Dr. Yeaman Clip #2: Introduction & High School Advice

I'm Sam Yeaman, I'm an assistant professor at the University of Calgary and I work in **bioinformatics** and **genomics**. We use genomics to try to understand **evolution**, broadly, and adaptation. And specifically, one of the main projects we work on a lot is looking at **adaptation** in trees—so, how trees deal with climate—but a lot of trees grow across really wide ranges in their environment. So, lodgepole pine grows in really cold areas and in really hot areas and in the same species [lodgepole pine] we are trying to see how differences in adaptation arise and how we can use genomics to study those.

One of the main things to do to use genomics and study trees, we actually use sample populations from across the range of a species. We sample seeds from way up north and seeds from down south and sometimes grow them in a **common garden** and see if there's differences in **phenotypes**—traits—like cold injury for example. And then we'll sequence the genomes—the **exomes**, so the exome is the part of the genome that's coding for particular genes—and we'll use techniques called **sequence capture** that focus in on just those coding regions in the genome, and allow us to sequence all of the letters of that genetic code from individuals from these northern populations or the southern populations. And so we'll approach this by sequencing hundreds of individuals from across the range of the species and then scan through those millions and millions and millions of base pairs of DNA code looking for associations between variants in the code that tend to be strongly associated with the environment. So we look for, say, if there's one letter that's in the code that's found more commonly up north and then others found more commonly down south, we'll try to see if that's happening much more than expected by chance and use statistics to kind of parse out those patterns in the genome and then use that to learn something about how the trees are adapting to their environment.

So, that's the main approach we use with genomics and that uses Illumina sequencing, which is short read sequencing. And we have a genome assembly and we basically sequence all the reads from these genomes in small little pieces and then align them back to the reference assembly and use that alignment to make calls about, say, 'this individual is a T/this individual is a G at this particular site'. And that allows us to really see all these enormous files that show all

sorts of millions and millions of letters for the code for each individual and then use computers and high-throughput statistics to really study that [enormous files of code].

Life is massively full of unanswered questions and it's like the last frontier is the frontier of knowledge. We're coming up against that frontier every day and pushing it just a little bit and it's a fascinating process to be a scientist because you're confronting the mystery inherent to life and really trying to get at that little by little and understand these fundamental questions about how life works or why it does things this way or that way. I find that an eternally interesting process and I think if you're interested in kind of answering and asking questions and confronting the fundamental mystery of life, science is a wonderful choice of a discipline to apply yourself to. And I find myself drawn to it every day when I get up and go to work.