**Genomics 101 transcript: How do pharmaceutical companies use genomic data for drug discovery?**

**James Duboff**

**Naimah:** How do pharmaceutical companies use genomic data for drug discovery? Today, I’m joined by James Duboff, a Strategic Partnerships Director here at Genomics England, to find out more.

So James, first of all, what is genomic data, and how does this relate to our genes?

**James:** Let’s start with a simple explanation of what we mean by genomic data and our genes. So, every cell in our body contains a complete copy of our genome. Now, genome is kind of a mini instruction manual that describes exactly how to make you. Now, those instructions are written in a language called DNA, which is over 99 percent identical in every single human on the planet, so you and I are actually genomically very, very similar. The differences, however, are called variants, and they’re what make us unique. Now, some of those variants can actually be very dangerous, and they can code for things like rare genetic diseases or even cancer. So, we need to read in detail exactly what’s going on in your DNA and in your genome to see where changes are and where those variants really are, and we do this by sequencing the genome. So, if you get a DNA sequence, that’s effectively an electronic readout of your genomic data, which is your genome in computational form.

Now, understanding that and working with that is still a relatively new field, so what we try and do is connect the genomic data, your genome, with health information, such as hospital records and what you’re presenting with in clinic, if you’re in a patient setting, and look at those together to give context to those variants in the genome. So, genomic research is actually where we look at how genes and physical outcomes could be linked. So thinking of, you know, biology and physiology term, what does a variant exactly do and how might it cause a disease.

**Naimah:** You mentioned both the genome and whole genome sequencing, and if our listeners aren’t too sure exactly what they are, they can listen to some of our other explainer episodes with Greg Elgar, who explains these concepts. So James, next could you tell me why are pharma and biotech companies interested in genomic data?

**James:** Ultimately, pharma and biotech companies are interested in genomic data because that really tells them what’s going on within the blueprint or that mini instruction manual of an individual. So, pharma and biotech have dedicated research teams that focus on genomic research, and they look through genetic databases across the world, such as Genomics England and others, to really understand the role of the genome in their target disease areas. By looking at those, that helps them develop new drugs and tools to specifically diagnose, treat and also even cure these diseases.

**Naimah:** So, how exactly do they do that? Can you explain it in some simple steps?

**James:** I think there are four key areas that they need to focus on. So, starting with the first, where, whereabouts on a genome should they focus? Now, the way that a pharma company would do this, or any researcher really, is by taking two populations of people. So, you’d take a population who have a known disease, and you’d compare that to people without. Now if you’re looking at the genomes of people with the disease and those without the disease, you can kind of play spot the difference between those two, and understand whereabouts on the genome variants appear for the disease population and not for the healthy or undiseased control group. Now, when you do that, you can kind of pinpoint exactly whereabouts you see variants only in that patient population. That helps you identify your target, and that’s known as target identification, which is essentially pinpointing that spot on the genome that’s linked only to the disease. Once you know that, you can use that as a potential target for a new drug.

So, once you’ve found that variant, the next step was, what does that variant do? Is it potentially overproducing something? Is it activating a promoter and therefore making more and more and more of a gene product that, you know, might be toxic inside a person if you have too much? Even too much of a good thing could be a bad thing. So, is that the case? Or does that variant cause an underproduction or something to just be not actually made by your body at all? So, if that variant kind of interferes with a piece of genetic code, it could stop that gene from producing anything, and therefore you might be effectively detrimented and deprived of that particular gene product. And both of those, an overproduction or an underproduction, could lead to a disease. So, to understand that in more detail, you might need to look at gene products as well.

The next step, once you know whereabouts in the genome you’re looking and what exactly a variant does, the next step really for a pharma company is how could you fix that. So, if you’re looking at too little of something – so, if a variant stops a gene from actually developing into a gene product then you might need a drug to boost or to compensate for that, so potentially a supplement or having some kind of drug that can get the body to make more of that product. If on the other hand your body is making too much of something in a way that could be toxic, you kind of want a drug to reduce those levels, so a drug that could potentially breakdown that gene product so that you don’t have too much of it, or stop it from working effectively, so that it doesn’t seem as if you have too much of it, or otherwise prevent it being made altogether.

Now, one example of this prevention is actually a gene silencing drug, or an ASO, as they’re effectively known, which can be used as a genetic mask. So, that sits on top of a gene and hides it, so the body can’t actually make that dangerous varied gene product. Now, if you’re going to make something like that, you need to be absolutely sure that masking that entire gene and stopping even a varied form of it isn’t dangerous, so that last step really is making sure that your drug is safe and wouldn’t cause any other issues. So, traditionally, that would have been done using animal models as kind of a surrogate organism, but now using genomic databases, you can use human genomics as kind of real world examples of applying say a genetic mask and hiding an entire gene or genetic section, and you can look through genetic databases to have a look for individuals who are alive and hopefully healthy in the population, who don’t express a certain gene. So, if you can find people who are healthy, who don’t have that gene or have variants that stop that gene from being produced, you kind of can be confident that you can make a drug to cover that and it would be considered safe.

**Naimah:** Okay, so that’s really interesting. So, what you’re saying is, by using human genomic data, we can test the impact and safety of gene targeting drugs directly in humans.

**James:** Yes, exactly. So, you can ask that question of would hiding that gene entirely cause any other health issues or any adverse effects really from a drug that hides it. And the really useful thing about that is that we’d know the impact of a gene targeting drug before you’d say start a clinical trial, so that really stacks the odds in your favour of the drug working safely, which is really powerful for a drug company that would otherwise invest a lot of money in a clinical trial that could be a risky endeavour for the company and also for participants. So, this is very useful for patients, and also fundamentally it’s a lot more useful for a company to be assessing safety using humans and human genomics directly as opposed to using a surrogate organism like a mouse, which many people would argue is not a good reflection of what would happen in humans.

**Naimah:** Can you tell me briefly if genomics can be applied to other stages of the drug delivery pipeline?

**James:** Yes, in fact genomics can be applied all along the drug discovery and development R&D pipeline. So, as an example, biomarker identification. A biomarker is a biological product or a chemical signal that’s associated with a disease, that you can find and monitor inside the

body. So, you can look at an increase in that biomarker or a decrease in that to monitor whether a drug is working as you’d expect. Is the drug increasing levels of something being produced, or is it decreasing that product being produced? And you can use that to understand whether it’s possible to potentially develop that treatment, would that treatment actually work. So, that’s really important in monitoring drug impact and also understanding clinical endpoints for a trial.

You can also look at biomarker identification to look for genes and variants that are associated with a disease that could help you understand who best to enrol in a clinical trial. So, clinical trial recruitment is another key area, where if you involve the genome in your enrolment criteria, you can essentially just recruit the most suitable people where you know the drug will work best, and also you’re sure that the drug would be most safe and effective at treating their condition. And then actually to go a step further on the clinical trial point, clinical genomic datasets are actually really useful, if you think about it, in the opportunity to recontact participants too where they’ve consented. So, what I mean by that is, a pharma company could directly find and recruit optimal patients with either a rare disease or a cancer where their drug would help most, based solely on their genome, and that’s a really, really exciting point, because that offers the opportunity for pharma to both develop a drug based on that genomic dataset, but also then deliver the drugs to treat those same exact people.

**Naimah:** So, how do pharmaceutical companies access this data?

**James:** Well, there are different datasets, and each different dataset has a different population within those, and each of them have their own consent models and governance rules on how that data can be used and who can access it, and how they access it. So, some of these datasets just hold genomic data, while others would have additional biochemical data and also health information potentially on participants. So, depending on the different types of data, there’ll be different access limitations and restrictions. So, some entities and some datasets can be simply downloaded, and that could be very useful for pharma and biotech companies, because that means that they could use them inhouse. Other datasets and groupings of genomic data and libraries of sorts would operate a TRE or an SDE model, so that stands for a trusted research environment or a secure data environment, and these are essentially – you could consider them as libraries, like a reading library, where you can come in and read the books but not take out those books, or genomes in this case.

**Naimah:** Can you tell me, what impact does the use of genomic data for drug discovery have on the public or patients?

**James:** Oh, there’s huge impact on drug discovery, and ultimately genomic research really helps drug companies make better treatments for patients and the public. So, we’ve already seen the benefits of genomics used in drug discovery, and I think we will do more and more as DNA sequencing is used more in clinic, and also that’s going to keep happening the more cost keeps dropping, which is making genomic medicine really and genomic healthcare increasingly feasible at scale. So, 20 years ago, it cost over £100 million and it took years to sequence a genome, but today you can sequence a genome within a few hours for under £1,000.

**Naimah:** What are the benefits of having your genomes sequenced in a healthcare setting?

**James:** Ultimately, genomics enable a faster and more accurate diagnosis. That enables early intervention, which can really maximise the treatment impact and improve outcomes. So, what I mean by early intervention, if you can give a drug before someone shows symptoms then you could prevent them ever getting the disease, so that’s moving towards preventative medicine, which is really exciting and absolutely enabled through genomics. So, genomics really help pharma companies make also better drugs and target the underlying disease directly rather than just addressing symptoms, so this helps them make more effective and safe treatments to really improve overall outcomes for patients.

Another thing to think about is that some drugs are already on the market but used for different reasons. Genomics can help pinpoint the root cause of that disease within a genomic setting, so that can highlight repurposing opportunities for existing drugs. Now, existing drugs are those that have already been proven safe in humans and approved for use, albeit potentially in a different setting. Now, if a drug could be shown by genomic research to be targeting the same root cause within the same biological pathway, they could very easily be repositioned and applied in an entirely new disease.

So, I guess to finish, through genomics, drug development can help us move towards precision healthcare, and by that I mean making targeted treatments for specific patients. That will be far more effective and have significantly fewer side effects. In the case of participants in clinical genomics sequencing programmes open to researchers, that also means matchmaking opportunities for companies to diagnose and treat unique patients. In the case of ultra rare conditions, that means they can create a treatment specifically for that patient and then work with their doctors to deliver the brand new drug just to them, to ultimately save lives.

**Naimah:** That was James Duboff explaining how pharmaceutical companies can use genomic data for drug discovery. If you’d like to hear more explainer episodes like this, you can find them on our website at www.genomicsengland.co.uk. Thank you for listening.