It's In Your Blood
Season 2 Episode 1 Transcript
The Team: Alex Nguyen Ba, Adrine de Souza, Moez Dawood, Kortni Kindree, Katie Partington, Ziyi Dai, and Jerome Freudenberg
Facilitated by the Altas of Variant Effect Alliance

Season 2 Episode 1 Transcript

00;00;00;04 - 00;00;28;18

Alex Nguyen Ba: Blood has always fascinated us. In the ancient world, it was seen as the very essence of life, the center of health and disease. Even today, we've linked it to our genes by calling it “bloodline” as a way to trace our family roots. Myths and stories are filled with blood. We depict creatures that need to drink blood to survive, or show rituals where blood is shed to seal unbreakable oaths.

00;00;28;21 - 00;01;02;13

But have you ever ask yourself why there's so many myths surrounding this? Could it be because so many of our most mysterious genetic diseases are tied to blood disorders? Setting aside cancers that affect blood cells or infectious agents such as malaria and HIV, close to 2% of Americans are affected by blood disorders. That's millions of people whose lives are affected every single day by the way their blood cells form, function, or fail.

00;01;02;15 - 00;01;26;28

And while people suffering from blood disorders are living longer than ever before, thanks to countless hours and efforts from scientists and doctors all around the world, rapid diagnosis is still key to improving the quality of life and reducing the health effects of complications. And part of what makes blood so captivating is how universal it is, even amongst animals.

00;01;27;01 - 00;01;56;23

Every single heartbeat is pushing our blood through our bodies, delivering oxygen and nutrients that keep us alive. At the same time, its main components, the blood cells, have a fragile balance. Immune cells must be kept in check, and other processes such as blood clotting have to be controlled. For scientists, blood is a puzzle that has held fascination for centuries due to its remarkable ways at which we can study it.

00;01;56;25 - 00;02;22;24

We can see the cells under a microscope. We can extract the cells, reinfuse them, and even see their development. All of that in the promise of treating those that suffer from a blood disease. But here's where the story takes a turn. What if we could do more than treat these disorders?

What if we could cure them? You might have heard of patients being cured from HIV after a bone marrow transplant.

00;02;22;26 - 00;02;47;22

How does that work? What if there was a way to do something like this? Some form of new technology. But for inherited blood disorders like sickle cell anemia.

[Intro music plays]

00;02;47;24 - 00;03;12;11

Welcome back to the VUS podcast, a series initiated by the outreach effort of the Atlas of Variant Effects Alliance. I'm Alex, a professor at the University of Toronto. And today, the team Adrine, Moez, Kortni, Katie, Ziyi, Jerome, and I will walk you through how MAVEs are advancing the interpretation of blood disease causing variants. This episode marks the first of the second season of The VUS podcast

00;03;12;11 - 00;03;48;11

with many great interviews coming up. Today, our two hosts, Jerome and Katie, will walk you through how modern technology has been harnessed to modify genes within blood cells by coupling this technology with basic science that brought understanding of health and disease, blood function and development. Scientists have developed surprising cures for age old genetic disorders affecting our blood. As always, we would like to note that this podcast focuses on the realistic portrayal of the basic and translational research behind rare diseases, with the hope that it will lead to better patient care

00;03;48;13 - 00;04;08;14

but this podcast is not meant to be medical advice.

Jerome Freudenberg: Hi everyone, I'm Jerome

Katie Partington: and I'm Katie, and today we're going to be learning about blood diseases and how doctors have used our understanding of how blood is formed in our body to treat common and rare blood disorders.

Jerome Freudenberg: Did you know that some genetic blood disorders can now be cured?

00;04;08;17 - 00;04;27;02

Katie Partington: But how can scientists do that? All of our cells carry the same germline variants. And so let's say I was born with sickle cell anemia, a disease that alters the shape of red blood cells. Would you have to remove all my blood? And then somehow, I guess, correct the variant that's causing the disease?

00;04;27;05 - 00;04;47;29

Jerome Freudenberg: Well, not quite. And this is why it's not only important to know the mutations that cause disease, but also understand how the disease occurs. To learn more, we invited a guest to teach us about blood diseases, development, and how modern technology is enabling us to cure genetic disease in grown adults. Thank you, Doctor Sankaran, for joining us today.

00;04;48;02 - 00;04;57;24

Can you begin by introducing yourself for the audience?

Vijay Sankaran: Hi. I'm Vijay Sankaran from Boston Children's Hospital and Harvard Medical School, as well as the Broad Institute.

00;04;57;26 - 00;05;33;22

Jerome Freudenberg: So your lab studies how human genetic variation influences blood and immune cell production and their relationship with human health and disease. Can you tell us about how you got started on this?

Vijay Sankaran: Yeah. Thanks so much for that question. Maybe I'll talk a little bit about things that I did actually about 20 years ago or so as an MD PhD student. I got very interested in a problem of why the patients with the most common monogenic disorders in the world, and this includes sickle cell disease and beta thalassemia, have such variable clinical presentations.

00;05;33;25 - 00;05;46;11

We knew that there are some patients who were really severely affected and there was other patients who were very mildly impacted, and this had been an observation that had been studied for decades.

00;05;46;13 - 00;06;08;16

Katie Partington: This variable clinical presentation is also called variable expressivity, and means that a specific mutation might not have the same effect in all people. The effect of a mutation can be influenced by a variety of other biological traits that differ between people, and which are broadly called context. The effect of a mutation might depend on the context.

00;06;08;19 - 00;06;14;06

Jerome Freudenberg: So in this case, were you able to identify a reason for the difference in clinical presentation?

00;06;14;09 - 00;06;37;03

Vijay Sankaran: One factor that kept coming up was the amount of fetal hemoglobin that patients had. And it turns out during human development there is a switch from a fetal form of hemoglobin to an adult form of hemoglobin. This occurs during normal gestation and shortly after birth. But it's the adult hemoglobin molecules that are actually mutated in sickle cell disease and thalassemia.

00;06;37;06 - 00;06;51;15

So if you have more of the fetal form of hemoglobin, you can actually compensate for either the mutated hemoglobin in sickle cell disease or for the defective production of beta hemoglobin in beta thalassemia.

00;06;51;17 - 00;06;59;22

Jerome Freudenberg: So basically, if we could block the switch or the repression of fetal hemoglobin, we could essentially have a therapy for sickle cell?

00;06;59;24 - 00;07;21;25

Vijay Sankaran: It was that idea that really got me excited. And it turns out that during those decades of work setting this process, people had actually developed small molecule therapies for sickle cell disease, not curative therapies, but therapies for sickle cell disease. This was

hydroxyurea was really developed with the idea that that could boost fetal hemoglobin levels in patients with sickle cell disease.

00;07;21;27 - 00;07;46;22

And so I actually met some patients when I was a medical student who had sickle cell disease and thalassemia and saw the impact that hydroxyurea could have on them. And at the time, I was working in a lab that had studied this decades before, but had moved on to other problems. And so suddenly I was like, wow, this is an amazing opportunity to use all these emerging technologies to try to study this problem.

00;07;46;24 - 00;08;13;21

It was probably a little bit of naivete. And as chance would have it, I spent years failing during my PhD. But I was really fortunate because during the tail end of my PhD, there had just been the first genome wide association studies being applied. It turns out that we got involved in one of these genome wide association studies for fetal hemoglobin levels, and what underlies intra individual variation in fetal hemoglobin levels.

00;08;13;24 - 00;08;39;18

And we were able to show in a cohort of patients with sickle cell disease that these same factors that affected fetal hemoglobin levels in non-anemic individuals actually also impacted the variation in fetal hemoglobin levels in these patients with sickle cell. This then resulted in the identification of a factor BCL11A, which was an important regulator of this fetal to adult hemoglobin switch.

00;08;39;18 - 00;09;04;09

so if you remove BCL11A, it's a key repressor of the fetal hemoglobin genes. And then you get reactivation of fetal hemoglobin.

00;08;47;21 - 00;08;54;10

Jerome Freudenberg: That's in principle, right? Have people tried this reactivation since you and your colleagues made this discovery?

00;08;54;12 - 00;09;09;27

Vijay Sankaran: Fast forward now 15 plus years from when we made those initial discoveries. This is now the target of now what's an FDA approved therapy for sickle cell disease and thalassemia Casgevy. And so that's incredibly exciting to see.

00;09;10;00 - 00;09;18;18

Jerome Freudenberg: That's really incredible. It really is so exciting to see how much progress has been made from initial gene identification to treatment now.

00;09;18;21 - 00;09;30;03

Vijay Sankaran: As I look back, it's just hard to envision how this could have happened. But it's really incredible to just see the things that we had done lay the groundwork for such incredible developments.

00;09;30;05 - 00;09;50;15

Jerome Freudenberg: All of this came from observing a variant in the human population. Understanding what that variant did to the gene, and finally having a good grasp of the molecular biology that underlied the effect of that variant on the phenotype. But if I understand things correctly, this is all stemming from the discovery of a single variant from a genome wide association study.

00;09;50;17 - 00;10;02;03

As you know, our podcast tries to place human variation in the lens of these multiplex assays to produce variant effect maps. Before the show, you told me that you're looking into this. Can you talk a little bit about that?

00;10;02;05 - 00;10;23;18

Vijay Sankaran: We're really bad at prosecuting how variation impacts the biological processes, and that's really motivated us to think about how we can scale this up. Could we actually scale that up in a way that we could think not just about doing one variant that, you know, undoubtedly has had an impact on patients, but could we actually scale this up?

00;10;23;23 - 00;10;30;11

It's frustrating on the other end to see this only apply to a single variant to a single disease. And so could we do better?

00;10;30;13 - 00;10;52;26

Jerome Freudenberg: So doing multiplex assays helps us go from a single variant association to connecting many potential variants to many diseases via their biological processes, even if they don't show up in the genome wide association studies. But the effect that a variant has on a cell depends on what type of cell we are looking at. In this case, we need to know how variants have an impact specifically in blood cell development.

00;10;52;28 - 00;11;04;28

Katie Partington: I think a lot of people will have vague memories of blood consisting of red blood cells and white blood cells, but it's a bit more complicated than that. These cells come from blood stem cells.

00;11;05;01 - 00;11;26;05

Vijay Sankaran: I think what's become apparent as we've been studying blood and immune cell development in general is there's just a tremendous number of cells states that these blood stem cells transit through as they're differentiating and as they go to produce the millions of red blood cells, white blood cells, and platelets that you're producing every second right now, as we're sitting here.

00;11;26;07 - 00;11;52;29

Katie Partington: Blood cells are produced in the bone marrow through a process of cellular differentiation. A stem cell can become one of two types of progenitor cells, one which will eventually become the lymphocytes involved in adaptive immunity, and one that will become red blood cells and cells involved in innate immunity. This developmental history and the relationship between cells as they undergo differentiation is what we call the lineage.

00;11;53;01 - 00;12;00;14

What's really cool about cells is that the differentiation process is sometimes irreversible, something we call cell fate.

00;12;00;16 - 00;12;06;05

Jerome Freudenberg: And so how are all these different cell types and states connected to blood disease?

00;12;06;08 - 00;12;35;05

Vijay Sankaran: To me, what's clear is, as we've been studying individual genes or variants that have an effect on this process, their effects tend to be very cell state specific. So for instance, we've studied diseases that affect the ribosomes. But it turns out those ribosomal protein mutations that cause diseases like Diamond-Blackfan anemia, they actually affect the specification and lineage commitment of blood stem cells and progenitors to become red cells and not to other lineages.

00;12;35;07 - 00;12;53;10

And so there's an exquisite cell type specificity. So if you did not capture that stage where those cells are undergoing lineage commitment, you would miss that whole aspect of biology that's so important. And so that's what really motivated us to think about, well, how do we capture this diversity of cell state.

00;12;53;12 - 00;13;14;20

Jerome Freudenberg: So essentially mutations can affect the functionality of a cell but also how it determines its cell lineage. So when scientists are creating these variant effect maps for a gene, they can't just do it in any cell. They have to do it in the correct cell type. So how many blood diseases are there considering the complex effects of mutations on cells throughout different cell states?

00;13;14;22 - 00;13;32;26

Vijay Sankaran: Yeah, it's a really great question because I think it sort of depends on how you define these things, right? Because I remember when I was thinking about sickle cell disease and beta thalassemia, which we're still thinking about, we often would cite it as the most common monogenic disorder in the world. But how do you define that? Right? And I think that that's hard.

00;13;32;26 - 00;13;56;28

And as you think about looking at the combination of common and rare genetic variation, which I think we're probably going to see a spectrum of variation. So I would sort of argue it in a couple of ways, which is I would say there are dozens to hundreds of rare monogenic blood diseases that I, as a hematologist, will see. Some are a little bit more common than others, and some are fairly rare.

00;13;57;00 - 00;14;27;26

For instance, we thought a lot about these bone marrow failure syndromes I've given one example, Diamond-Blackfan anemia that affect a few individuals in every million live births. But I would argue that to me, this is just the tip of the iceberg. We know that there is an incredible contribution of genetic variation that underlies most of the blood cell phenotypes. Even just common measure, like what our red cell count is, what our white blood cell counts are, so forth.

00;14;28;00 - 00;14;48;23

And we thought a lot about some of that genetic variation and I think it plays an important role in modifying both blood diseases and cancers. And I would say the other end of the spectrum that we've thought about is even just genetic variation that impacts blood production to put people at risk for acquiring things like blood cancers and much of the heritable variation

00;14;48;23 - 00;15;03;29

and much of the heritable risk seems to be in common genetic variation. So I say that only because, to me, it's a hard question to answer, because I would say almost everybody has genetic variation that might affect risk and so forth.

00;15;04;01 - 00;15;15;10

Jerome Freudenberg: Oh okay, so some genetic variants lead to rare disease, but other variants can simply help explain characteristics like how many red blood cells a person has which also affects their health.

00;15;15;13 - 00;15;29;04

Katie Partington: Even within the general population that has not been diagnosed with a genetic disease, genetic variants could help explain some of the variation within that population. In fact, genetic variation is what makes each of us who we are.

00;15;29;06 - 00;15;37;11

Jerome Freudenberg: So given how much genetic variation there is, this can't just be studied one variant at a time. How did you go about trying to tackle this challenge?

00;15;37;14 - 00;16;01;03

Vijay Sankaran: So you know when I started our group, I naively said, we're going to do a couple of things. We're going to be involved in doing some human genetics and discovering rare variants that cause different blood and immune diseases. And we're also going to continue to study blood production. We're going to try to manipulate human cells try to really understand how this different variation can impact this.

00;16;01;06 - 00;16;29;15

And what was really exciting when we started was CRISPR-Cas9 had just been described a couple of years earlier. People were starting to apply it to largely human cell lines. We started to apply it to primarily human blood stem cells and some of the progenitors that differentiate from those cells. And we could then in vitro or even in vivo, manipulate the cells in different ways and start to differentiate them.

00;16;29;16 - 00;16;50;20

But the problem was one of scale. We were keeping up with just the amazing work being done in the field and where people started to do imprecise genome editing, you know, creating random insertions and deletions at a desired site. But you couldn't really control that. There was, of course, all these generation of genome editing machinery that continued to evolve.

00;16;50;22 - 00;17;11;22

And so as we continue to incorporate some of this largely in a low throughput fashion in these blood stem cells that we could get from healthy human individuals and manipulate different ways. The thought was always in our head, could we somehow scale this up? Could we do more than just one gene, one regulatory element at a time?

00;17;11;24 - 00;17;23;04

Katie Partington: We talked about regulatory elements in a previous episode and how variant effect maps can decipher non-coding DNA. Keep this in mind as we'll touch on this a little later in this episode.

00;17;23;07 - 00;17;53;13

Vijay Sankaran: And it was really exciting to see the developments that started with Cas9 screening approaches in cell lines, and then evolved into doing things like base editing and prime editing, largely in cell lines. And those developments were really exciting. But the challenge for us was those could be applied in convenient cell models. But how do you apply those in the systems that we love to study blood stem cells and some of the differentiated progeny of these cells, and how can you actually apply this?

00;17;53;15 - 00;18;12;11

Could we start to think about applying some of these tools to primary blood and immune cells, and really look at the outcome of these, both in vitro and eventually in vivo? We started initially thinking about cutting Cas9 approaches and scaling this to target different genes.

00;18;12;13 - 00;18;18;17

Jerome Freudenberg: But my understanding is that this doesn't really give you the allelic precision that you'd ideally have, right?

00;18;18;19 - 00;18;47;14

Vijay Sankaran: Right. We know that most human genetic variation are likely single nucleotide variants. There's certainly a role for important structural variants and other types of variation. And so it was that motivation to think about how do we apply this. That really led us to realize that base editing could be applied in these primary blood and immune cells. And it took a little bit of development and a little bit of trial and error and some failure.

00;18;47;16 - 00;18;49;18

Jerome Freudenberg: So what ended up working?

00;18;49;20 - 00;19;13;23

Vijay Sankaran: We realized that you could, at least for individual edits, deliver bae editor protein pretty effectively to the cells. And this is actually what's happening in clinical trials or companies like Beam Therapeutics and others that are using this approach to target particular edits and introduce them to cells. But that, of course, introduces the challenge, okay, you can deliver a protein, you can create a single edit, but how are you going to scale this up?

00;19;13;25 - 00;19;19;13

Jerome Freudenberg: Basically, you need a way to introduce and test every possible edit.

00;19;19;16 - 00;19;34;16

Katie Partington: Scaling these assays up to be able to measure multiple variants simultaneously is so important, and a big reason people use MAVEs. One researcher who has been working to address the problem of scale is Doctor John Doench. Can you introduce yourself, John?

00;19;34;19 - 00;19;42;03

John Doench: Hi, I'm John Doench. I'm an institute scientist at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts.

00;19;42;05 - 00;19;46;16

Katie Partington: How did you first get interested in genetics and performing large screens?

00;19;46;19 - 00;20;00;24

John Doench: I was a PhD student in Phil Sharp's lab at MIT, and this is when RNAi had first been discovered. And pretty quickly it became clear that it was going to be a great tool for interrogating gene function.

00;20;00;27 - 00;20;26;19

Jerome Freudenberg: RNAi is an abbreviation for RNA interference. This works by introducing double stranded RNA into the cell, which then utilizes existing cellular machinery to cleave the complementary mRNA molecule and prevent it from being translated. An important difference between RNAi and CRISPR is that RNAi can be used to change mRNA levels, whereas CRISPR screens enable making changes to DNA.

00;20;26;22 - 00;20;28;14

Katie Partington: So how did you use RNAi?

00;20;28;15 - 00;20;55;22

John Doench: At the time, the Human Genome Project was wrapping up. Now we have a list of all the parts in the human genome, but we don't know what most of them do. And we don't know how dysfunction in those genes leads to disease. And so having a method to perturb genes and see how is the cell different than it was when the gene was there, struck me as a very powerful approach to systematically explore biology.

00;20;55;25 - 00;21;05;19

I was leaving my PhD and then got into the field of screening. And at that time, you know, RNAi was the only game in town. So it's what we all went with.

00;21;05;21 - 00;21;10;22

Katie Partington: But RNAi is no longer really the method of choice. What changed?

00;21;10;24 - 00;21;38;06

John Doench: The widespread off target effects of RNAi were becoming pretty apparent. Fast forward to the early 20 tens, and CRISPR technology comes along in the diagram of how I'm disrupting the cell, I'm moving the scissors from cutting the mRNA to cutting the DNA. So the ability to manipulate genetic information at multiple levels now has just really accelerated the specificity of CRISPR.

00;21;38;10 - 00;22;01;10

It is a lot better. You can do more with it, since it's entirely exogenous. You can modify it to be a transcriptional activator or transcriptional repressor, or as I think we're going to talk about largely here today, to edit DNA at the nucleotide level. So like base editing and prime editing approaches. So it's a really, really useful tool that one can apply in a lot of different ways.

00;22;01;12 - 00;22;06;29

Katie Partington: John mentioned you can use CRISPR to make changes at the nucleotide level. How does that work?

00;22;07;01 - 00;22;29;29

Jerome Freudenberg: CRISPR editing involves an enzyme called Cas9 that cuts both strands of DNA. A guide RNA forms a complex with the Cas9 molecule and localizes the complex to a specific location in the genome that matches the guide RNA. And this is where Cas9 makes a cut. In CRISPR knockout screens, this cut is incorrectly repaired by the cell, which leads to that gene being inactivated or knocked out.

00;22;30;02 - 00;22;50;29

For applications like base and prime editing, the Cas9 enzyme is modified so that it only makes a cut in one strand of DNA, or modified such that it doesn't cut it at all. In the case of base editing, a deaminase enzyme is attached to the Cas9 molecule and is able to convert a C to a T or an A to a G with the help of the cell's repair machinery.

00;22;51;01 - 00;23;04;14

In prime editing, the guide not only targets the Cas9 to the correct location in the genome, but also contains a template for repair that is reverse transcribed into DNA. For prime editing, this guide is called a peg RNA.

00;23;04;17 - 00;23;14;02

Katie Partington: So, unlike your traditional CRISPR knockout screen, base editing enables single nucleotide changes, and prime editing enables changes encoded by the peg RNA?

00;23;14;05 - 00;23;15;23

Jerome Freudenberg: Yeah, exactly.

00;23;15;26 - 00;23;19;15

Katie Partington: So how are these tools used in screens?

00;23;19;17 - 00;23;41;15

John Doench: So technologies that really look at the nucleotide level enable you to answer very different questions than one does with a knockout screen. For base editing it allows you to ask questions at the level of the nucleotide or the level of the amino acid. So the ability to tile across a gene of interest rather than screen many, many, many genes at a time.

00;23;41;17 - 00;24;06;20

There are other technologies as well, the most obvious one being prime editing. Now, we first started with base editing because prime editing didn't exist yet, but at least at a broad level, they're both technologies that allow you to instead of thinking about genes as the discrete unit of information, to think more about amino acids, or even in the non-coding space nucleotides as the level of information that you're trying to understand.

00;24;06;23 - 00;24;14;07

Katie Partington: I've heard that base and prime editing are less efficient than CRISPR knockout screens. What causes this difference in efficiency?

00;24;14;09 - 00;24;39;24

John Doench: In terms of cell by cell efficiency, meaning, when I introduce a reagent into cells, what fraction of cells do the thing that I want them to do? Cas9 based knockout is still the most efficient of the three technologies that we're discussing here. It doesn't require any other moving parts. Cas9 evolved to be a nuclease, and we're just taking advantage of that.

00;24;39;28 - 00;25;06;20

It evolved to defend against phages. And now we're just cutting different DNA. So the very high fraction of cells that receive a reagent, will use that reagent in the way that we wanted. I'd say the

next most efficient are base editors. Base editors obviously have more moving parts. One or more domains are fused to a Cas protein, and those enzymes have different efficiencies.

00;25;06;23 - 00;25;29;24

Likewise, when you're trying to change bases in a mammalian genome, we have mechanisms that are looking for that and trying to prevent it from happening, which, you know, in the long run is a very good thing from the standpoint of the health of our genome. But from the standpoint of introducing changes, you have to go through some intermediates that a cell again, is on the lookout for and is trying to reverse.

00;25;29;26 - 00;25;50;25

So you either need to play tricks with the cell or you need to accept lower efficiency. The good thing about any of these technologies is that they're a one way street. Meaning once you introduce your base edit or your knockout or your prime edit, there's no way for the cell to really go back once it's installed in the genome.

00;25;50;25 - 00;25;52;22

It's installed in the genome.

00;25;52;25 - 00;25;58;12

Katie Partington: How does the efficiency of these methods impact what applications they can be used for?

00;25;58;15 - 00;26;24;22

John Doench: I think the real measure for efficiency is can one use this technology for loss of function screens or loss of signal screens? Can something undergo negative selection in a pooled setting? And can you measure that if you want to do positive selection? An example of that would be cells that survive challenge with a small molecule or survive challenge with the virus that'll otherwise kill them.

00;26;24;24 - 00;26;29;20

Jerome Freudenberg: Here, John is talking about multiplexed assays that can produce variant effect maps.

00;26;29;23 - 00;26;38;21

John Doench: In order to see negative selection phenotypes, where your window is a lot smaller. And so the efficiency of the technology is really critical.

00;26;38;23 - 00;27;05;09

Jerome Freudenberg: Let's unpack that for a second. Why would a variant decreasing in frequency require a higher efficiency editing approach? The reason is subtle, but essentially the assay relies on a variant not being seen after the editing as a marker for loss of function. But if the editing has low efficiency, then the scientist has to determine whether the variant is lethal or whether they simply never made it in the first place.

00;27;05;12 - 00;27;32;14

John Doench: I think I've seen enough base editing screens to say like, yeah, if your system is optimized well, you can reasonably expect to do a negative selection screen with base editing. With prime editing the third on the list, I'd say it's fair to say as of today, it is the least efficient of those three technologies. But it's also the one that is iterating the most rapidly right now, being the technique that is the newest relative to base editing

00;27;32;14 - 00;28;08;02

and again, nuclease Cas9. Prime editing has the most moving parts, in addition to needing to recruit enzymes, in this case reverse transcriptase. There's also more complicated aspects to the RNA itself. You're not only using the RNA as a guide, but the RNA has another part to it the pegRNA that encodes specific information and how well that information is transferred is another source where some sequences are good and some sequences aren't so good. And likewise, the same challenge that base editing has of the cell

00;28;08;08 - 00;28;17;27

realizing that it has a mutation and trying to counteract that mutation with endogenous repair pathways. That's something that prime editing also suffers from.

00;28;18;00 - 00;28;23;25

Katie Partington: So for these screens, can you just sequence the guides in order to know which edit a given cell got?

00;28;23;28 - 00;28;40;13

John Doench: So with Crispr nuclease screens, again, because you can assume that the efficiency was quite high, all you need to do is read out the guide RNA. You don't need to actually read out the endogenous locus and figure out if the thing was knocked out.

00;28;40;16 - 00;28;49;07

Jerome Freudenberg: Geneticists use the term knockout to mean gene deletion or complete inactivation, a frequent outcome with CRISPR Cas9.

00;28;49;09 - 00;29;10;24

John Doench: So because of that, it's very easy to screen lots and lots of these things in high throughput and have a lot of confidence in the results. With base editing. The main challenge is that you don't really know the edit. You can make a reasonable prediction of what the edit was going to be. With base editing there are numerous sources

00;29;11;18 - 00;29;35;27

of I don't want to quite call it off target editing, but at least ambiguous editing you're using in a base editor. There might be an A at position two that shows some editing that's outside the canonical window of where editing would occur along the length of the guide RNA. But it can still happen for sure. And if the effect of that edit is really potent, then maybe that's what's driving your phenotype.

00;29;35;27 - 00;30;02;01

So what you can't do is just sequence the guide RNA and assume that you know what happened at the endogenous locus. I think the same is true with prime editing less because of the ambiguity of the potential outcomes per se. But just from an efficiency standpoint, you can't assume that a cell that has a particular pegRNA necessarily introduced the edit that you hoped it did.

00;30;02;03 - 00;30;10;15

Katie Partington: So if you're not sure if an edit was made or which edit was made, how do you interpret the results of a base editing or prime editing screen?

00;30;10;18 - 00;30;33;01

John Doench: So how one deals with that in a screening campaign has varied. With base editing, one approach that we've liked is to do a large scale screen, read out the guide RNA, but know that we're not going to interpret the hits until we individually validate a guide RNA and determine, well, what change did it actually make at the endogenous locus?

00;30;33;03 - 00;30;51;02

For example, let's say you're scanning across a protein of interest and you realize like, oh, I got a bunch of hits that are all localized to this region, I'm going to go study those, and I'm going to individually validate the hits that arose from that region. But, you know, nothing else interesting happen anywhere else. So I don't need to individually validate those.

00;30;51;08 - 00;31;13;04

Same is true of prime editing. Now, if you rely solely on the abundance of pegRNA to determine what happened in the screen, your analyzes are going to have a lot of false positives and false negatives. The other, and I've seen this used with both base and prime editing, is to read out the screen, not by sequencing the guide RNA, but rather to sequence the actual region of interest.

00;31;13;04 - 00;31;39;01

You know, if you're introducing edits. The problem with that is it doesn't scale terribly well. Introducing edits into a thousand places in the genome. Then you're doing a thousand plex PCR, which is possible, but not trivial. So, again, I think it's more a question of how do you tether the realistic expectations of a technology to the ambition of the question that you're asking and the scale of the model system that is feasible?

00;31;39;03 - 00;31;54;00

Katie Partington: Okay. So the current state of the field is to use these screens as a sort of first pass, and then you can validate any important findings from a larger screen for a smaller set of variants. Can you give an example of a screen you've done and something you've learned from it?

00;31;54;03 - 00;32;18;05

John Doench: There was a screen that I was involved in, which was work done in collaboration with Scott Armstrong's group, where they were concurrently with a clinical trial to study small molecule inhibitors of MEN1 where they were seeing what resistance mutations were appearing in patients MEN1 is a chromatin modifier that is often disrupted in a number of leukemias.

00;32;18;08 - 00;32;38;20

Jerome Freudenberg: Chromatin modifiers help cells determine which genes should be turned on and turned off. Like Doctor Sankaran explained earlier, healthy cells need to follow specific gene expression programs and this needs to be well controlled. You wouldn't want an eye cell in the wrong place. Chromatin modifiers help this process, and of course it can frequently go wrong in cancer.

00;32;38;22 - 00;33;21;10

John Doench: And it's thought that MEN1 inhibitors will be particularly potent for types of leukemias that are addicted to a particular signaling cascade or a particular chromatin state that MEN1 is involved in. So this was identified as a target by cancer genomics efforts. So like, okay, well, this is a gene that that we should go after. And then a number of small molecules and other modalities including some peptides that have been developed to try to target Men1 with the idea that these leukemias will be particularly sensitive to MEN1 inhibition the same time as clinical trials were going on that we're testing these MEN1 inhibitors in patients, people who were also then

00;33;21;10 - 00;33;56;01

working on well, can we look at the resistance landscape in cell culture models of this leukemia? They were also doing a parallel based editing screen, and the two were lining up quite nicely. And obviously one would then wants to get to the point where we are understanding resistance mechanisms, not in people, but where we're able to understand these things far earlier in the clinical development pathway, so we can understand what particular small molecules are less

likely to give rise to resistance, or when they do, what does resistance look like? And what could next line therapies look like?

00;33;56;03 - 00;34;08;27

Katie Partington: Wow, that's so cool. It would be great if we could use this approach for even more genes. What areas do you think there are still gaps where advances are needed that are preventing these technologies from being deployed more broadly?

00;34;08;29 - 00;34;35;04

John Doench: We're going to continue to see developments in sort of the three legs of the stool that any particular screen stands on. One of them is the perturbation, you know, how are you manipulating, how are you changing the cell in some way or another? CRISPR screens are only about ten years old, and the efficiency gains and the spread of biological questions and manipulations that one has been able to use has been amazing.

00;34;35;06 - 00;35;01;20

We're going to continue to see development in new types of perturbations, but also just greater efficiency of perturbations, so that you're able to get more information out of a cell reliably. The second is the assay itself. There will continue to be new assays developed, but just as importantly is how do we take those existing assays and how do we make them more scalable, which is oftentimes a question of cost.

00;35;01;27 - 00;35;32;29

How do we make them more reliable? They're easy for a new lab to pick up and run effectively, and the data that are generated from them can be put in a place where it doesn't just exist, but it is functionally usable. I think that's still a huge bottleneck in biology. And then third is the model system. Do we have cells that are accurate models of the various diseases that we want to study, especially for diseases that are not cell autonomous?

00;35;33;01 - 00;35;52;12

You can assess a lot of cancer phenotypes by is the cancer cell still alive. Yes or no. But even then, if you're not thinking about the role that the immune system plays in cancer, you're missing

a pretty big part of the picture. And so understanding the role of the immune system in cancer, that requires, by definition, multiple cell types coming together into a screen.

00;35;52;14 - 00;36;19;09

How do we think about other areas of biology, neurodegenerative diseases? What are the model systems that allow us to interrogate that? Schizophrenia, autism, diabetes? There's all sorts of higher order complexity diseases for which what is the model system and how do we wrestle that to the ground in a way where screening even protein coding genes, much less regulatory elements, is actually meaningful?

00;36;19;09 - 00;36;32;17

So I think continued development of model systems that reflect the actual disease biology we want to study. That's going to have to continue for any of these perturbational technologies to have an impact even more broadly.

00;36;32;19 - 00;36;36;10

Katie Partington: Where do you see the field headed in the next ten years?

John Doench: Well, I don't

00;36;36;10 - 00;37;02;28

know where we are going to be, but where I hope we are is that there is more systematizing of these approaches across the community and it's not just a wild west of one off publications that then don't actually have any impact on clinical treatment, and our ability to interpret genome sequences is not moving nearly as fast as our ability to generate genome sequences.

00;37;02;28 - 00;37;37;09

Now, it has to be that way, right? It couldn't be the reverse. But, you know, I'd like to see those lines move much more in parallel rather than one of them be a hockey stick and the other increasing linearly. So I think where we're going to have the most impact is scientists working together, working together with journals, working together with funding organizations, working together with software engineers and data scientists to really take the stuff that each of us does individually in the lab and move it towards a common purpose.

00;37;37;11 - 00;37;49;17

I'm optimistic that over the next ten years or so, we're going to get to a place where the work that we do in the lab is reaching its way into clinical practice a lot faster than it is now.

00;37;49;20 - 00;37;53;25

Katie Partington: Thanks a lot, Doctor Doench. This was very inspiring.

00;37;53;25 - 00;38;10;06

Jerome Freudenberg: Actually. I think Doctor Sankaran had a story to tell about how these techniques can be applied for uncovering real mechanisms of disease. Now that we've heard about how based and prime editing can be used in screens, let's see it in practice. Vijay, can you describe some of the findings from a screen you've performed?

00;38;10;09 - 00;38;37;13

Vijay Sankaran: Maybe I'll just give you a couple of examples where it's actually been really useful for us. So I mentioned to you earlier, you know, when we started our group, we were really interested in trying to understand both common genetic variation, but also some of the rare genetic variation. And I think one of the most powerful lessons that happened is one of the first mutations I discovered, which, incidentally, was actually something I had initially found when I was an intern in pediatrics.

00;38;37;20 - 00;38;53;16

I found a mutation in a patient who had a disorder that was attributed largely to these mutations in ribosomal proteins, but it turned out this patient had a mutation and a transcription factor, GATA1, and that caused the disease Diamond-Blackfan anemia in that patient.

00;38;53;19 - 00;39;05;13

Katie Partington: A transcription factor influences transcription levels of nearby genes by binding to specific DNA sequences and recruiting other proteins that increase or decrease the level of transcription.

00;39;05;15 - 00;39;30;27

Vijay Sankaran: Since that time, lots of groups have reported GATA1 mutations, causing a range of different phenotypes. Everything from patients having low platelet counts to patients having transfusion dependent anemias to a whole spectrum of different blood diseases. And so the hard part right now is that if you go in, if you have a blood disorder, you'll often get GATA1 sequenced.

00;39;30;27 - 00;39;53;01

But there are many, many variants of uncertain significance that, as with many other genes that are tested for that come out of such screens. And so it happened that we actually did a base editing screen in GATA1. The benefit is that we could actually get all the cell states that are relevant. So we could see the blood stem cells differentiating to produce red blood cells.

00;39;53;01 - 00;40;16;21

We could see them producing the platelet precursors called megakaryocytes that spit out platelets from the bone marrow. We could produce all these and other types of white blood cells. And importantly, we could couple the base editing screen we did with single cell RNA sequencing readouts. The real benefit there is that you don't have to have a marker to pick out your red blood cells or your platelets.

00;40;16;23 - 00;40;40;05

You can just be agnostic to it and say, let me look at where different cells are and what cell states those are occupying, and how each variant that you've recreated will allow you to impact one lineage or another or all lineages. And so using that kind of data, we were able to generate a really nice screen of different mutations in GATA1.

00;40;40;07 - 00;41;04;02

So as we were doing the screen, I get a call from one of my colleagues and he says, Vijay, you know, I have a patient with a variant of uncertain significance in GATA1. And I said, oh, really? And he knew we had done work on GATA1. But, you know, the typical thing would be, all right, let me see if I can find somebody to do an experiment on this particular variant that you have identified.

00;41;04;05 - 00;41;23;07

The wonderful part is we'd already done the screen and we could just look up in the data we had. What is the impact of that variant? And so that was really satisfying because within a few minutes we could actually prosecute that variant. And of course we had to do lots of work to follow up and really say that indeed, that was a pathogenic variant.

00;41;23;07 - 00;41;37;19

But I think that that's just an important lesson. I think as we get more and more of this kind of data, we'll be able to turn what were problems that would take months, maybe longer for somebody to really follow up on and and actually prosecute.

00;41;37;22 - 00;41;51;27

Jerome Freudenberg: Well, that's a wonderful story. It really shows the direct impact that some of this work can have. So what about other types of screens? Can you tell me about the one that you performed on a sequence that didn't encode a protein, but was for a promoter?

00;41;52;00 - 00;42;18;16

Vijay Sankaran: One of the problems we've been interested in is how these fetal hemoglobin genes are regulated. I told you, that's where I started my career. And we continue to think about that problem. And it turns out that there's been lots of single nucleotide variants in the promoters of these genes that have been characterized. But the impact of these different variants are on fetal hemoglobin levels has not been well characterized because typically people were describing single individuals or single patients

00;42;18;18 - 00;42;43;10

with these variants. We were able to take these blood stem and progenitor cells, perform a screening for all variants in the promoter and then look at their effects, and that was valuable because then we could really Side-By-Side compare what is the effect of one region versus another. And I think that we're just starting to think about how we can really use this kind of data.

00;42;43;10 - 00;43;06;27

But I think it's already valuable, right? Because some of these nucleotides are being targeted with ongoing therapies you'd love to know which might be a better target than another. And if you've only gotten that data from some patient that has a certain level of fetal hemoglobin that could be dependent upon their genetic background or a lot of other factors, whereas you doing isogenic screen really allows you to cross compare in the same genetic background.

00;43;06;27 - 00;43;17;12

What is the impact of a variant on fetal hemoglobin levels as compared to another variant? And so I think it also just opens up a lot of opportunities that hopefully will be therapeutically valuable.

00;43;17;14 - 00;43;37;01

Jerome Freudenberg: So this ties us back to another episode that we did on MPRAs. A lot of people seem to focus on making variant effect maps for coding variants, but non-coding variants are a particularly interesting group of variants that we haven't been able to decipher as well as coding ones. Vijay, can you tell us more about how non-coding variants are implicated with disease?

00;43;37;03 - 00;44;03;18

Vijay Sankaran: I think that what we found is for many of the rare diseases we've worked on, we see a tremendous amount of coding variation causing these diseases. And of course, I do think that we'll probably increasingly recognize non-coding variation. But as we just look at how much of a rare disease can we explain with some of the sequencing we've done, I think much of it falls into the coding space.

00;44;03;20 - 00;44;37;15

On the other hand, for some of the more complex phenotypes, such as blood counts or other phenotypes that we've looked at, or even looking at just common variation that impacts the risk of disease, much of that seems to be non-coding. So I think it really emphasizes how even if we're thinking about rare diseases, we have to think about some of the common genetic variation in the genetic background that individuals have that influences the rare, that sort of more highly penetrant phenotypes that might be due to, you know, coding variations.

00;44;37;17 - 00;45;05;28

I think that non-coding variation is going to have a huge role in many of these diseases, whether as a modifier or as a primary driver of the phenotype that an individual has. In studies that we've done in hematopoiesis, we know that much of this seems to map to regulatory regions that are important for blood cell production, where many master transcription factors seem to bind.

00;45;06;00 - 00;45;16;00

But I think you bring up a great point, which is how do you tackle this? How do you go after this? And I'm not sure that we have a great answer at the moment, but I do feel optimistic.

00;45;16;02 - 00;45;21;03

Jerome Freudenberg: Thank you, Doctor Sankar. And hopefully the non-coding problem will be solved in the near future.

00;45;21;05 - 00;45;32;16

Katie Partington: Most of the genome is non-coding and that seems really daunting to me. In the search for disease causing variation, Doctor Doench explains the challenges associated with studying the non-coding genome.

00;45;32;19 - 00;45;59;29

John Doench: We have ways of manipulating proteins not only at the DNA level, but at the RNA level, at the protein level themselves. And we're still cataloging and understanding all of that. So the non-coding genome is in orders of magnitude larger problem, not just because there's orders of magnitude more non-coding genome then protein coding, but also the number of tools that we have to manipulate it are by definition smaller.

00;45;59;29 - 00;46;31;26

And I think that CRISPR technologies are, on some level, the only game in town from a perturbation standpoint. Obviously, the ENCODE project has made tremendous strides in cataloging, but I'm talking about systems that allow us to actually manipulate the DNA. But if you just want to generically study non-coding DNA, what is the readout that you use? Well, it's

hard to come up with something. If you were to do RNA seq on, you know, all 3 billion single base changes that one could make,

00;46;31;26 - 00;46;53;20

that would be a rather expensive endeavor and probably not the way to go figure this out. So to me, the question is really how do you nominate hypotheses for what non-coding DNA do you want to study in what cell type, why? And then how will you read out that that non-coding DNA, the change you made to that non-coding DNA was impactful?

00;46;53;23 - 00;47;30;11

Katie Partington: Thank you, Doctor Doench, for this wonderful interview on modern genetic editing technologies.

Alex Nguyen Ba: Let's thank Doctor Sankaran and Doctor Doench for joining us today. And so it turns out that non-coding variation, despite its importance, is just harder to study. Not only is the code less clear, the tools are not as robust, and the readout is fundamentally more complicated. That being said, I think this episode gave a glimpse of what is being done to address this and how we've used novel technologies to interrogate coding and non-coding variation in important disorders.

00;47;30;14 - 00;47;53;25

And so in this episode, we've explored how genetic tools are empowering scientists in the fight against blood genetic disorders. But these tools are especially powerful when you couple them with knowledge about how blood develops and what the mechanism of disease actually is. This means that variant effect mapping is only one of the many steps required for a future of precision medicine.

00;47;53;28 - 00;47;58;24

But until then, now is the time for the rapid fire questions.

00;47;58;26 - 00;48;01;04

Moez Dawood: Are you ready?

Vijay Sankaran: Absolutely.

John Doench: Okay.

00;48;01;06 - 00;48;14;05

Moez Dawood: DNA or RNA?

Vijay Sankaran: DNA

John Doench: RNA

Moez Dawood: R or Python

Vijay Sankaran: R

John Doench: Python

Moez Dawood: Short read or long read?

Vijay Sankaran: Short read.

John Doench: Short read.

Moez Dawood: Single cell or bulk?

Vijay Sankaran: Single cell.

John Doench: Bulk.

00;48;14;07 - 00;48;32;10

Moez Dawood: Knock in or knock out?

Vijay Sankaran: Knockout.

John Doench: Knockout.

Moez Dawood: Golden Gate or Gibson?

Vijay Sankaran: Gibson.

John Doench: Golden gate for the win

Moez Dawood: Coding or non-coding?

Vijay Sankaran: Non-coding. But I love it all.

John Doench: Coding

Moez Dawood: In vitro or in vivo?

Vijay Sankaran: In vivo.

John Doench: In vitro.

Moez Dawood: Favorite model organism?

00;48;32;14 - 00;48;46;08

Vijay Sankaran: Humans

John Doench: A375 cells

Moez Dawood: Favorite scientist?

Vijay Sankaran: Sean Carroll.

John Doench: Phil Sharp.

Moez Dawood: Favorite media?

John Doench: DMEM.

Vijay Sankaran: We have some nice serum free media for culturing blood stem cells.

Moez Dawood: Favorite sequencing technology?

00;48;46;10 - 00;49;07;11

Vijay Sankaran: Oxford nanopore technologies.

John Doench: Illumina.

Moez Dawood: Favorite piece of lab equipment?

Vijay Sankaran: A Pipetman.

John Doench: Vortexer.

Moez Dawood: Favorite science movie/media?

Vijay Sankaran: The night science podcast.

John Doench: I can't believe I'm going to say this, but, Armageddon.

Moez Dawood: Something outside the lab that you wish you could multiplex.

Vijay Sankaran: Reading books

John Doench: Folding laundry.

00;49;07;13 - 00;49;43;28

Moez Dawood: What is the smartest thing you've ever done in the lab?

Vijay Sankaran: I don't know that I've ever done anything really smart, but I think the thing I have been really fortunate to experience in the lab is working with other people.

John Doench: When I was doing a lot of cell culture, I glued a magnet to the vacuum aspirator so I wouldn't need to change my glass aspirator every time. It would just like magnetically stick to the side and be ready to reuse.

Moez Dawood: Then the last question here what is the stupidest thing that you've ever done in the lab?

Vijay Sankaran: I have set fire to my bench in a lab.

John Doench: I poured an agarose gel but didn't put the comb in

00;49;43;28 - 00;49;57;22

Alex Nguyen Ba: To all you listeners. I hope you enjoyed today's episode. Don't forget to tune in for the next episodes of Variants and Us, where we'll discuss cool topics such as genetic diagnoses, neurological disorders, and human variance in the context of human diversity.

[outro music plays]