Francis Collins

The G Word Transcript Podcast

**Chris:** Hello, and welcome to the G Word. My name is Chris Wigley, and I’m the Chief Executive here at Genomics England. On today’s episode I’m joined by something of a legend in the field of genomics, Francis Collins. We’ll be discussing the original human genome project which he co-led, the hunt for genetic causes or rare and common diseases, his advice to multiple presidents and the future of genomics and even humanity! So let’s dive in...

**Chris:** Francis, welcome to the pod.

**Francis:** Oh, Chris, thanks. It's great to be with you. We got a lot we can talk about.

**Chris:** We sure do. We may need to have an annual slot where we come back to you and pick up on some more topics in depth.

So, tell us a little bit about how you got into this whole field in the first place – you know, Francis, the kid, how did you decide what to study at university, whether to get into medicine, and so on? Give us a little bit of the flavour of what that looked like.

**Francis:** Well, it was a very non-linear pathway, and when I talked to young people, I encouraged them to think of linear pathways as a mistake because you could never be sure what's going to happen.

I grew up on a farm in Virginia, in the United States. Nobody in my family had any interest in science. I was homeschooled till the sixth grade because my parents were really good educators, and they were not happy with the local schools. So, I learned a love about learning. It was just like something new every day, but not so much about science.

Ultimately, I did go to public school and in a chemistry class got excited about science as something that I might want to spend my life on, where you could use these tools of the scientific method to discover things about how nature works. And that seemed something that I might want to invest my life in.

So, I went off to college and majored in chemistry. I ignored biology, Chris, because it seemed much too messy. It is, by the way, but it's –

**Chris:** That’s a fair assessment.

**Francis:** I finished and went off to graduate school in Chemistry. My PhD was in quantum mechanics, but along the way kind of realised maybe I'd missed out on this other branch of science called life science, and I discovered this molecule called DNA, and I did not really appreciate until then - just how elegant and how digital it was. For this part of my brain that wanted to see something that you could write code for, and it was mathematical, like here it was. So, I had this rather significant, ‘oh my gosh, what am I doing?’ moment, decided I'm going to have to change pathways dramatically and study life science in a way that opens as many doors as possible. So I'll just go to medical school. Now, why the University of North Carolina let me in with that story, I'm not sure, but they did, and in my first year as a medical student, it all sort of began to make sense. Yeah, I do love studying the biology of the human, medicine is fascinating, but I still love this genetics thing, so I want to somehow figure out how to put those together. Which seemed a little improbable in 1973 but, you know, things happen.

**Chris:** Amazing story from the fields of Virginia to the labs and the clinics. Lots of people have been on this podcast and described the structure of DNA as beautiful. You are the first one who's described it as digital, which I love. So instead of ones and zeros, we've got A's and C's and T's and G's. Tell us a little bit more about how it struck you as a digital structure, that seems quite a strong insight.

**Francis:** Well, for somebody who loved chemistry, this was the most amazing molecule I had ever imagined to be possible. It was an information molecule.

Certainly, its structure was appealing and beautiful. I will agree with that. The simplicity and the way in which, as Watson and Crick pointed out, it immediately suggests a means for itself. Even now I'm in awe, of just how phenomenally elegant this molecule is. But it is digital, it is information.

It is the way in which all living things actually have an instruction book that allows them to do amazing functions and to pass that information onto the next generation. It kind of does it all. Why wouldn't you want to study this molecule for the rest of your life, but particularly, as somebody now in medical training, maybe you could not just study it, but figure out how you could use that information to reduce suffering and prolong lives and avoid terrible outcomes.

 So that was the dream. Now, back in the 1970s, most people would say, well, you know, you might have some research there, but it's never really going to have much impact in medicine except maybe in paediatrics, that would maybe be a chance of it, we'll see. We could do cytogenetics then after all.

**Chris:** Absolutely. Eric Topple was the first ever guest on this podcast, and in researching that one, I came across his graduate thesis from 1975 on the applicability of genetics in medicine, and he was saying it took a little longer than he was hoping for it to get there, but it got there in the end.

**Francis:** That would be fascinating to read at this point. My PhD dissertation was called semi classical theory of vibrationally inelastic scattering. So, it didn't have many implications for -

**Chris:** A crowd pleaser. So, this is not the place for a potted history of genetics and genomics. But if we very broadly placed the kind of 1980s in a context of, it's around a century or so since Mendel had figured out that there was something, some kind of trait that was getting passed from one set of peas to another set of peas, generation by generation.

We start to figure out that that trait has a physical form in the form of a molecule. By the middle of the 20th century, we've figured out, thanks to Rosalyn Franklin, Crick and Watson, that that's this amazing double helix, but we're still really, figuring out, I guess, the relationship between what's happening in our DNA and what's happening in our bodies. And you were part of that groundbreaking work in the eighties to actually start making those associations in a meaningful way for some of those very Mendelian kinds of diseases like Huntington's, cystic fibrosis, help us get a bit of a sense because now I think we take things like Illumina NovaSeq machines or NovaSeq X’s for granted. You know, you extract the DNA, chuck it in one end and, you know, out comes the bam file the other end, you know, hallelujah, right? But in the eighties, how did you actually do that work and was there an aha moment or is it just the kind of classic stories of the grind for years and eventually you get there? How did you make those associations?

**Francis:** It was a grind, all right. My medical training resulted in my primarily getting interested in diseases that I saw first in the clinic. And when I saw patients with cystic fibrosis during my training, it was so frustrating to see here was a disease that everybody understood was very clearly Mendelian and autosomal recessive, and yet we had not a clue about what the responsible gene was, what its function was, and what to do about it. So, it was all symptomatic treatment, and it wasn't working all that well. So, as I got to the point of being ready to set up my own lab at the University of Michigan in 1984, this seemed like a topic that was worth putting a lot of effort into, although very high risk.

Because at that point, the idea that you might discover the gene that was responsible for a Mendelian disease without knowing what its function was, seemed almost out of reach. The genome was just too big and scary, imagine that you could go wandering in that jungle and actually find something. I even had a picture of myself taken in a Michigan hay loft holding up a needle to try to connect with why this was such a hard job. It was taking long, so we started.

Now the methods that you had at that point, again, graduate students today cannot imagine how little we knew by the way, there was no internet. So, what would you do? You would try to see, could you map the gene to a place on a chromosome? How would you do that? You'd have to look at lots of families. You'd have a set of markers which were very clumsy and were scored using something called a southern blot that most people today have never heard of and which was extremely demanding and very low throughput. Ultimately you would find a variable piece of DNA that tended to predict in the family who had the disease and who didn't and that meant that must be close by the locus on the same chromosome.

By 1985, it was clear that the cystic fibrosis gene was on the long arm of chromosome seven. That was a big advance. Woo-hoo. But it was still a zillion miles away from actually finding the mutation. My lab developed something called chromosome jumping where we could move along a chromosome, not just little itty-bitty step at a time by walking, but we would jump over a distance, which we thought was fairly clever. It was very demanding to implement into practice. And we tried to travel from something that looked like it was close to the gene itself without knowing really how far to go. And you were landing every time in totally uncharted territory. There was no database of the genome for much of anything except a few little islands here and there that had been studied because people had an interest in them, like haemoglobin genes, for instance.

So, it was pretty daunting to take this on and over the course of three or four years, we could see we were kind of moving in the right direction, but we had no idea how much further we had to go. One of the things that happened right then was my main competitor was Lap Chi Choi at the Hospital for Sick Kids in Toronto and he was doing the walking thing and we were doing the jumping thing, and I met him in a meeting and we sat in a lovely sunlit patio and talked about our shared frustrations about how hard this was and decided you know, why are we competing? Let's just join our labs together. The goal is to find the darn thing, and if we can get there faster, we would jump, and then he could walk from where we landed while we jumped to the next one and we could fill out the territory, and we started that in 1987.

Yes, there was an ‘aha moment’. It was May of 1989. Well, Lap-Chee, who became a very good friend, and I were both at a meeting in New Haven. He had set up a fax machine because that's how you communicated in his dorm room. And at the end of every day's meeting, we would run off to Lap-Chee's dorm room, see what data had emerged from my lab and his lab that day. That was the night where here was evidence. There was a three base pair deletion in the middle of an Exxon of a gene that nobody knew anything about, and you saw it in about 70% of cystic fibrosis chromosomes. Not a hundred percent. It's heterogeneity here. That's okay. But importantly, you never saw both copies of that in a normal person. That was sort of the proof that it was actually causative.

That was the so-called Delta F 5 0 8 mutation. And I knew that night that had to be Lap Chi was a little bit nervous because we had so many disappointments. Took another month or so to convince him, and then we published all of that in early September and it was kind of a big deal. It was the first time a positionally cloned gene had actually turned out to be right. And it suggested that we might be able to do this for thousands of other diseases that we knew were Mendelian, but we didn't have a cause.

**Chris:** Extraordinary work. And I love the data being sent by fax. You know, that it's very cutting edge, send the whole genome, that way it might take a bit of paper. And so I guess I'm probably unfairly drawing out little strands from that story. But the strand about we can't do this by ourselves, we need to collaborate, and the strand about, we need to actually map this and then understand how a given person's genome, let's say a cystic fibrosis sufferer is different to a quote unquote normal genome.

If such thing a exists, I can't help but see the seeds there of the human genome project. How did you move from that kind of focus on a given clinical indication like cystic fibrosis to thinking, actually, hold on, big audacious idea, what if we did this for everything? How did that happen and how many photos of yourself with needles or other great kind of communications techniques did you take to persuade people you should do that?

**Francis:** It had already begun this conversation about, should we just go ahead and map the whole genome? That was sort of starting in the mid-1980s, but it was kind of hypothetical. There was a famous Cold Spring Harbor meeting that suggested maybe it was worth doing, but of course it was very controversial, and most scientists were opposed of the idea.

I was a big fan, because it seemed to me this was so brutal, finding just one disease gene and probably one of the easiest ones, given how frequent cystic fibrosis is and how very well behaved it is, as far as its Mendelian inheritance. Suppose we want to tackle something that's not monogenic, let's go after cardiovascular disease or Alzheimer's, or, oh my gosh, diabetes, which my lab ended up working on for the next 30 years.

It was going to be impossible. Basically, if you really wanted to see this little glimmer of hope for one disease get applied to lots of others, we've got to have a better way of doing this. So, I became a very loud mouthed, outspoken proponent of the Genome project by 1989, and just quickly, because I think with hindsight we tend to forget that these things were controversial because everyone sees the Human Genome project as this kind of, you know, tablet of stone moment.

**Chris:** What was the opposition? Was it just, it would cost too much, it'd be too difficult? Like what was the debate? There were sort of four objections. One is it wouldn't be that interesting anyway because most of it’s non-coding, as if that wouldn't be interesting. The junk DNA.

**Francis:** Yeah, the junk DNA you know, just look at the CDNAs and you've got the whole thing. No, not true. Another was, yeah, technically it's just not going to be feasible. It will fail because at that point you had a really big challenge to read out a hundred base pairs in an afternoon. I mean, that was a really good day if you succeeded at that, and they were actually right.

Another complaint was it was going to cost too much, and that there were estimates of what it would cost, but they would be wrong, and this would become a big black hole into which all funding for biomedical research would get sucked and everything else would suffer as a result. And people with grants were particularly concerned about what that meant for them and the most offensive one was that it was just so mind numbingly boring, that only mediocre scientists would want to work on it, and so, by definition, become a sloppy job outcome that nobody could actually trust. That one I found offensive because I was one of the people who wanted to work on it.

**Chris:** That's amazing. Well, it's great to have that sense of, you know, the richness of the debate at the time. So, despite being, you know, mediocre, you sort of ploughed on, somehow blind everyone up around this.

There's probably a whole separate podcast and lots of great books and TV programs have been made about the tensions in the Human Genome Project. You know, Craig Vendor and others, and you are on the public side. And so, for listeners, do check that out. It's an amazing story. I don't think we have time for it today, but I do want to pull out one thread from that, which is around the nature of the data and the insights that came out of that.

There was talk at the time about should we be patenting genes? You know, if a company finds a new gene, should it be able to defend that as intellectual property, or actually does the data belong to each of us as individuals about ourselves, or to us as society about ourselves as a species? Just bring to life for us some of your thinking around that and why you felt it was so important that this was a public resource.

**Francis:** I'm glad you brought that up because I think in many ways that was one of the most substantive contributions of the Human Genome Project. Not just getting the data but developing this ethical stance that the data should be accessible for free to everybody right away, and it was not necessarily going to turn out that way.

In the early 1990s, people were piling claims on snippets of DNA bits and pieces of cDNA so-called ESTs. And it looked as if, if that continued in various ways, the whole genome would be tied up in this incredible meshwork of overlapping patent claims that would make it really hard for it to be used. So, the public project had some substantive conversations about this, and I will say John Sulston and Bob Waterston, who were working on this sequence of the roundworm C. elegans were particularly influential in saying, guys, this is nuts. I know you as scientists are used to having, you know, private access to the data that you're generating so you can play with it and publish papers about it before anybody else does. That's not the right attitude this time.

This is a community project. The community should be getting the data as soon as possible, and there was a famous meeting in Bermuda where we pulled together all of the public sequencing principal investigators and had it out about what was going to be the deal. Craig Venter was there, but he left early, and on the last day, John and Bob stood up with my encouragement and led a discussion about this, at the end of which the decision was made, human genome will be deposited every 24 hours.

**Chris:** Wow. I didn't realise it was that frequently.

 **Francis:** Yeah. As soon as you had an assembly that was more than a kilo base, it went into public domain and that was our best defence against people claiming intellectual property because once it's in the public domain, it becomes prior art, and so we specifically tried to undercut the gold rush that was going on, tried to grab parts of the genome because it was another 20 years before the Supreme Court finally agreed that DNA sequence without a clear connection to a possible product should not be patented. We tried to make that decision scientifically before the lawyers got there.

**Chris:** The whole sort of possession is nine tenths of the law. In this case, dispossession is nine tenths of the law. Right. It's sort of giving it away. I think that is, as you say, one of the enduring legacies of the program. We're kind of 20 years on from the publication of the full first draft.

What are the other elements of that legacy that you see, are there any that you find surprising or that you're frustrated that you still haven't come to light, or how do you see that now?

**Francis:** Well, it has been 20 years, and I think a lot of predictions were made in 2003 about what was going to happen next.

I made some myself, I keep my all my PowerPoints so I can go back and look at them and either feel vindicated or embarrassed. There's a little bit of both. Uh, certainly the tendency is to overestimate the immediate consequences of something like this, and a lot of people did that. Like, okay, when you go to your doctor next year, everything will be different.

Well, it took a little longer than that. There's also, though, a tendency, I think, to underestimate the longer-term consequences, and I think that has applied here as well. Certainly, the lessons we learned there about the importance of team efforts to try to develop these kinds of data sets. Those immediately got translated into lots of other really exciting contributions, starting with things like HapMap, which then became a thousand genomes, trying to really map out the genetic variation picture.

We still have work to do there in terms of more diversity in the DNA samples, and there's a big push to try to do better now, particularly with Africa, but those also were data sets that were profoundly valuable and made it possible to begin to do things like genome-wide association studies, which finally began to open the window into our ability to see what's going on with polygenic disease.

On top of that, I think the application to cancer, which I was really hoping would be an early consequence. There were people who thought, it's going to be hopeless, there's too much noise, there's going to be all of these random mutations in a cancer cell. But we figured out how to deal with that by looking for things that were really drivers and not passengers. Cancer became a huge consequence outcome of the genome project and the sequencing ability.

One should of course mention that the drive to improve the technology, which was really a consequence of the Human Genome Project and which NIH and a few other places put a lot of money into to try to be sure those technologies got advanced, has led to the most significant advance in any technology in human history in terms of speed and reduced cost and accuracy. People often say, you know, there are three things you're looking for in advance in technology. Is it faster or is it better? Is it cheaper? You're only going to get two of those. We got all three for sequencing dramatic ways, more so than anything, even including computer advances.

So that consequence opens up all these doors to the other things we can talk about now, some of which are clinically applicable, a lot of which are just what you can do in the laboratory, where a small lab like mine now can quite readily decide to take on an effort to do single cell biology and discern what hundreds of thousands of cells are up to by sequencing their RNAs and their chromatin and having a very good picture that, oh, that cell is doing this, and that cell is doing that. Unimaginable, if we didn't have this kind of really fast, really cheap technology.

**Chris:** Yeah, absolutely, you mentioned a couple of things there that I'd love to pick up on.

One is around the diversity of the datasets and therefore the kind of applicability of the insights to all of the different communities that we serve. The original human genome was mostly from a guy from Buffalo, New York, a few other bits contributed.

Francis: By someone who was African American, which is interesting.

Chris: You were involved in setting up H3Africa initiative on that note to try and bring that level of diversity into these databases. We at Genomics England have a diverse data initiative to try and make sure that we, our databases, reflect the communities that we serve.

There's recently been both a lot of work and some publications around the human pan genome, so going from a reference genome to multiple, like a few dozen representing different communities. There's a push to get up to 350 or so. This is an area where we could get very technical very quickly about the relationship between a box that you take on a census form about self-identified ethnicity and what's actually happening in our DNA, and so on.

Maybe, just at a high level, I'd love to get your views on when do you think we can say we're done here? What does good look like? Is it that we need to get to kind of 10 billion reference genomes? Is it a dynamic reference genome, which is driven by a machine learning assessment of you and who you are similar to, or different to, or what does good look like? What should we be shooting for?

**Francis:** I don't know that we'll ever exhaust the needs to go even a little deeper. Even if we had the complete genome sequence of blood samples from everybody on the planet, we'd still want to know, well, wait a minute, what's going on with somatic changes that occurred during life, there's a lot of interesting biology as well there, so we're not going to run out. But I think, for starters, what's happened particularly in the last few years with the ability to sequence across those very messy parts of the genome that have long repeats, the centromeres, the telomeres, the so-called telomere to telomere program led by Adam Philippe here at NIH and a bunch of other folks. That's been pretty exciting because there are people out there who sort of began to raise their eyebrows about how many times we celebrated the completion of the human genome. We were tending to do that every two or three years with some slight increment. I think at least when you have one genome that's telomere to telomere, you can court and say, okay, we do have one genome. We can have that celebration.

But the pan genome is another great advance in that regard, and I agree more of those will be better as references. And it does provide input for all kinds of artificial intelligence measures to try to assess what might be there. We haven't measured yet. I just saw today a paper, which is looking at 233 primate genomes and using those to try to assess what nucleotides in the human genome are unlikely to vary without consequences. If you have something that's been conserved, over the course of 10 or 12 billion years and in your genome, it's not conserved, that might be something that would add to your interest about whether that's a variable of unknown significance or whether it might be actually something that could be causing trouble. A lot to learn there, of course.

**Chris:** Wow. That evolutionary dimension is a whole other lens. Right, and I guess another development over the last 20 years since the Human Genome Project has been, as you say, to go from one to a thousand, almost exactly that midpoint. 10 years ago, Genomics England kicked off the 100,000 Genomes Project. ‘*All of us’* in the US are now doing a million genomes. That volume has grown and that's allowed a whole lot of research that we've alluded to. But maybe just to bring together these topics of diversity and kind of public engagement with science and public ownership of data, that ‘*All of us’* program has been admirable in many ways, not least the way that it has engaged with historically underrepresented communities, in the way that it's gone about recruiting and talking to and shaping the program with those kinds of folks.

I was talking recently to Andrea who's the chief data officer of ‘*All of us’*, who's brilliant, and she pointed out to me something, that if you're American, is probably more obvious than it was to me, which is that while engaging with Native American populations, those were actually international treaties because the reservations are not part of US land. Maybe bring to life for us a bit the thinking behind that, because there's so much effort has gone into that, why it was so important to you and the others who put the program together and what you see as the benefits of that kind of engagement and that kind of generation of trust, I guess, with those communities.

**Francis:** Well, I'm glad you brought it up. I'm a big fan of ‘*All of us’*. The idea for that emerged in some of our minds 20 years ago when it was totally impractical. And then coming back about 10 years ago, particularly with some help from enthusiastic support from President Obama who really loved the concept of what this could teach us about precision medicine, and we got it off the ground. But there were many deep discussions about, okay, what is the character of this that is going to be sustainable, maximally useful in terms of its scientific and medical implications? But it has to have the confidence of the public, and so lots of participation, discussions about how do you win the involvement of participants so they don't feel like human subjects, which is a terrible term that we probably ought to abolish, but it's still out there.

Again, a great concern about health disparities in the United States where we have so many different populations, but our clinical efforts tend to focus on white people, and that's just not good. So, let's break the mould there. And the goal was set to have at least 50% of the enrolees racial and ethnic minorities and some people thought, you're never going to achieve that, because there's a lot of distrust in many of those communities, oftentimes with good reason about healthcare. We have achieved that with a great deal of really ambitious efforts to reach out, take the time to listen, not just talk, and then also make it clear that this is a program where you're going to hear a lot about yourself.

This is not a program where you give a blood sample and that's it. You're going to be told, okay, here's what we have learned about you. Here's your genome sequence result. And here is the number of a genetic counsellor, if you think there's something in there you want to hear more about. And we knew going forward there were going to be people for whom serious mutations with actionable findings were going to turn up and we wanted to be sure those people had what they needed to follow up.

We also made sure that the people involved were consented for recontact for all kinds of other follow-up studies that might turn out to be quite useful and could be done much more efficiently with an already motivated group of a million people than what traditionally happens, at least in the US where it's like three or four years before we even get started with a clinical trial. And all of that I think has come to pass pretty well since I'm a member, I get the communications from ‘*All of us’*, so I know I'm a participant. I know what they know about me, and they've been pretty good about sharing all of that with me. And I hit my genome sequence now because they have done and released to researchers now, 245,000 whole genome sequences, which I think at the moment, probably won't last long, is the largest collection of whole genome sequences in a public project. Researchers, that's the other thing, needed to have access to this, but in a way that protected the privacy and confidentiality and that got worked out pretty effectively I think.

There are thousands of researchers who are now authorised to go digging into the dataset. What I'm particularly excited about now is the opportunity to layer other studies on top of this. So, for instance, I'm fascinated with this variant in APOL1 that turns out to be very frequent in Africans and African Americans, and was a protective variant against traps, but now is associated with a high risk of kidney disease and may account for the fact that African Americans are very much overrepresented in dialysis units because of this genetic risk factor. There may even be a path forward to coming up with a preventive treatment for that. So that's the kind of thing ‘*All of us’* is just beautifully situated to be able to look at.

I also think because we ask these people to share their electronic health records, to fill out all kinds of questionnaires, to wear Fitbits with all that data being collected. We're learning a lot about what happens when you don't just have a diagnosis of a disease where you're just walking around. How can we practice prevention more effectively? So much of what we've done in medical research is to focus on disease. Let's have a bit more of a focus on how to stop it from happening.

**Chris:** Yeah, absolutely. Those elements that you've mentioned there. Absolutely inspiring research efforts, the original name for ‘*All of us’* was the Precision Medicine Initiative.

When we think about the impact that work, the genetics and genomic medicine, which has already been practiced in the States, in England and other places, what do you see as the kind of state of play in genetic or genomic medicine today? Where do you think we should be trying to take it next?

**Francis:** Yeah, well, I think it's had profound impact on cancer. I think at this point, anybody who has a new diagnosis of cancer would want to have their cancer genome sequence to see what are the driver mutations, and then try to do the best job of picking the therapeutic intervention that's going to be specific for your cancer and not some general generic kind of cancer that is already, I think, becoming almost in most places the expectation. That's pretty profound, and it's one of the reasons I think we're seeing deaths from cancer continue to drop a little bit each year, and that we can promise that they should drop even more substantially.

I think in newborn screening, particularly for kids who have uncertain diagnoses, this has been a profound advance that has made it possible to get answers sometimes in a couple of days that in the past might have taken four or five years of intense frustration, so that ought to be available, I think pretty much everywhere, although it's not quite yet.

I think the big challenge is what about the average person who seems to be doing okay? The 35-year-old who goes to their doc for a regular checkup. Has genomics already changed the way that that medical care is being approached? You could say, yes, in terms of things like vaccines, because genomics is also the reason that we were so quick in getting a Covid 19 vaccine, but in terms of common disease which is, again my lab mostly works on is diabetes, where I think we're getting close to the point. Where things like polygenic risk scores could become really actionable as far as giving each person a personalised recommendation about how to stay healthy instead of the one size fits all, which people mostly ignore, and I think *All of us* in Genomics England and the Biobank that are all sort of collecting this kind of data are going to put us in a better position to be able to say, have we reached that point where that information is going to actually result in better outcomes? We just need to have that data. I think, before we rush into it.

Certainly, when it comes to surveillance, I'm just as excited as a lot of people are about cell-free DNA and whether that's going to become the dominant way for doing cancer surveillance, but I don't want to go there prematurely in case it turns out that you cause a lot of anxiety and a lot of follow studies and you don't actually help anybody.

So, the question is, does that really work for very early cancers where, you know, early diagnosis may be a way in which you can save lives? That's not entirely clear yet. We need better data on that front.

**Chris:** Well, we're doing 140,000 people in the NHS-Galleri trial in the UK as we speak. Som we're going to get a read out on that pretty soon.

**Francis:** Watching that closely.

No, pharmacogenomics I think is the other area that I had hoped by now maybe we'd be a little further along, because there are certainly compelling circumstances where knowing your genotype is going to optimise the choice of the right drug at the right dose at the right time. But it hasn't found its way quite into general practice.

Again, I think a project like *All of us* maybe will help with that. Everybody's getting their pharmacogenomics. I know that I'm a slow metaboliser for a particular gene that I probably would want to know about if I ended up needing a particular therapy. But you know the problem, Chris, and it's frustrating that we haven't really made much headway on these, studies have shown that when you have research studies that demonstrate something is clinically beneficial, the timetable before it actually is implemented in the standard of care is about 20 years. That's just not acceptable. It is not acceptable.

**Chris:** There are these apocryphal stories that were true with this stethoscope, but I have no basis of knowing whether that is true or not.

So let alone something like polygenic risk score. I think I'm right in saying that the participants in ‘*All of us’* get the polygenic risk score as effectively as a PDF. Right, and it's like you take it to your physician and say, here is my pharmacogenetic set of insights. I agree with you. I think there's always more science to do, but at least for a proportion of drug gene pairs, the science is there.

Actually, getting that to a physician's fingertips at the point of diagnosis and at the point of prescription is just insanely difficult in terms of systems and integration and all of these wonderful things that we have to wrestle with. Interesting.

So, we're starting as we talk about this, to build a picture of pieces of, let's say, emerging insider, emerging benefit from genomics in healthcare. So, we're starting to see a path towards pharmacogenetics, pharmacogenomics, helping us prescribe better medicines to individual people. You mentioned things like circulating DNA is where we can just look at someone's blood and say, oh, you know, there's maybe a pre-cancerous thing happening here.

We're starting to get to technologies that we haven't touched on yet, but I'd love to, around things like individualised cancer vaccines, so developing treatments, therapeutics off the back of your specific tumour, if we sort of lift our heads and think maybe 10, even 20 years in advance, what kinds of benefits should we be pursuing at this point? What does the next wave look like?

**Francis:** Well, we touched on some of them, but we haven't talked about gene therapy, and we should. Because, after all, there are some 7,000 genetic diseases where we know the specific DNA mutation, but only about 500 of them currently have any kind of therapy. We should be making every effort to change all of that, and you can see it's starting to happen.

Sickle cell disease, which I studied intensively as a postdoc, it's now a curable disease. If you're lucky enough to be in a clinical trial that's using one of the Crispr or the gene therapy vectors, and you have, you know, some way to pay for the 2 million dollars or whatever it's going to cost to get you there, and if you're willing to go through a very toxic experience over a month of having your bone marrow ablated in order to make room for the cells that have now been genetically corrected. We got to come up with a more extrapolatable fashion way to do that, especially since most people with sickle cell disease are in West Africa and not in the United States.

We should be working on that. My dream, Chris, is that we come up with, in the next decade, a scalable approach to every genetic disease where you know the mutation. And it involves the apparatus, which is going to be some kind of a prime editor built upon Crispr and look at the stuff David Lewis is doing. It's just breathtaking.

The hard part's going to be the delivery system. How do you deliver that apparatus to the right tissue safely and a one-time intravenous infusion? That ought to be the goal without any need for ex vivo. This is in vivo and you get the result you're looking for safely and you have to get to the brain too, because a lot of these are conditions where the brain is affected.

That ought to be our goal. I don't see why we couldn't get there. I don't think the delivery systems, will be other kinds of viral vectors. I think there'll be something, maybe like a nanoparticle that actually is less likely to have an immune system response. But that ought to be such a goal. I mean, when you add it up, even though these are all rare diseases, that's a lot of people. So that's one thing I want to see.

**Chris:** That's a big shift, right? From diagnosis to therapeutic development on the right. On the basis of those insights, could you see a world in which, you know, at the moment in loads of hospitals, you have diagnostic labs, right? We get some blood from the patient, we send it down to the lab on the next floor, they extract the DNA, they do something, they sequence it. Could we get to a world in which we're manufacturing therapeutics in the lab next door? You know, we build an extension on the side of that wing of the hospital where we start churning out these individualised therapeutics.

**Francis:** Yes, I think that ought to be the goal as well. Regulators will, you know, have a little bit of anxiety about this, and it has to be done so that it really is high quality GMP material. But I don't see why that could not be done in a distributed way, and it would need to be for so many rare conditions that you want to be able to treat right at the site where the patient is.

So, yeah, let's try for that too. In terms of cancer, I do think we're going to make a lot of advances there and I'm glad you brought up cancer vaccines with mRNA vaccine platforms. Now we're already starting to see how you can do the cycle time so much faster, so you don't end up with a cancer vaccine, but the patient's already dead this time. You got a cancer vaccine you might have in three months or even less. It would have the chance for that patient to really fire up the immune system to go after it. I think we're going to start to see some pretty interesting solid tumour metastatic disease examples where this actually works. At least I'm going to be very disappointed if that's not the case.

Then for all of the other common chronic diseases, heart disease, diabetes, let's not forget, we need to do more about prevention. And again, part of that will be coming up with prevention schemes that are specific for the individual which are more likely to have people pay attention to and some of them will be new insights about pathways that are coming out of things like diabetes. You know, 15 years ago we knew about two genes that played a role in type two diabetes risk. As of right now, we have 622 loci in the human genome that carry a variant that confers risk. Most of them non-coding but pointing us to really interesting pathways that we did not know about before, many of which are actionable.

**Chris:** And let me push you on the common disease piece because we've talked about using genomics for large scale, insights about wellness prevention, personalised approach to life. We've talked about curative therapeutics and manufacturing, those for rare diseases, Mendelian diseases, cancer. If I join those lines, do you think we can also get to a point where we can be generating genomically driven curative therapies for common diseases, cardiovascular, diabetes, you mentioned, or even neurodegenerative in later life?

**Francis:** I think we should be able to do that. I mean, and we've kind of got examples from a whole lot less data. I mean, look at statins, which have undoubtedly prevented oodles and oodles of heart attacks and strokes which came out of genetics. Remember that was how we learned about that, actually from studying relatively rare genetic disease familiar hypercholesterolemia. If you look at people with sort of polygenic cardiovascular disease and do a GWAS (genome-wide association study), yeah, the target shows up the HMG-CoA reductase, but it's not a strong signal. That should tell you that all these other GWAS that have lots of signals, somewhere in there, there might be some that are gold and that's what I think we're all trying to discover right now. I think that will happen. For diabetes, I hope for Alzheimer's disease, that we get beyond a total focus on the amyloid hypothesis, although it's gratifying to see we're getting some benefit finally there. In terms of clinical interventions, as long as you start really early, but there's got to be others synergistic targets that are under investigation coming out of genomics that perhaps with combination therapy we could do even better.

**Chris:** Amazing. I'm conscious we only have a couple of minutes left before you need to. Go and put on a cape and advise presidents and do other things. Let me maybe bring us back to kind of the human condition for a second, because if we think about this world whose outlines we've just sketched, where there are newborn sequencing so that potentially many most babies are born with their DNA on file. Maybe it's appropriate to look at different insights from that at different points in your life. There's no real benefit as a newborn baby to understanding your Alzheimer's risk necessary. We have stronger diagnosis; we have individualised wellness plans. We have personalised therapeutics over the course of one's life. That all sounds great, right? You know, healthy and golden age, but I was having a really interesting conversation recently about, at the moment, there are kind of two categories of people. If you want to look at this way, sick people and well people, and we're going to live in a world where it's a lot more shades of grey than that, right?

We understand some of our risk factors where maybe somewhere on a spectrum, on loads of different dimensions, between being well or having a sort of clinical level of issue. How do we best live rich and fulfilling lives? Not anxious and kind of stressed lives in that world as humans?

**Francis:** I think the data would show that people adjust pretty well to getting information about their future health risks from genetics. We've done a lot of those studies initially. If somebody hears they had an unexpected risk, there's a little bit of a time of being shaken up by it, but you accommodate. I think our goal, as you just outlined, is a noble one.

It is not just to extend lifespan, although that would be nice, and I don't see us likely to get much beyond a hundred or 110 years anytime in the near future, but also health span where people have the maximum opportunity during their life not to be in some way knocked back by a chronic illness that makes the quality of life and the ability to achieve what you want to achieve a whole lot more difficult. If we could actually focus on that, that would be a good thing, but we can't just focus on it for the rich countries. The meeting I'm about to go to is a meeting to try to see how we could inspire the development of centres of excellence in genomics in Africa, and I think the time is right for that.

We are citizens of the whole planet, and if we care about our planet's future, we have to care about all those other folks out there who need these kinds of advances to happen to them as well. By 20, 30, 40, 2% of the people under age 30 will be in Africa, and we need to be sure that they have economic opportunity and a job opportunity and a health opportunity to flourish, otherwise we've not done our job, so let's not forget that that's part of the whole enterprise. We get kind of focused sometimes about our own neighbourhoods: make the neighbourhood the planet.

**Chris:** There you go. On that note of universal human flourishing, there's no better note to end on. Thank you so much, Francis, for taking the time and thank you for everything you've done for, let's just put it boldly, humanity over the course of this incredibly rich career you've had.

Thank you so much.

**Francis:** Great to talk to you Chris, and thanks for everything you're doing for Genomics England. It's great to talk to you.

Chris: Thank you to Francis for joining me today. If you enjoyed this podcast, please do give us a 5-star rating, as it helps others to find the pod. And if you’d like to hear more like this, please subscribe to the G Word on your favourite podcast app. Thanks for listening.