E. coli Variations in Washington State

[Announcer] This program is presented by the Centers for Disease Control and Prevention.

[Sarah Gregory] I have Dr. Gillian Tarr with me today. She's a postdoctoral epidemiologist with the University of Calgary. We'll be discussing *E. coli* in rural Washington state. Welcome Dr. Tarr.

[Gillian Tarr] Hi. Thanks for having me.

[Sarah Gregory] Your study is about *E.coli* and how it's affecting rural communities. Most people probably know that it's a serious disease, since it's been in the news so much, but what actually is it?

[Gillian Tarr] Well, *E. coli* is a gram-negative bacteria and there're actually a lot of types of it, most of which are completely harmless and incredibly common. The one that I was studying is called *E. coli* O157:H7 and it's one of what are called shiga toxin–producing *E. coli,* or STEC. O157 is the most commonly reported STEC serogroup, but there are many others. In the 1980s, *E. coli* O157 was linked to a disease called hemolytic uremic syndrome, what I'll refer to as HUS. And that's a pretty serious disease that can lead to kidney failure. And both O157 infection and HUS are most common in kids under the age of five, but all age groups are really at risk. And O157 is transmitted by the fecal-oral route and is often associated with foodborne outbreaks, and so that's where you'll see most of those news reports coming from.

[Sarah Gregory] The title of your study includes the words "geogenomic segregations and temporal trends." What do these terms mean?

[Gillian Tarr] Yeah, that refers to two different things, but first I think it would be helpful to define phylogenetic lineage, because that's really key to understanding these other terms in the context of our study.

[Sarah Gregory] Okay.

[Gillian Tarr] So, the evolution of *E. coli* O157:H7 through time can be summarized using what's called a phylogenetic tree. So, it's an evolutionary tree, and the more related two individual bacteria isolates are to each other, the closer they are on that tree. And so we can group sets of closely related isolates into what we call lineages, which are like the big main branches of the tree. And there are three of these phylogenetic lineages that are most common in human O157 cases in Washington state. So, when we say "geogenomic segregation," that's just a succinct way of saying that different O157 phylogenetic lineages are found in distinct geographic areas. They're not all found in the same proportion across space. The "temporal trends" part refers to the changes in the O157 population that we observed during the ten-year period study that we studied.

[Sarah Gregory] Okay. So, what were you looking for in your study?

[Gillian Tarr] There were really three things that we wanted to look at. The main thing, the first thing, was to investigate spatial distribution. And so, in Washington state, people had noticed

that *E. coli* O157 isolates with the same PFGE pattern, and PFGE is kind of like the bacteria's fingerprints, had been isolated from human cases in the same geographic areas over multiple years. One of my collaborators, Phil Tarr, had also seen something similar using whole-genome sequencing in a pair of Missouri cases that were separated by several years. And so this really sparked the question of whether strains were being maintained in particular geographic areas over long stretches of time. It's a hypothesis that's consistent with the increased risk of O157 in rural communities, as well as previous literature looking at persistence of O157 on farms. So that was the main goal, to test that hypothesis. We also wanted to look at two other things: One, had the makeup of the *E. coli* O157 population in Washington changed over time, and two, were there any risk factors for O157 generally that were actually more associated with some phylogenetic lineages than others.

[Sarah Gregory] How did you go about studying this?

[Gillian Tarr] I worked closely with the Washington State Department of Health. Hanna Oltean is one of the people that I collaborated with. We identified all the laboratory-confirmed *E. coli* O157:H7 cases that had been reported to the states from 2005 through 2014, so a ten-year period. The bacteria from each of those cases had been isolated and stored during the case investigation, and so my collaborators, Tom Besser and Smriti Shringi, at Washington State University were able to type those isolates and determine their phylogenetic lineage. So, and then we used a variety of methods to look at spatial distribution, temporal trends, and risk factors.

[Sarah Gregory] And what did you find? Were some areas more affected than others and were results tied to specific foods, such as raw fruits and vegetables or raw milk?

[Gillian Tarr] Well, we found a number of things. One was that a lineage that we call 2B was segregated in an area of the state without large-scale agriculture production. And so, when I say "segregated," I mean that, sort of, it was in one area of the state out of proportion with how it was found in the rest of the state. The other two major clinical lineages were really centered in regions with large cattle production operations. And that was expected, because cattle are the primary *E. coli* O157 reservoir, so it makes sense that we would see more of those lineages in areas that are more at risk of animal-human transmission. But the segregation of lineage 2B is really interesting because it's not found in a really agriculture-intensive area. And, what's more, we saw it across the entire period, so it wasn't just a single outbreak that was causing the signal.

You know, in 2005, we really didn't have much lineage 2B in the state, and then there was a spike, and then it dropped back down, but it never got back down to baseline. And so, that was really interesting. And during that period, we also saw a big shift in the O157 population. So, at the beginning of the period, lineage 1B was pretty much all that was in the state. There was a little bit of lineage 2A, a little bit of lineage 2B, but over time, lineage 2A really grew, and by the end of the period, in 2014, there were actually more lineage 2A cases than there were 1B cases. So, kind of ignoring the specific numbers and letters of the lineages, what, what we found was that we went from a very homogenous population to a much more heterogeneous population, over a ten year period. And what you were referring to around raw fruits and vegetables and raw milk, that was the risk-factor analysis that we did. I want to make sure that I emphasize, though,

that the results showing increased risk of lineage 2A with raw fruits and vegetables and increased risk of lineage 2B with raw milk were from an exploratory analysis. So, that means that we didn't start out with a hypothesis, we conducted a lot of tests. And that's because ours is the first study to look at preferential exposures by lineage, so we didn't know what to expect.

[Sarah Gregory] Okay. Have you drawn any conclusions from this study? And do you suggest any "next steps"?

[Gillian Tarr] Yeah, I think the big take-home message here is that it looks like the local environment has a sizable role in *E. coli* O157:H7 epidemiology, even though when we hear about it, we often hear about these large disseminated food vehicles, like spinach and the like. A lot of the, the drivers of risk look like they may actually come from local sources. And so, there are a couple of possible explanations for that. One is that a particular O157 strain could get into a reservoir in an area, circulate locally, and just occasionally cross over to a person and cause infection. That's one explanation. Another explanation is that lineages have a, sort of, preferential vehicle of infection, and particular vehicles are more common in some areas. So that's why we did the exploratory analysis looking at the association between common sources of infection and lineages. The other big conclusion is that the statewide composition of *E. coli* O157 is definitely not static. Like I was saying, we saw that big shift from a homogenous population to a more heterogeneous population, and that's interesting sort of ecologically, but also it could be important clinically, because different lineages tend to each have a dominant shiga toxin genotype and those genotypes are associated with more severe disease, some of them, i.e., more HUS.

And I, about next steps, you know, there are lots that are possible. You know, we started with using these lineages to sort of group isolates together. Whole genome sequencing would really let us look more deeply at the, the associations and the, between isolates and really trace that path to see if that hypothesis about, sort of, a bug getting into a local population and just circulating and occasionally causing cases, is, is true. That could help public health by helping to find a main reservoir for, for lineages that are segregated and persistent in an area. And, of course, if a main reservoir was found, then there could be steps to help control disease coming from that reservoir. And then, of course, there's the issue of what do other areas look like. Do other states or provinces or countries have pockets of segregated O157? Is it always lineage 2B that's segregated? You know, what's similar about those areas? What's different? And the same goes with, sort of, that temporal trend that we saw with lineage 2A kind of displacing lineage 1B, to some extent. Is that also present elsewhere? And we'd also like to look more closely at these associations between particular lineages and particular risk factors, like raw milk. You know, is that a real signal that we saw? Is that, sort of, a signal from a real effect? It would take more investigation to find that out.

[Sarah Gregory] So, next steps, a lot of next steps are further studies?

[Gillian Tarr] Yes. Yeah, yeah. There a lot of directions that this research can take us that could really help improve public health in regard to O157.

[Sarah Gregory] Okay. Great. Well, would you like to tell us about your job? What do you do and how your expertise was used in this particular study?

[Gillian Tarr] Sure, yeah. I conducted this project as part of my doctoral work in epidemiology at the University of Washington, and I've now moved on to a postdoctoral fellowship in pediatric enteric infections at the University of Calgary. Broadly, epidemiology is the science concerned with the distribution, causes, outcomes, and prevention of disease. And that disease can be something infectious, like I study, or it can be cancer, gun violence, etc. So, most of my work has been divided between the epidemiology of vaccine preventable diseases and zoonotic diseases, diseases that are transmitted between animals and humans, like *E. coli* O157. And this study was focused on the distribution of O157, and there are a lot of factors that can impact the distribution of a zoonotic disease. So, we really had to look at the causal model and think about which of those factors we would need to account for and how best to ask the question with the data available. And what I think this project really exemplified is, is the power of collaboration, which is really invaluable in a field like zoonotic disease epidemiology, where you need human and animal health professionals, microbiology, public health, methodologic experts, etc. And so, drawing on all of that expertise turns what started out as a vague idea about looking at distribution into a really well-formed hypothesis that was tested with solid data and some really powerful methods.

[Sarah Gregory] Thank you so very much, Dr. Tarr. I've been talking with Dr. Gillian Tarr about her January 2018 article, Geogenomic Segregation and Temporal Trends of Human Pathogenic *Escherichia coli* O157:H7, Washington. Listeners can read the article online at cdc.gov/eid.

I'm Sarah Logan Gregory for *Emerging Infectious diseases*.

[Announcer] For the most accurate health information, visit [cdc.gov](cdc.gov) or call 1-800-CDC-INFO.