Polymerase Chain Reaction (PCR) and Reverse-Transcriptase PCR (RT-PCR)

**Polymerase chain reaction (PCR) and reverse-transcriptase PCR (RT-PCR)**

Our DNA contains loads of information, neatly stacked to insanely small sizes, able to fit within a cell nucleus. A single DNA molecule has two strands, which wrap one around one another to form a double helix. Each single strand of DNA is composed of a sequence of four types of nucleotides - which are the individual letters or building blocks of DNA. Nucleotides of DNA are made up of a sugar - deoxyribose, a phosphate, and one of the four nucleobases - adenine, cytosine, guanine, and thymine - or A, C, G, T for short. The nucleotides on one strand form hydrogen bonds to complementary nucleotides on the other strand; specifically, A bonds with T via two hydrogen bonds, and C bonds with G, via three hydrogen bonds. Additionally, the two DNA strands also have a “direction” - meaning, one of them runs from the 3’ end to the 5’ end, while the other one runs from the 5’ end to the 3’ end. Kind of like two snakes coiled up together, but facing in different directions. Every single protein of our body is encoded through combinations of just four nucleotides. When there are errors in our genetic information, diseases occur. And let’s be honest, we were always interested in knowing what was written in our DNA.

Polymerase chain reaction, or PCR for short, is a technique used in molecular biology to amplify a segment of DNA. Let’s take a step back. A single copy of DNA is not very much DNA. So to work with DNA, we basically make lots and lots of lots of copies of it, so that it’s easier to analyze. For example, if we want to visualize it, we can use a technique called gel electrophoresis.

PCR is based on DNA replication, a process that our cells normally use to duplicate their genetic material before dividing in two identical daughter cells. So first of all, we’re going to need a machine called a thermal cycler - that’s where the PCR magic happens. You can think of it like a cauldron filled with a solution, where genetic wizards add the ingredients. The ingredients are the DNA that we wish to multiply, an enzyme called Taq polymerase, specific primers, that bind to the DNA, and a mixture of free nucleotides - A, T, C and G. Throw everything in the thermal cycler, wave your magic wand, and the process begins.

So, let’s say that we have a long, double stranded DNA molecule, and we’re interested in the highlighted part. These two strands would be the template strands.

5’ - T T C **A G G T C A C A G T C C T G T A T G C C T A T G** T C C- 3’

3’ - A A G **T C C A G T G T C A G G A C A T A C G G A T A C** A G G- 5’

The first step of PCR is denaturation - meaning that we heat up our ingredients to exactly 96 degrees Celsius - almost as hot as boiling water. This breaks open all bonds between the two strands of DNA, so that they can separate from one another.

The second step is called annealing. Here’s where we need primers, and for our highlighted sequence, the primers would look like this:

 <----3’-  **C G G A T A C**- 5’

    5’ -  **A G G T C A C** - 3’ --->

During the annealing step, we cool everything down to around 55 degrees Celsius, and this allows the primers to bind to their complementary sequences on the single stranded DNA.

The third step is called extension - we want to make the environment as optimal as possible for Taq polymerase to do its job. Ok, see, Taq polymerase is actually a really cool enzyme. We get it from a bacteria called *Thermus aquaticus* that grows in geothermal hot springs. So not only can Taq polymerase withstand the heat during the denaturation step of PCR, but it also functions best at around 72 degrees Celsius. So during extension, we heat everything back up to 72 degrees, and Taq polymerase latches onto the primers, grabs some free nucleotides, and assembles the new DNA strands.

And that’s pretty much all there is to it! A whole PCR cycle lasts roughly 10 minutes, and at the end we’ve doubled the amount of DNA because each template strand of DNA is now double stranded! So, now, we just continue cycling these three steps a couple of dozen times - after every step, and the quantity of DNA gets doubled each cycle. 2 strands, then 4 strands, then 8, then 16, then 32, then 64, then 128, so on and so forth. So PCR actually multiplies DNA at an exponential rate - after around 6 hours and 40 cycles, we’ll theoretically have 2^40, or 1,099,511,627,776 copies - easier to remember as a whole lot of DNA that we can further analyze.

There are a lot of variations of PCR, but let’s discuss one that has been used recently for diagnosing a viral infection. Specifically, what if you wanted to diagnose someone with a COVID-19? To do that you would collect a nasopharyngeal swab which is where you take a long Q-tip with only one soft end, and twirl it a few times inside a nostril to get enough secretions on it to be analyzed. Now, getting a good sample is crucial, so make sure you really get in there; if no virus gets on the swab, then PCR can't detect it! To diagnose Covid-19 what you would look for is a virus called severe acute respiratory syndrome coronavirus 2, or SARS-CoV-2.

Now, SARS-CoV2 is an RNA virus, not DNA, so to detect you would need to use a particular type of PCR called reverse-transcriptase PCR; or RT-PCR for short. This technique is a little different, because it starts with RNA, instead of DNA. If you remember, one difference between RNA and DNA is that RNA doesn’t have thymine; instead, it has uracil. RNA is complementary to DNA, though, so if an RNA strand has a nucleotide sequence, 3’-AAG UCC AGU-5’, then the complementary DNA, or cDNA, strand would be 5’- TTC AGG TCA -3’.

To perform RT-PCR some of the extracted sample from the nasopharyngeal swab is added to a solution containing an enzyme called reverse transcriptase, nucleotides, and primers that are complementary to a specific SARS-CoV-2 target sequence. If the viral RNA is present, the primers attach to the RNA strand, and then reverse transcriptase synthesizes a cDNA strand. Once we've got the cDNA, the steps of RT-PCR are pretty much the same as PCR; Taq polymerase is used to amplify the cDNA through denaturation, annealing, and extension steps. After several rounds, if there was SARS-CoV-2 RNA in the original sample you’ll have amplified DNA from that sample that can be detected.

Alright, as a quick recap. Polymerase chain reaction is a technique used to massively multiply DNA, so that it can be analyzed using other techniques. It requires a thermal cycler, filled with the DNA sample to be multiplied, thermostable Taq polymerase, primers with which we select what gets multiplied, and free nucleotides of all 4 types - A, T, C and G. The first step is denaturation, then there’s annealing, and finally extension. PCR doubles the amount of DNA that we’re interested in looking at in each cycle.

***Author:****Luka Puzigaća, MD*

***Editor:****Viviana Popa, Rishi Desai, MD, MPH*