Transcription: FDA Podcast - Whole Genome Sequencing

Welcome to this podcast on FDA's use of an advanced technology called whole genome sequencing. I'm Sue Kelly, a health communications specialist at the FDA and I'm speaking today with Dr. Eric Brown, the director of FDA's Division of Microbiology and Dr. Marc Allard, senior biomedical research services officer.

Brown - "Ah well Sue, thank you for the opportunity, we are delighted to meet with you this morning."

Kelly - **"Thank you for joining us today and my first question is: How is whole genome sequencing used by the FDA today?"**

Brown - "Whole genome sequencing was first developed here and adapted for foods in about 2008/2009 and it has grown to a point from where it is no longer in the developmental phase. Now it's being used regularly since about 2012/2013. Since then, we have been using it regularly to help safeguard the food supply. The way we use it is to augment investigations from our outbreak responders, our first-line outbreak responders, to support them with better information about traceback of contamination in the food supply and we also use it to help our Office of Compliance, to help companies who may have problems find the source of those problems very quickly, to help keep food safer rather than have it go out into the food supply in a contaminated form."

Kelly - **"Just for people who aren't as familiar with this technology, this is essentially a genetic fingerprint of bacteria. Is that the right way to describe it?"**

Allard - "Yes, it's the whole genome of bacteria, so if you think about a common human genome project with the sequence all of the genetic data that are present in a human cell and this is the equivalent in bacteria. Bacteria are much smaller genomes but they're still in the 3-5 million bases range and so when you sequence that you have all of the genetic makeup and that may also include mobile elements, plasmids[i] and phages[ii] that are extra chromosomal pieces of DNA and with this information, instead of using a small amount of genetic information to subtype, you have a huge amount. This provides the ability for high resolution and clear distinction in characterization of the bacteria."

Brown -"The certainty is unprecedented and, as Marc pointed out, because we have so many data points now to compare we can get back down to the source despite where it may have come from, including any place around the globe and including many food imports, which for us has always been a great challenge to try to source track contamination of imports because there are so many potential sources around the world where the same food can come from. So to have this tool in our hands now and to localize contamination to parts of the world very quickly, this really goes a long way to augment the very few inspectors that we have out at any given time. So we are very excited about the laboratory being able to support that role."

Kelly - **"Now where is this technology in its development? Will it be advancing from here?"**

Allard - "So I'd just like to say I've been sequencing since the mid-80s and every 3-5 years we've had a new sequencer for all this time and there is no indication that things are slowing down. There will continue to be technological advances and one of the roles of the genetic laboratory is to test the new genome sequencers and see if they are really faster, better, and cheaper or maybe more mobile. We are looking at some very small sequencers that could fit in a pocket. We could have a lab in a briefcase that could go out to the consumer safety officer and actually do field testing. This is the future vision and it will take time to test that equipment. But at the same time you have to decide when to use it, if you are going to use it to get your feet wet, and actually start using it. And so in 2013, really late 2012, we started implementing desktop sequencers in state and federal laboratories so that they could test the isolates, the foodborne pathogens that they collected in either surveillance or during inspection. So we have a working network of roughly 30 laboratories, including state and federal as well as many international partners that are also collaborators."

Kelly – "That's actually a good segue to a question I was going to ask you a little later but what has been the FDA's role in spreading this technology in addition to working with, my understanding is that you have been spreading it, sharing it with state labs across the country but now you're also sort of looking at the global need. Could you say a little about that?"

Brown - "Sure, I'll let Marc explain exactly how we interface with states and international partners. But what we really started then is you're talking about the onset of the GenomeTrakr, which is the network we spearheaded in late 2012. First nationally and now internationally, we partnered with our federal partners in health, including the National Institutes of Health, National Library of Medicine, where the data is housed and curated, that's called NCBI (National Center for Biotechnology Information) and that's run by Dr. David Lipman at NIH; they've been kind enough to help curate all the data. We've worked with the CDC, who's providing clinical data from PulseNet into the database constantly, and our partners at the USDA FSIS who are working on, of course, those other commodities that they regulate as well as that they're uploading from. So starting in late 2012 the concept was born that we could do a distributed network of whole genome sequencers around the country and really start to make a difference in our ability to pinpoint sources very quickly, almost instantly, as soon as an outbreak or contamination event would happen. So in doing that we are asking everybody to upload their genomes to this source, to the library over at NIH. It is fully public accessible, it is easily accessible, and the data is there for us, for regulators, for public health experts, for the public, and for industry to see the same data that we see in real time. This open source availability has been one of the great innovations of the network, allowing everybody to have equal access at any given time to do analysis and to add further strains to the analysis and so the GenomeTrakr has been growing ever since. Now I'm going to let Marc tell you about its current size and its expansion efforts from there."

Allard – "I just wanted to, so you have a broader perspective on how we develop and apply new technology. In general there's three phases; the first phase is testing equipment in the laboratory and we're always doing this. We're also testing the new sequencers that become publicly available for purchase and if it looks like it has a good application for purchase, then we would do a small round robin study with half a dozen volunteers from state and federal labs and then after that phase, if it is working well, then they either adopt it or they don't adopt it. So we are already through those three phases on

the first desktop instruments and so this is no longer research and development, this is regulatory decisions are being made every day and that the genetic data is helping guide decisions and helping focus the investigations. So where we are is essentially moving that process mostly out of the R&D part and into the regulatory aspects of the FDA and so our partners in ORA (Office of Regulatory Affairs) are essentially doing more and more on their own and taking over that developing network. The other piece is that it started in the research laboratories, but there are many other FDA programs that could benefit from having genomes of their pathogens and we even have some human work done on the White Oak campus. And so some of this is just getting more genomics being added to other applications. So, for example, we have the Center for Veterinary Medicine that includes antimicrobial resistance and monitoring through the NARMS (National Antimicrobial Resistance Monitoring Systems) program, that includes the Vet-LIRN (Veterinary Laboratory Investigation and Response Network), the FERN (Food Emergency Response Network), and the CAP (Cooperative Agreement Program) labs. There are many programs that could be expanded into doing genomics. The international piece is we have provided funding for one World Health Organization laboratory to set up a pilot program that was in Buenos Aries in the Malbran Institute and we continue to work with them to spread to more international laboratories to act a s a center, a regional center of excellence. But the goal is that we would like to see genomics laboratories throughout the world in many regions, so we are working carefully with the World Health Organization and the Food and Agriculture Association to help provide guidance for when countries should start getting involved in genomics and what the best application is, and to which parts of our system that are easily adopted to bring more genomics to the global community.

Brown -"I think we are happy to say too that the GenomeTrakr model with this open source data and shared uploading to single databases that are instantly shared, curated, and cross-referenced in real time. So these 3-4 databases around the world will be able to take all the data and you will be able to see it all at any time at any one of them. This model is really being adopted and accepted by the World Health Organization and other global leaders in public health so we are really excited that we can support WHO, FAO (Food and Agriculture Organization of the United Nations), and OIE in their efforts to really learn how to best deploy whole genome sequencing to developing and more developed countries globally and not just here in the U.S. or western Europe, because we get our food from all over the world and we need to have good surveillance and good traceability from all points in the world where food comes from."

Allard - "Now, I'd also like to point out our partners at the National Institute of Health and National Center for Biotechnology and Information, which holds all the published genetic information, this is the NCBI that Eric referred to. They've been building tools where they do a first rapid interpretation and build a phylogeny[iii] that says how the isolates are related and what is the closest related sample that anyone else has uploaded into the database to look for a match. We built this system with major contributions; NCBI is a full partner in helping analyze and setting up all of these databases. Now, it isn't just foodborne pathogens that are being collected and compared on a daily basis. They've expanded, NIH has a much broader mandate and so they are comparing any human pathogen. So the basic tools that were set up to use for foodborne safety are also now being used for global surveillance of human infectious diseases."

Brown – "And we'll begin to further encompass community-acquired and hospital-acquired infections as well in the clinical setting so we are very excited about that."

Allard - "That piece may expand a lot faster over the next coming year and that's not a piece that at least the Center for Food Safety and Applied Nutrition, that we focus on the food supply and the foodborne pathogens, but this other piece that's heavily supported by NIH and NCBI that's going to continue to be expanded as well."

Kelly - **"Has that started yet, Dr. Allard?"**

Allard – "Yes, it has. Really it was the last year or two years that it's been going but they're adding a lot more. It seems that there is a new species that is being added every month. It's an old infectious disease, but a new genomic database to directly compare those samples."

Kelly – "That's a good segue to my next question, what is ahead in 2017 on the whole genome sequencing front?"

Brown - "Well I think in our case there is multiple avenues that we want to continue to pursue. Perhaps most importantly we want to continue to serve our own agency and to serve the foods program as best as we can. And to do that we want to bring whole genome sequencing around more to the front-line users like the people in compliance, our people in the outbreak response units, so that they can have these tools at their desktop, at their disposal, so they can go in and make these queries and look at these relationships of outbreak strains or non-outbreak strains quickly and as best as they can. So we will continue to try to adapt tools for their use and continue to expand the database with food pathogens so that we can continue to support our role here at FDA. In addition to that, we have a big year planned with our international partners at the World Health Organization. We're having a meeting, actually staring tomorrow, in Washington to try to better prepare the landscape for the deployment of the technology to the developing world and then in July we're supposed to go to Geneva, to the WHO headquarters, and have a two-day outreach session with more than 60 countries that will be onsite to learn more about the technology. So we are very excited about that too in terms of our international outreach. Finally, I know from my perspective, it's been good to see how whole genome sequencing can serve as one microbiological workflow tool. And that is a tool that really displaces numerous other separate tests that each cost a lot of money for a long time. If we wanted to know about the virulence of a strain of a pathogen, if we wanted to know its serotype, if we wanted to know its drug-resistances, all of those were separate tests and very expensive but genomics is providing one avenue, one workflow, so that with that sequence we can tell you: Does it have the genes for drug resistance to various antibiotics, what virulence factors does it have inside of it, what is its serotype, how do we classify it. All of this can be done now very quickly using computational tools rather than wet laboratory tools. We are going to continue to drive towards one micro workflow using this amazing tool that we have at our disposal. "

Kelly – "This is just sort of a side question. When we share the technology, are we actually providing sequencers?"

Allard - " No no, we started in the early days when we were doing the pilot study and we are still providing some support for state and federal laboratories. When we began, the best way to get a new application, a new technology into a state laboratory, is we wrote a research and collaborative agreement between the FDA and the state and we provided a sequencer and a staff and an intern and all the reagents they needed for the year and then we also did additional support. For example we did additional training to the interns, documentation, and standard operating procedures for how to handle the data, upload the data, move the data, interpret the data and those continue. Those initial starts developed into harmonization between the Centers for Disease Control and the USDA FSIS so that laboratories that were collecting from different areas were all in harmony in the quality of the data and uploading of the data.

Kelly – "And you just answered my next question because I know other agencies do whole genome sequencing. I was just going to ask you how your coordinate your efforts?"

Brown – "So that is a very good question, and now that sequencing has permeated all the major public health agencies there is a good bit of coordination. And so we actually have formal steering committees between our agencies that involve management and some of the upper leadership to help develop policy decisions about who will sequence what, how we will share that data, and how we will transfer that information, not just to the database but also to the public. And so we are working hard, we have something called Gen-FS, which is our steering committee for next generation sequencing with CDC, FSIS, and NIH. So we're doing that and the other thing to remember too is that we are making sure we stay in our lanes in what we sequence and that's very powerful because when you have a database that's populated with huge amounts of clinical data from sick people from the Centers for Disease Control and combine that with many thousands of strains from environmental and food data from farms, factories, facilities, and contaminated foods from FDA and FSIS, now you've got a recipe for real success and you can start doing comparisons that are rich and robust and will likely give you information fairly quickly on the sources of some of these events. So it has been a really great experience working with our federal partners and expanding the tool across clinical, food, environmental, and facility health. So it is truly, as Marc said, driving us towards a one health approach for what we are trying to do."

Kelly – "So having the information about human illness and foodborne bacteria in the same database would allow you to tie a particular bacteria not only to its source, but also to the person it has made sick?"

Allard – "The initial matching is a key piece and having the whole genome means that if they're genetically identical you know they're connected in some way. So just imagine a source of an ingredient could have been mailed from one contaminated place to two separate industries. One industry, their preventative controls caught the problem either throughout the food or decontaminated it and it's safe, but another facility did not catch it, their preventative controls failed, their hygienic process failed, and it actually had this one pathogen become a resident problem and then all of the food, or a great majority of the food, that is coming out of that facility is contaminated. So it doesn't always give the source. You have to do a trace forward and traceback or full investigation to really understand the root cause of where the contamination entered the food supply. Having that genetic match means it's connected in

some way and that allows investigators and the FDA to act quickly rather than if there is some ambiguity. Maybe they have a pathogen but maybe it's not out in the public. Maybe no one has gotten sick yet. Maybe it's not really a problem in that it is a single point source but because you have a database you're collecting this and you can say, oh this matches what we saw from the same facility a year ago, which is worrying because it suggests that they have an ongoing resident contamination in the facility that is not under control."

Kelly – "Could it be used to match the bacteria to a person?"

Brown – "I think the CDC uses it for that purpose, to really help support their epidemiological surveillance program. If you remember the CDC has PulseNet and I think the GenomeTrakr and PulseNet work well together to provide all the strains necessary to do a proper surveillance."

Allard – "And the fact you can match it to a person hasn't been lost on the lawyers who want to litigate and reimburse people who have gotten sick from a particular food. We see they are aware of the research and databases we have been building."

Kelly – "And PulseNet is a similar tool?"

Brown – "Yeah, PulseNet actually started in the mid-90s using just a little older tool called PulseField Gel Electrophoresis and that was a fingerprint tool that was used by cutting the DNA into smaller or larger pieces. Now as we are moving from PFGE to whole genome sequencing, it was literally changing our ability to look into an outbreak, much like going from using a light telescope in your backyard to the Hubble, as I like to say. That kind of jump in resolution, so I think it's fair to say the CDC is very excited about whole genome sequencing for the kind of resolution they are getting for linking human illness back to food source as well."

Allard – "The PulseNet now does have a genomic component and all of the clinical isolates that have been sequenced are uploaded in to the NCBI database, which is where GenomeTrakr is part of that pathogen detection network and all of the foodborne isolates can also be downloaded by anyone, including the CDC, to be added into their database. So all the genomic sequences are being shared across the agencies."

Kelly – "Now if somebody was listening, consumers listening to this podcast who are not as familiar with whole genome sequencing, what is the most important thing for the average person to know about this technology and how it affects them?"

Brown – "Well, I'd say for the average consumer, every consumer has to eat and drink and because of that you want to know that your foods and beverages are as safe as they can possibly be and I'm happy to say that here in the United States you can rest assured that they are some of the safest in the world. Having said that, for us to be able to leverage this new technology that got its start in human bio sciences, and bring it over to food safety in this way has been extraordinary and one thing we'd like to tell the public is that we are leveraging this technology to help further ensure and enhance that safety

you have been accustomed to. Hopefully, if there is an event and something does get through, we will find the source much more quickly and remove it from the U.S. food supply so that our consumers can be as safe as they can possibly be."

Allard –"We are seeing that foodborne outbreaks even over the last year tend to be smaller and tend to be identified much more quickly, and it is because if you have certainty in a match you can even have just a few people that have gotten sick, and if we act on that information and those matches quickly, you can ask them if they have had exposure to those foods that have been identified and help remove them from the food supply. Some of the future tools are: more people having eyes on the data, better ways to respond rapidly, integrate activity between the agencies. These are important pieces that will continue to speed things up and to see more. In the past, because they didn't have high resolution tools there were a lot of things that would slowly get people sick. They called that background. What that meant was: They knew something was in the food, but they didn't know where, and they couldn't identify it because it was too spread apart, and maybe it was common genetic fingerprinting types. But with genomics you can see each outbreak and its separateness and its linkage to the food and environment and to clinical isolates. This key tool, what we used to call as background is going to be all identified to smaller specific outbreaks that are occurring, and as we clean up those industries that are having problems and contaminating the food supply this will also drop down."

Brown – "I just want to close by saying: One example that Marc was saying was in 2010, we had a massive shell-egg outbreak, one of the largest in the country. We weren't online ready to go yet but we did a retrospective analysis of that outbreak and had we been online we were able to show that not only could whole genome sequencing pinpoint the two farms that were leaking contaminated eggs into the food supply but it could actually separate and distinguish which farms were responsible for which illnesses. We knew at that point the kind of power we had in our hands here. So we worked hard to get it up and running full time as quickly as possible and really I think we were really honored and surprised to see that the president put out in his impact report that the genome tracker effort in whole genome sequencing was one of the most influential government programs that he listed for his tenure as president for the last 8 years. We were really honored and delighted to see that; it meant a lot."

Kelly – **"Okay, thank you so much. Thank you both."**

Allard – "Thank you."

Brown – "Thank you Sue, appreciate the opportunity."

---

[i] Plasmids: A Plasmid is an independent small circular, double stranded DNA molecule that carries only a few genes.
[ii] Phages: A phages is a virus that infects and replicates within bacteria.
[iii] Phylogeny: Much like a family tree, a phylogeny, also known as a phylogenetic tree, traces the evolutionary history and relationships of groups of organisms.