Of Known Significance

An exclusive Clinical OMICs conversation with the director of Genomes2People, Robert Green, M.D., MPH

By Julianna LeMieux, Ph.D. (https://www.clinicalomics.com/author/juliannalemieux/) - March 25, 2020

Robert Green, M.D., MPH, is director of the Boston-based Genomes2People Research Program and new Preventive Genomics Clinic at Brigham and Women's Hospital, the Broad Institute, and Harvard Medical School. Through his work there on such research initiatives as the REVEAL Study, the MedSeq, and BabySeq Projects, Green has become one of the leading authorities, if not the leading authority, on the health benefits and economic effects of genomic medicine. He recently spoke with *Clinical OMICs* Senior Editor Julianna LeMieux about this work and current landscape for genomic-driven medicine.

Julianna LeMieux, senior editor, Clinical OMICs: What has been the focus of your work in genomics?

Robert Green, M.D., MPH, director, Genomes2 People: I am a clinical geneticist, with additional training in public health and implementation science, which means that I am constantly thinking about how genomics will be beneficial to the greatest number of people. This requires attention to the individual patient or family that is in front of me in the examining room, but also a focus on how new technologies and new clinical practices can improve health in the population at large.

This focus on population genomics is pretty new. There has been plenty of attention and research around the ways in which genomics can assist with diagnosis, prognosis, and treatment of heritable conditions and with personalizing the treatment of cancer, but far less scientific attention to the potential for genomics to provide risk information that could prevent disease and maintain wellness. Our most recent scientific studies in the Genomes2People Research Program have been examining the amazing wealth of information that is available, even in a healthy individual, if a human genome is comprehensively sequenced and interpreted. And the data we have generated over the past decade have encouraged us to recently create the Brigham Preventive Genomics Clinic, the world's first clinical service to offer comprehensive DNA sequencing and interpretation to healthy adults and their children.

So, what is the current state of technologies for sequencing and interpreting the human genome?

Green: There is the technology for sequencing a genome, and then there is the technology and process—best practice, if you will—for interpreting that genome. Both have dramatically evolved in the past decade and are continuing to improve. But they are really two different workflows, and because the technology for sequencing has drawn so much excited attention over the past nearly 20 years, since the first draft of the Human Genome Project in 2001, in our conversations with patients, with media, and even among ourselves, we tend to give shorter shrift to the interpretation of the genome.

To put that in perspective, we have frequently talked about how genomes can now be sequenced for under $1,000 and by some companies more recently for much less than this. But these rates for sequencing typically offer little or no clinical interpretation. For most vendors of very low-cost sequencing, specific genes and variants are not even examined. In contrast our MedSeq and BabySeq Projects have not only offered sequencing, but also in-depth, careful analysis of thousands of disease-associated genes.

So, when you are talking about clinical sequencing, it helps to talk about all of the different elements of the interpretive step and how many genes you are planning to analyze. Are you going to select 10 or 50 or 147 or 5000 genes to interpret? And does your interpretive process involve a trained molecular geneticist? Does it involve a manual curation step where a variant curation scientist takes difficult-to-interpret variants and goes into the literature and explores what the papers have written about this gene and this variance?

Remember, there are all sorts of weird boobytraps in analyzing variants for pathogenicity in the genome... it's not as simple as just finding where you have a predicted loss of function. For example, there are some conditions, like some of the inherited cardiomyopathies, where a loss of function mutation does not cause the disease, but a particular missense mutation, with a single change in one amino acid, does.

What are some of the new clinical developments in genomic testing that we are going to be hearing about this year?.

Green: When we talk about genomic testing, we are generally talking about two huge buckets of testing. One we call indication-based testing. This is when there is something wrong with you or your family and we are trying to figure out what it is, or

when there is a known mutation running in your family and we are trying to determine if you inherited it.

The other bucket, which we are starting to hear more about, is pre-dispositional or preventive genomics, where you do not have something obviously wrong with you, but we want to find out if you are carrying a DNA change that puts you at risk for a specific condition—either a rare condition that you are at risk for by virtue of a DNA change in a single gene, or a more common condition where you have a constellation of many genetic markers that puts you at higher risk.

My work has been demonstrating that a surprisingly high percentage of people are walking around with a pathogenic or likely pathogenic mutation in a disease-associated gene for a rare condition. This discovery is going to fundamentally change what it means to talk about "rare" genetic diseases because while they remain individually rare, in the aggregate they are not at all rare!

At the same time, polygenic risks that were first promoted by consumer-facing companies as early as 2007 are coming into maturity, in which thousands or even millions of markers contribute to defining a small group of people for each common complex disease—heart disease, atrial fibrillation, type 2 diabetes (as example)—where a small percentage of people are at a significantly higher risk.

These are some ways in which genetics in the province of rare disease is becoming more and more salient to everyday medicine.

You mentioned that your work showed that many, many people have a pathogenic or likely pathogenic variant. Can you expand on that?

Green: When you think about the DNA changes underlying rare genetic diseases, we have tended to think about 1 or 2% of the population at best, who are carrying these. But it turns out that while these diseases are individually rare, they are far more common in the aggregate than people realize. Perhaps this is because our medical system is fragmented. All the single-gene changes that put you at risk for cancer get referred to the cancer doctors and seen at the cancer centers; all the ones that put you at risk for cardiomyopathy go to the cardiomyopathy doctors; and all the ones that put you at risk for familial hypercholesterolemia get ignored altogether because internists are convinced they can manage cholesterol levels without bothering with genetics.

So, there is a tremendous amount of genetic risk for supposedly rare disease that is out there. As part of our MedSeq Project, we asked the question that if money were no object and we sequenced 100 healthy people and simply asked—if you look at every well-established disease-associated gene and you applied the very best practices in variant curation—how many of them would be found to be carrying well-characterized pathogenic, or likely pathogenic, mutations for a dominantly heritable condition?

In MedSeq, we looked at over 5,000 genes for each person, which was tremendously difficult, time-consuming and labor-intensive. We looked at them extremely carefully and found that 20% of these 100 people were carrying such a mutation. Twenty percent!

Then, in our BabySeq Project, we looked at about 150 such newborn babies, perfectly healthy, cherubic, baby boys and girls. We looked at a smaller gene set—closer to 2,000 genes—because we wanted to restrict ourselves in this case to conditions that

manifested in childhood and adolescence. We still found 11% of these healthy babies were carrying pathogenic variations in these genes.

This discovery has shaken up the field. We have discovered that a predisposition to rare diseases is, in fact, not rare at all. In addition to this we have found that such a comprehensive analysis of the genomes reveals that over 90% of humans are carrying pathogenic changes in genes for a recessive condition which would have implications for their reproductive health. And on top of all this, there are polygenic risk scores. If you just take the top six conditions that have polygenic risk scores, our colleagues have calculated that 50% of men and women, approximately, are going to have an odds ratio of greater than 2.5 for one of those conditions.

It's important to remember that all of these markers are risk factors, not diagnostic markers, but with this kind of frequency, all of a sudden, genetics is everywhere. And I think we have to ask ourselves why genetics is not a more important part of our day-to-day medical care.

What are some of the challenges to incorporating genetics in everyday medical care that you are addressing at Genomes2People?

Green: One of the barriers we are seeking to overcome in our implementation science research is this very widely shared belief that when people seek genetic information, they will be devastated by finding out their results. Our experimental data have debunked this concern over and over again in very carefully controlled studies. Another obstacle is the idea that primary care docs cannot manage genomic information, and we have evidence to rebut this as well in our MedSeq Project. For example, we were the first ones to do a single-page summary of an entire whole genome sequence in which we demonstrated that you could give practicing primary care doctors this information and they would manage it pretty well.

It turns out there are a lot of "myths" about why genomics cannot, or should not, be part of day-to-day medicine, and we are systematically developing high quality evidence to disprove these myths.

Will one bottleneck be relieved by designing ways to present this information simpler, so almost anyone can understand it?

Green: Yes, this is one of the things that consumer-facing genetic companies have done much better than the medical system. But in order to keep the cost down, they have typically done a much more superficial number of gene markers and, in some cases, a much more superficial job of analyzing those.

In order to improve the process, what changes would you make in the information that is given to people?

Green: I would increase the sophistication of the interpretation.

So, deeper or more impactful interpretation?

Green: I would say a more sophisticated interpretation, because what we are really asking is when your computer spits out a marker, it says, okay, you have a variant here in this disease predisposition gene, let us say BRCA2 or MYHP3. You have a change in your DNA. The next question is, is this meaningful or not? Is this a problem or not?

Because there are millions of DNA changes that we have decided do not matter, and there are very elaborate systems for categorizing these. There are five categories: pathogenic, likely pathogenic, variant of uncertain significance, likely benign, and benign.

There are algorithms for doing this and, of course, the first thing any new biotech expert might say is that if there is an algorithm for curating these variants, let us just put it in the computer and have the computer do it. Up to a point you can do that. But the state of current best practice in clinical molecular genetics interpretation does still require some thoughtful time, what you might consider to be manual curation, before the final report is generated.

Now, you see that this manual curation step is in tension with how many genes you want to evaluate. Because if you only evaluate 10 genes, you can do those algorithms and do that tiny bit of manual curation on 10 genes, and you can offer that to a lot of people. But if you want to really evaluate each genome for 5,000 disease-associated genes that you have to put through your algorithm and find markers in almost all of them that will require manual curation, you have bogged the system down in the manual curation step.

This tension will gradually be resolved because we are doing a better and better job of categorizing all the variants that are being found, and agreeing on how they should be categorized, and updating that agreement in powerful ways. This will lend itself to faster and faster algorithms as time goes on, and I'm thrilled to say that we have a new NIH grant to take genomes that have already been sequenced in the Framingham and Jackson Heart Studies and try to build faster, better interpretation pipelines that will reduce the manual curation component. At the same time, this will be the first research grant to return genomic results to research participants in a population-based epidemiological study, and the very first to do so in a population of African Americans in Jackson, Mississippi.