously the next step is to genotype some Venezuelans." And so we had a lot of Venezuelans but we chose a grouping that could

be particularly informative for linkage, to grow up right away and get tested. So that was early in 83 that we started growing; we started growing them as soon as we got the LOD score. We got the blots done and I remember getting the data for that just before going to a meeting of this group, which, whatever it was called at the time -

PH Was that the one in Rochester?

JG No, it was the one in Chicago; it was in Chicago. Pretty sure it was in Chicago. I know it was on April 2nd, if I remember that correctly. I'm going to have to look back on all these things and see how well I really remember them. Anyway, I'm pretty sure it was in Chicago and I know I got the data just before, and the data was sufficient, it was really only on two adjacent small families with several children, and many of those children were not affected so in the probability calculations you got something added to the LOD score, but not very much. But the reason I had chosen it was that the two families were connected back two generations, so you could get events that way. And lo and behold the data were consistent and now when you looked at it, you could see it, but it only raised the LOD score to 2.2. So I remember reporting at that Chicago meeting that the LOD score could go up or it could go down, but it was, you know, it was not yet significant. It was interesting. I was clearly more excited about growing more Venezuelans, and knew that, I figure through my off hand calculations of not knowing that much about LOD scores, that if you got to 2.2 you had about a 50/50 shot: it was either going to go over 3, or it was going to go away.

And so I got back from that meeting and, with Rudy, we set up to do the entire Venezuelan pedigree. We grew, Marianne Anderson was the tissue culture tech who had developed the lymphoblast technique, and she set up to grow all the lines for DNA, and then Rudy and I set out to run lots on the entire pedigree, which was, oh at that time, hundreds of individuals. I can't remember exactly, but it was a lot more blots than we would normally run. And the probe that we were using was one of those original, long pieces of DNA called G8 that didn't have any repeat sequences. The reason it was particularly good is because it actually detected multiple RFLPs. It detected two RFLPs with Hind III and two with EcoRI. And it was the Hind III that we were using most effectively, although the [s.l. EcoR1 worked as well. But we had to set up, for both those enzymes to run all those blots, because I really wanted to get all the data at once. I didn't want to go through this next step of having to go up a little bit and down a little bit; I wanted it all at once. So we ran all of that in June of 83; got all those blots together, got the data back at the end of July, because getting the blots through this cycle took a lot longer then. You know, you couldn't order radioactivity and have it there the next day back then; you had to plan everything ahead.

So we got it all done and in July the data came off just before I was supposed to go to a meeting in Aspen, at the Aspen Institute. And my wife and I, who at that point had a 2 year old, had to drive to Ottawa to drop the 2 year old off with the grandparents so that we could go to Aspen, and I got the data off just before having to leave for driving to Canada. And this time you could really see it on the pedigree that it was absolutely linked. I didn't need a LOD score to know that this was absolutely linked. So I spent that drive to Ottawa recognising that everything had changed all of a sudden, and the whole plan had to change from then on, as to exactly what you're going to do. Because the first step that was supposed to take years and years and years was already over. So I don't remember that trip very much in terms of the drive; I remember the line down the middle of the road; that was about it. That was great. So then we got to Aspen; I had sent the data off to Mike Conneally's lab; Mike was in Yellowstone on a lake somewhere, and the lab in Indiana calculated the LOD score and obviously they knew what was going on. They tried to get in touch with Mike; the park ranger had to go pull him off the lake or whatever, to tell him. I was sitting in a, the way the Aspen meetings worked, they put you up in somebody's ski condo during the meeting, and so Marie and I were in this ski condo getting ready to go out to eat dinner, and I mean it was old for me because I already knew, but Mike had just found out and Mike called Nancy. And I hadn't told Nancy because I wanted the LOD score. You know,

I didn't want to prematurely say something and then it turn out I didn't know how to calculate or something.

So Nancy called me and I ended up spending probably 3 hours on the phone with Nancy, who turned out, was at a Banbury conference at NIH, sitting in that little phone booth at the back of the Banbury room, so that by the time I got off the phone all the restaurants in Aspen were closed. I ended up, I found one Chinese restaurant that had lemon chicken. The only time in my life I ever had lemon chicken, and it was good. But that was great, it was a very, you know, an experience. And then we submitted, so there was a Human Genetics meeting in LA in the middle of August, that we talked about it just to the committee that, at that point there were committees that dealt with different chromosomes and stuff; we talked to them there, we put an abstract in to the American Society for Human Genetics as a late-breaking abstract; did a presentation; presentation was supposed to be private, not to be discussed in the press; not everyone recognised that so there was some premature discussion of it. Anyway, it was a great experience; very exciting.

PH So, summer of 1983 you had mapped the gene but how did you plan a strategy then from going from mapping to actually isolating the gene?

JG Okay, so there were, so the first thing that happened: we had the gene. We had it with a reasonable knowledge, a definite knowledge that it was the gene in terms of it's location in the Venezuelan pedigree; and a reasonable knowledge that it was in the American pedigree, even though that, by itself, did not achieve a LOD score of 3, given the LOD score and given the Venezuelans it had to be that gene. We didn't know if that was going to be the case in other Huntington's families, so the first thing that we did, in fact, was to get in touch with the various other groups that were doing Huntington studies, like Susan Folstein at Hopkins, yourself at Cardiff; Ira Shoulson at Rochester, and try to get DNA from as many more families as we could both to confirm that there wasn't non allelic heterogeneity, because as soon as we published there was heavy pressure to apply the probe to predictive testing. And there was actually an exchange in Nature over it, because John Edwards didn't like the fact that we didn't want to let people immediately… we made the probe available for research but we didn't want to make it immediately available for diagnostic testing given the dangers involved in that in Huntington's disease, and the fact that we only knew that it was linked in two families. So we got more families collected.

At the same time we knew we needed to figure out where it was on chromosome 4 in order to be able to narrow in, and we set up to do that in a variety of ways: we collaborated with people who had somatic cell hybrids that had pieces of chromosome 4; we went out and collected, again by connection to clinical geneticists, in this case, cases of Wolf Hirschhorn syndrome because they had various deletions in the short arm of chromosome 4; and we collaborated with people who were just in the sort of early days of doing fluorescent in situ hybridisations. And all of those data kind of converged on the idea that the probe was out on the tip of the short arm of chromosome 4. That took, that actually took a couple of years to get that all together, partly because, as I said, our blots stopped working when we ran out of that membrane, and so getting more data on all of the families took longer than we expected; the FISH technique wasn't well worked out so getting that took longer than expected; collecting the Wolf Hirschhorn people took longer than expected. But in the end, within a couple of years, we knew that we were dealing with the tip of chromosome 4 and we knew that the distance between the original probe and the HD gene was about 4% recombination, whatever that turned out to be in physical distance; it was about 4% recombination.

So in the meantime, in thinking about that we would fairly soon know an approximate location on the chromosome and know an approximate distance from the original probe, we thought about two things: one is the idea of getting a flanking marker on the other side of the gene to be able to narrow in; and the other was of what techniques are we going to actually use to narrow it? And at that point I was faced with two problems: one problem is that we got the linkage so early that there weren't any

techniques that had been worked out to move from probe to gene; you didn't at that point walk along the chromosome or jump along the chromosome, or anything along the chromosome. And the other problem I had was one of, I had a small grant as part of this centre to do this work, and it was no longer sufficient to actually do the work, but I wasn't in a position to get another grant because I already had one on that topic, and couldn't really write another one. Nancy again helped out a little bit there in the sense that the Venezuelans offered the opportunity to get a different kind of grant. Because they were such big families I applied for a grant to just build linkage maps of RFLPs, but of course had to actually build those linkage maps.

So we did things on chromosome 21 and various places, but it helped to create some critical mass in the lab. But it was really clear that there needed to be a concerted effort to develop technologies that would be applicable. And the ones that were natural for my lab, I had actually been exploring as part of a muscular dystrophy project, because I also was working on muscular dystrophy before that gene was cloned; and it sort of flowed from where I'd come from in Toronto and then David Housman's lab: there was always a strong emphasis on somatic cell genetics associated with the work, and so just like I had developed this technique for cloning DNA based on where it came from, by cloning somatic cell hybrids that had only a piece of human chromosome, I'd also been working on a technique for selecting, for inserting selectable markers into the chromosomes and then producing hybrids that had regions]. So we were going to produce hybrids that pieces of DNA from the short arm of chromosome 4, and with Nancy's help and the Foundation's help, a number of other labs were located that potential technologies that would be useful. Charlie Cantor's lab had been working on pulsed field gels as a way of separating big DNAs; Hans Lehrach had been working on combination of technologies, just about every technology you can think of, his lab worked on in some way, but he also was working on pulsed field gels; he was working on cloning CpG islands and libraries as a way of getting to the front end of genes; he was working on a number of things.

And so a number of labs got together for a meeting that included those two and included John Wasmuth who was expert in somatic cell hybrids made in the traditional way by fusion. And he happened to be very focused on chromosome 5 at the time, but the idea was let's get him switched over and work on 4. And Dave Housman was there, and basically created a collaboration and said, "Everybody apply their own particular technique and we can get to this disease gene."

MM But that was later. That was around -

JG That was around 86, which is right around when we knew the location.

PH And this was roughly the time when you joined Jim on the project?

MM I joined in 85 -

JG Yeah, about a year before.

MM - and there wasn't really a consortium then.

JG Even in 86; it was just a meeting. It wasn't formal.

MM I joined in January 85. I remember I'd come for an interview and I'd gone over to say hello to Housman and he had leaked this story of the linkage in the hallway [laughs]. And then I saw Jim afterwards who, you didn't become apoplectic [laughter] but I think you were quite surprised. But I didn't arrive actually to start work until 85. So, but I knew that it was linked at the time I came and I said, "I just have to come and work on this project because it's not published yet, but I know there's a place to start on this. And this just looks so doable. This is it." And I had experience, a lot of molecular experience in making libraries, cosmid libraries, and other kinds of large insert libraries, and cDNA libraries, expressed sequences and DNA sequencing, and all this stuff I had wanted to learn I had learnt in Dick Flavell's lab. And so I came in 85 and at that point, Jim, if I recall, one of the things that I

set out to do was almost immediately actually, was to start working with somatic cell hybrids, and making clones and single copy probes from somatic cell hybrids. And there were a number of -

JG Eighty ten.

MM Eighty ten! There were a number of people that, I think were supposed to be working on other projects: Debbie Zucker on muscular dystrophy -

JG Debbie Zucker was on muscular dystrophy; we were using prophase fusion to put markers in the X chromosome.

MM Exactly. And so a lot of these things were sort of really just getting started, and I think the somatic cell hybrid work to try and identify and map on a panel, somatic cell hybrid panel, the locations of single copy markers, was something that I started to work on and to put this panel together. And I remember when John Wasmuth got involved, in a real way John Wasmuth got involved, because John's lab was one of the first, we were making hybrids to segregate from some of the 4p minus and Wolf Hirschhorn… Mary Ann Anderson did a lot of that work; and then I would do the molecular characterisation to find the right hybrid that had the right derivative chromosome, and was it going to be good for the panel? Conrad Gilliam was working on finding a flanking marker and doing the genetic part -

JG Yeah, he was just trying to find new polymorphic markers.

MM That's what he was doing, and it was very clear what had to be done, and it was just a question of getting enough new pieces of DNA from the right pieces of chromosome, roughly, and then doing some running, finding new polymorphisms and doing blots to both physically map them against a somatic cell hybrid panel, and also do genetic analysis. There were also efforts, as you said, to make reference pedigrees with the Venezuelan DNAs for broader linkage. And so it really took probably until 86 or 7 before there was a collaboration of techniques.

JG I think the meeting on techniques was in 86 and then there was a period during which it was unclear whether anything was going to consolidate into a collaboration. It took about a year. And some of the people who were at the original meeting decided not to be part of the collaboration, and others did, and then subsequently other people were invited in to participate, based again on the idea that they had something specific to add whether it be in the genetic area, the physical mapping area or the cloning area. But in the end, as things went on, everybody kind of moved out of their area and into everybody else's area, and it really became -

MM [laughs] As these things usually happen.

JG As they usually do. And it became one of everybody contributing multiply, essentially.

MM It was, I remember John's lab early on, and Hans' lab early on, being quite interactive in the sense that there were people in the lab that really were excited by exchanging things and doing something new with whatever it was you got; and that was terrific. And Conrad, if you remember, went to Hans' lab and -

JG Got us in trouble.

MM - got us in trouble, amongst other things.

JG Some of the other people thought that it wasn't good that he went there.

MM But never mind. It was very interactive; with a couple of the groups quite interactive.

PH Can I just ask: am I right in thinking that one factor which really held up progress from quite a long time was this, in retrospect, fairly obvious misclassification at the clinical level of some individuals

either as affected when they weren't, or perhaps vice versa, which kind of gave this false location at the absolute tip of the chromosome?

JG Right, so the markers, the original marker was in the terminal sub band of the short arm of chromosome 4.

MM As it turned out. [laugh]

JG And the gene was distal to it. So the target area for finding a flanking marker was not that high, but would depend on exactly how close the gene was to the telomere. So as we identified markers through a variety of means, two kinds of things happened: one is that we discovered that just beyond the original marker was a relative hot spot where a lot of recombination happened and once you got beyond that, you didn't see recombination. So you didn't have to travel very much physical distance to actually get to the point of having no recombination of the gene. And the other was that as more and more markers were put farther out on the chromosome, basically they didn't show any recombination except in a couple of people. And I remember specifically that there were particular problems and we in fact very early on made the decision as we recognised something funny was going on in terms of pushing the gene to the tip of the chromosome, not to consider any recombination that involved the non affected. We only would count recombination in someone that the clinician swore up and down had Huntington's disease; classical Huntington's disease.

And we had 3 people who were diagnosed with Huntington's disease: one from Spain and two from the US who showed recombination with every marker that was tested on those. And got them recollected; they still showed it. Got them relooked at; they still agreed they had Huntington's disease. And in the end that pushed us out to the very tip of the chromosome and there was a very large effort to clone the end of the chromosome; Hans' lab ultimately cloned it by using the telomere function for complementation in yeast of telomere function. He made a yeast artificial chromosome based on that telomere. John Wasmuth did a lot of cloning and work in the region. We tried to do polymorphism work in the region and in fact, there were some very interesting polymorphic repeats but it turned out that you couldn't, when you got a polymorphic allele, you couldn't tell whether that allele was on chromosome 4 or chromosome 21 or chromosome 14 because the tips of the chromosomes were very similar, and similarly polymorphic.

MM Peter's lab, Peter your lab did try looking for polymorphisms.

JG Peter's lab actually did come up with the flanking marker.

MM That's right.

JG Finally! [laughs] That pushed it back the other way, that then made us say, "Are these people really affected?"

PH And we also came up with strong disequilibrium for the less distant marker, and lack of disequilibrium at the tip.

JG Yup, so there were a lot of clues along the way that we should not maybe pay so much attention. But I think the fact that it was all cloned at that point -

MM Did it.

JG - and that it was finite.

MM And that it had no genes.

JG Well, that caused it to be looked at intensively and then the fact that it had no genes kind of turned us off.

MM I would like to add at this point that there really was a conscious decision to use two genetic strategies, and I don't know that all members, I think your lab… Russell [Snell] certainly appreciated

it… appreciated these two genetic strategies, one being, as you mentioned, linkage disequilibrium; but the other being association with haplotype, and haplotype analysis.

JG Well of the groups in the collaboration… I said that people did multiple things, but the two groups that did genetics were Peter's group and ours. And the connection with Russell was really -

MM Key.

JG - fantastic because he understood what we were trying to do, and we understood what he was trying to do, and it really meshed well.

MM You see, it's interesting Peter: you heard it at this meeting as well, and I hear it at every HD meeting I go to: that people don't really appreciate the application of these genetic strategies to find the gene. So the statement that a single family was used to find the gene: it can't be done with a single family, and the strategies of haplotype and linkage disequilibrium analyses. It was interesting because members of the consortium who worked very hard around HD to this day still don't understand actually that it was a purposeful set of genetic strategies, I think with the exception of Russell.

JG Right. I think that's true.

MM I mean, Russell for sure, but I'm not sure that other members understand, or understood, that these were specific strategies.

JG Yeah, because once we turned back to the inside, you know the physical cloners in the group obviously found a 2 mega base candidate region less attractive than a 100 kb candidate region on the telomere. But from the genetic point of view, this discovery of disequilibrium within that region did suggest a founder effect for the mutation -

MM And by your lab, Peter.

JG - so it was worthy of further look, and we again, at this point, within the group, we were generating between your lab and our lab, disequilibrium data, but no one had really gone beyond the statistical analysis of disequilibrium. So everybody was, in the

MM Oh, I remember distinctly.

JG - we didn't have it all together. And we had a conversation with Russell where there was a particular probe. I think we asked him to type it originally, and then we compared it on a phone call.

MM Exactly right. I called Russell, or faxed him; remember we used to fax things, Peter? Or faxed Russell and said, "Would you look at this?" I think we sent the probe.

JG It was the proximal marker.

MM It was the proximal marker I'd cloned out of the one the cosmid things, and asked him if he would run it in a BamH1 digest, I can't remember.

JG Yeah, it was BamH1.

MM And he got the data and he called me on the telephone -

JG It was great.

MM And I knew right then. Right there.

JG Right. The light bulbs went off. The way that the data worked: his data were different than our data, but they were different in such a way that you could explain it all as single recombination, historical recombination event among the [s.l. Welsh 1:08:58] having flipped off the bottom part but left the top part intact.

MM Getting, cutting that whole bottom part right off.

JG And that, just seeing it, visualising it in that way said, "You know, interpreting it as if this common region is where this HD gene is really is the best explanation and let's just go after that region."

MM Ignore those 3 people that don't fit: there's got to be something wrong with them.

JG Yeah.

PH So by this point had you actually cloned out the HD gene itself and, so to speak, put it to one side? Or did you go back and look intensively for the gene as well as for the expanded sequence?

JG So you do this one.

MM We had cloned out cDNAs along that roughly half a million base pair stretch of this shared haplotype. This region, this poor region, with a heavy focus on probably half of that, and then with this BamH1 polymorphism we started to ignore the bottom part of it, not realising at the time that the polymorphism was actually within the gene. We just knew that the mutation was distal to it, and so Christine Ambrose had cloned out an interesting transcript IT15, and another one immediately adjacent IT16, distal, and Mabel Duyao had cloned out several more just distal to those; a little small one called IT10 and a bigger one called IT11 that Chris had also cloned out, that turned out to be a kinase, although we didn't know it at the time. And so Glen Barnes had been looking for new polymorphisms; new multi allele polymorphisms, VNTRs, and had cloned out some repeat polymorphisms from the region. He had several; one of which was a CAG repeat. And there was talk at that time about CAG repeats.

JG David's lab had done the RED technique, and actually he had declared that Huntington's disease was not a repeat mutation.

MM That's right. They'd done RED and they hadn't found it; that was that. Actually he announced that at one of the meetings. So David, come on… And so Glen had found this polymorphism, and we knew it was a really good polymorphism and he couldn't get it to amplify very well; it was just a really juicy, rich region. Chris then cloned out a single copy piece of DNA that included the polymorphism and the iniqua 1 fragment that was quite long. If you recall, Alan Buckler had done all this really lovely exon trapping work and that was really the way we came upon cDNAs; that's really how we cloned cDNAs out of libraries. And we had made libraries from an American homozygote to be able to, to be sure to get something that would have HD sequence along the way. And so Chris had cloned out using these trapped exons, and Mabel these cDNAs from the region. The EcoR1 fragment turned out to have this polymorphism on it, and it turned out to have a piece of IT16 on it. And Glen couldn't get the polymorphism to act as a marker because it wouldn't amplify very well in his PCR assay. So Chris set about just trying to get the sequence of this long IT15, which was clearly a very big gene but we didn't know how long it was, and she ended up having to do direct DNA sequencing from a cosmid.

We had a normal chromosome 4 cosmid from the region, and an HD, the HD homozygote cosmid ,and she sequenced directly off the cosmid DNA and she found Glen's polymorphism and not only that, it was different on the HD cosmid than it was on the normal cosmid. And it was the end of her cDNA that she'd been trying to find the 5 prime end of. And she had virtually the entire 13 kb sequence done by hand from stitching cDNAs together, and she couldn't get the 5 prime in. Couldn't get the end of the it from a cDNA library but she did get it by doing this cosmid sequencing. And lo and behold there was this really, not only was it polymorphic, it was amazing. She came into the office with the wet film and said, "What do you think this is?" And it was absolutely beautiful. It was just absolutely beautiful. And I think we went down to your office -

JG Yup.

MM - and said, "We think we know what this is. What do you think this looks like?"

JG Yeah; it was great.

MM So we actually did some Southern blots then and Mabel Duyao got the PCR polymorphism that Glen couldn't get to work, to actually put it into being a real assay, with help. Do you remember? From David Nelson.

JG Oh yeah.

MM Because I had seen David Nelson at a meeting and David said they had been doing these very long repeat stretches because they were working on Fragile X, and they had some really neat tricks. And so I had emailed or called David, and asked him, "Could he send us some of his tricks?" because we were interested in getting some of these interesting polymorphisms to work. I think the term 'interesting' or something. And David immediately sent things and said, "So… why do you need to get this to work?" [laughter] "It is really, in fact, interesting?" And I ignored him, and he took whatever which way. But he was extremely helpful: he sent various tricks with salt and other things for the PCR buffers, and Mabel got this assay to work, and Glen did Southern blots and Mabel did PCR assays. She got the assay to work, and then I remember we had to triage the order of the experiments of where to look. And we discussed: should we look at homozygotes first? Should we look at families with homozygotes because you'd predict if this was right, you know, one thing; and if it wasn't right, who knows, it would just be random. Or should we do things with different ages of onset; or should we do; what should we do? It was, in hindsight, kind of funny because in the end we gave the show away because I, because you were in the cold room looking for DNAs [laughter] Everyone thought that was very strange.

JG Very strange, yes. [laughter]

MM But you were the only one who knew what some of the very, very oldest samples were. And in the end I think Mabel and Chris… Chris was checking the cDNA sequences to make sure this was really all open reading frame cDNA, and also running assays with Mabel. And I think Mabel and Chris ran, over a several week period, hundreds. Hundreds and hundreds of samples.

JG Definitely in the hundreds.

MM Hundreds and hundreds of samples. And it was really amazing. And with every new, do you remember gels you could only fit so many on? And with every little bunch it was like you were saying, we didn't have enough machines to run everything all at once, and so you'd get a little bolus of data and you'd think, "Oh well, it still looks good, but all right, with the next bunch maybe it won't look so good anymore." So it was a process, as you know.

PH But in the end, in the initial paper, you were able to really include a large number of patients and the homozygotes, and really just about everything that needed including, isn't it?

MM And that was really Mabel and Chris. It was, and the fact that the families had been collected; that there were cell lines on most of them -

JG There were so many families that had participated over the years, because, you know, I mentioned with the linkage, the original American family and the Venezuelans, but then subsequent to that we didn't just get families from Susan Folstein, and family's from Ira, we got families from most American sites that saw HD, contributed families; families came on their own, separate from any academic medical centre; and we got families from around the world that were contributed. So there was a lot. I mean, we had cases from New Guinea, and we had cases that were Druse.

MM So the haplotyping to narrow the region involved, in the end we had really good analysis on more than 80 families, and some analysis on chromosomes from probably 120 families. And I think that's an aspect of this whole story that really needs to be emphasized because without all these families it wouldn't… with the technologies possible at the time, you had to use haplotype analysis; and without all those different independent chromosomes, you couldn't have found it.

JG It's worth emphasizing for two reasons: one is I think the family participation is just incredibly important; it's what made the field. And I think that's definitely worth emphasizing. I think the other thing, though, in these days of genome wide snips and everything, haplotype analysis is not appreciated as it should be because it's just coming back, you know. The whole genome wide snip stuff was done as individual sites and a lot of information was thrown away by not thinking about haplotypes. And it's worth noting that, you know, snip haplotypes aren't the first haplotypes that have come along: people did it years before; they did it in HLA for a long time, and it was meaningful.

MM I agree. And I think the other thing I'd like to add is that even today, 20 years later, 20 years afterwards, after this excitement of trying to decide which batch you should run first, we still employ in consortium studies, the same families.

JG Yup; they're still participating.

MM The same samples; they're still participating, and we're still learning new things as we try and go after modifiers. And others; we and others, try and go after modifiers. And Peter, you would laugh, but there, to think perhaps that there are mouldy boxes in the cold room that say 'Welsh 1' and 'Welsh 2'. [laughs]

JG That's right; that's right.

MM And so on and so on.

JG 002, 004, 066. I have those. [laughs]

MM Exactly. So it's worth emphasizing that it's really, it's truly the gift that keeps on giving.

PH Well, Jim and Marcy, we could go on talking about this for the rest of the day [laughter] but I think we've got to stop. Are there any sort concluding things that you feel you'd like to say that need saying, and we haven't covered; or do you think we've given it a good going over? [pause] I think that the silence… it's pretty much covered.

JG It's pretty much covered.

MM I think we've pretty much covered it. I think the thing to bring out as well though but, it's actually come out at this meeting and I really like it, is the fact that the families participating keep participating and you can still keep learning through the participation of these families while adding new families. I think it's augmented by the role of the younger people. I mean, 20 years ago, I was much, I was 20 years younger then -

JG You were the young one; that's right.

MM Actually, I wasn't. I'm older than you are and I'm still older than you are, but I think -

JG I was young too. [laughs]

MM - but I think that what's so nice to see now is the enablement of the younger investigators and the young people in the families, who more and more have come to expect something from the research to feed back into their clinical care and management. They

JG They want a two-way partnership; not a one-way partnership.

MM Exactly. A seismic shift that I think has come about because of the lay organisations bringing people together, meetings like this; but the expectation that research can do that. And that's pretty cool. So I'll stop there, but I think that's really something that's not often appreciated.

PH Well, thank you both very much indeed.

MM Thank you, Peter.

JG Thank you.

MM Thank you for then and for now. [laughter]