Slide 1: Title slide

Hi, hello, my name is Cade Sterling, and I'm a senior biochemistry major in Dr. Clint Smith's lab. And today I'm going to tell you how to make coronaviruses.

Slide 2: Future coronavirus zoonoses are almost certain

And you might, rightfully, think why on earth would we want to make coronaviruses after the havoc SARS-CoV-2 has and is continuing to cause, but there's good news and bad news on that front.

The bad news is that coronaviruses have been with humans for centuries, they're emerging at what appears to be an increased rate, and on top of that we are constantly surrounded by non-human animal coronaviruses.

On the far left you will see the endemic human coronaviruses with their estimated dates of emergence in parentheses, and those are your common cold coronaviruses. In the center you'll see the highly pathogenic, epidemic coronavirus, all of which emerged in the last twenty years, and while it's a weak pattern, it does appear that coronaviruses are emerging at an increasing rate. And to our knowledge, all coronaviruses are of zoonotic origin.

In addition to many wild animals, especially birds and bats, there are coronaviruses in many animals we have constant contact with. And when you take into consideration habitat loss, increased human population, and increased human-animal interaction as a result it is hard to argue that coronavirus zoonoses will become less common.

And that term zoonosis basically means when an animal pathogen is able to jump to humans. Most zoonoses lead to dead end infections, but, occasionally, you get a virus like SARS-CoV-2 which is really good at infecting humans. So that's the bad news.

The good news is that the scientific capacity of humanity has never been higher, and I'm going to share with you a way we can make coronaviruses to study them, learn how they work, and ultimately break them.

Slide 3: Coronaviruses have the largest-known positive-sense, single-stranded RNA non-segmented genome

The first thing I want to do is orient you to the coronavirus genome. The coronavirus genome is maintained as positive-sense, single-stranded RNA which is functionally just mRNA so it can be translated immediately upon entry into a permissive host cell. So the genome is just one big

strand of RNA which is subdivided into several open reading frames that are beyond the scope of this discussion.

Slide 4: Reverse genetics allows us to go from genome to viral particles

The way we make viruses is using what is called a reverse genetics system, which is a fancy way of saying we take viral genetic material and end up with infectious viral particles.

Slide 5: Reverse genetics allows us to introduce mutations that change viral phenotypes

The other cool feature of reverse genetics is the ability to introduce specific mutations to proteins that modulate their structure and function.

Slide 6: Central dogma of biology

We can think of reverse genetics in terms of the central dogma -- DNA is transcribed into RNA, which is translated into Proteins. And of course, protein structure dictates protein function, and that structure is largely determined by the genomic encoding of protein primary structure.

Slide 7: Coronavirus RNA genomes complicate reverse genetics

Now this is complicated for coronaviruses, because coronavirus genomes are positive-sense, single-stranded RNA. When their genomes are replicated, it's by a special viral protein called an RNA-dependent RNA polymerase which transcribes RNA to RNA. But we want to make coronaviruses…

And RNA is not a very stable molecule. This is further complicated by the fact that coronaviruses have some of the largest RNA genomes ever observed at around 30 kilobases. It's more difficult to work with in the lab than DNA.

Slide 8: Reverse transcription generates DNA fragments

So, we subvert the central dogma a little bit with an enzyme called reverse transcriptase. This allows us to generate fragments of the viral genome as double-stranded DNA. At the bottom of the slide in green, you can actually see the ten fragments that we divided the mouse hepatitis virus genome into.

This is where things get interesting. Because now that we have DNA fragments of the viral genome, which presents us with both possibilities and a key problem. DNA fragments can be artificially synthesized at a reasonable cost, so we can introduce mutations to viral proteins by synthesizing one of these fragments. But, a fragmented coronavirus genome won't do us much good. So, we need a way to assemble these fragments.

Slide 9: Yeast quickly assemble plasmids via transformation-associated recombination (TAR)

That's where transformation-associated recombination (or TAR) cloning comes in. TAR leverages the high rates of homologous recombination in yeast to assemble large plasmids. You can think of transformation as a sort of horizontal gene transfer, it's basically the introduction of exogenous genetic material through a cell's membrane. Put even more simply, we're putting the DNA fragments inside yeast cells.

So once all the genome fragments and TAR vector are inside of the yeast, the homologous overlaps are treated as if they were a yeast chromosome which has been severely fragmented. The yeast goes to work connecting fragments by homologous recombination, and the result is a circular plasmid of double-stranded DNA with an assembled coronavirus genome.

Now I know you may be thinking, how on earth can you guarantee eleven pieces of DNA all end up in yeast… And the answer is, most don't. We incubate somewhere in the high millions to low billions of yeast with our DNA fragments, and only a few hundred end up with fully assembled genomes. But getting a few hundred is honestly pretty high efficiency, because we only need one.

Slide 10X: Yeast are selected post-transformation by histidine deficient media

In the photo you can actually see our yeast plates. Each of the dots is an individual yeast colony which should contain our plasmid. Those on the left are pink because that media was not given supplemental adenine, and adenine starved yeast turn pink.

Our yeast are histidine and adenine auxotrophs, which means they cannot synthesize the amino acid histidine or the nucleobase adenine on their own. Of the two, histidine deficiency is more detrimental which is why it is our selection agent. Our TAR vector encodes the gene for histidine synthesis, so a yeast which receives the TAR vector will be able to form a colony on histidine deficient media, and one which does not will perish.

Nevertheless, before moving to downstream steps, we want to verify that the plasmid has been assembled properly, and we do this using a junction verification assay.

Slide 11: Junction amplicons suggest proper viral fragment assembly with the TAR vector

We developed a multiplex PCR assay with twenty-two primers to generate eleven amplicons of unique size so they could be quickly distinguished via agarose gel electrophoresis.

We extracted total yeast DNA to use as PCR template, which should contain our assembled plasmid. On the left you can see a schematic of that plasmid with the viral genome fragments in green, TAR vector in grey, and the junction amplicons in purple. As you can see, the amplicons get increasingly larger by about a hundred base pairs.

On the right you can see the agarose gels where we ran the PCR product. If you take the time to count, you will see there are eleven unique bands. Using the molecular weight marker on the left we can confirm that each band is of the appropriate size. You may recognize the naming scheme of each lane as the viruses from Natalie's talk. These are actually the junction verification gels from those viruses.

It's worth noting that after this step we amplify our plasmid in bacteria to increase the amount of genetic material we have to work with. And we do that by transforming the bacteria with the plasmid and letting them grow.

So we have our viral genome as a double-stranded DNA plasmid, but if we want replication-competent virus, we need to get back to a positive-sense, single-stranded RNA genome.

Slide 12: CMV promotion allows for direct transfection of the assembled plasmid to establish infection

We decided to make our vector plasmid-launched. To do this, we placed a cytomegalovirus (or CMV) promoter immediately upstream of the coronavirus genome. This way, when we put the plasmid into mammalian cells an infection can be established. The terminology is a little different, but transfection is principally the same as transformation just with mammalian cells. It's still the uptake of exogenous genetic material.

Since the CMV promoter is taken from a DNA virus, it recruits host RNA polymerase II to transcribe the coronavirus genome which is then exported to the cytoplasm where it is translated as it would be in a natural infection, so the viral replication cycle can begin.

Slide 13: Overview of TAR viral recovery workflow

And now we have arrived. We have gone from the single-stranded, positive-sense RNA genome of coronavirus, to double-stranded DNA fragments, which were joined by TAR into a plasmid containing the viral genome. We amplified that plasmid in bacteria and then transfected it into mammalian cell culture where an infection was established. We can harvest the viral particles from that infection and study it using whatever methods we would like.

Great, but we wanted to make sure that this new process didn't lead to any unexpected changes in the virus, so we decided to characterize the wild-type virus generated by our system.

Slide 14: TAR-generated MHV-A59 has similar viral yield as the parental virus

This is the figure that really shows that virus produced using the TAR system have no observable deficit in viral replication or infection in cell culture.

To do this, we infected three cell cultures with the parental virus and three with our TAR-generated virus, and then we collected supernatants from those cultures at the indicated time points. We used a technique called plaque assay to enumerate the number of infectious viral particles in that sample which you see here.

Since these are wild-type viruses, we expect the infectious cycle to peak at around twelve hours and by twenty-four hours we'd expect the number of infectious particles to decrease as most of the cells have been killed and viral particles degrade over time.

That's exactly what we see here. You'll notice the Y axis is a log base ten scale, so the 6 is actually ten to the six or one-million plaque forming units per mL. The most notable thing, however, is that there was no statistically significant difference between parental or TAR-generated virus at any time point.

These data suggest that our new reverse genetic system produces viruses which are biologically equivalent to other wild-type viruses.

Slide 15: Acknowledgements

So I would like to thank Natalie Wilson, Debreiona Harris, and Dr. Clint Smith for their essential contributions to this project. I would also like to thank the University, the Department of Biology, and the Office of Undergraduate Research for their support. And finally, funding for this project was provided by the Appalachian College Association Ledford Scholarship, the Department of Biology, and the Henrietta Brown Croom Faculty-Student Research Award.

I'd be happy to take any questions you may have.