PCR Optimization Package

Lionel Tukei and Lily Torp
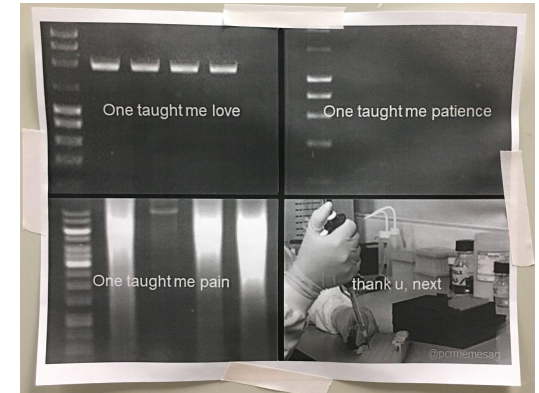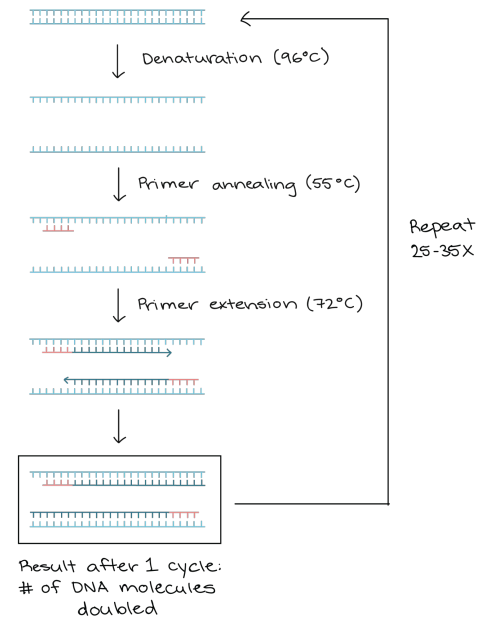
Background

What is PCR?

- *Polymerase chain reaction (PCR)*: a common molecular biology technique for amplifying DNA sequences.

What is the problem?

- **Many factors to consider** when designing PCR:
  - Enzyme cost per reaction ($$)
  - Primer annealing temperature
 - Primer-gene compatibility
  - Total PCR reaction time

Denaturation (96°C)

Primer annealing (55°C)

Primer extension (72°C)

Repeat 25-35X

Result after 1 cycle: # of DNA molecules doubled

One taught me love

One taught me patience

One taught me pain

thank u, next

PCR Protocol Optimizer

A python package for:

1. Checking gene and primer sequences for **errors** and **evaluating primer binding.**

2. Optimizing a PCR protocol for **cost**, **time**, or **both factors** depending on **user preference.**

Design
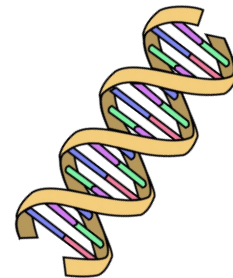
One Object:

1. **PCR**
   a. Gene, forward primer and reverse primer are input in 5'-3' format as strings. DNA template type is also specified (plasmid, lambda, BAC DNA, or genomic).
   b. Object stores information and contains relevant functions

Two Functions:

1. **Check**
   a. check() takes a PCR object (gene, forward primer, reverse primer) and checks sequences for errors and compatibility. User will fix any errors.
2. **Recommend**
   a. recommend() will optimize PCR based on user-defined factor ("time", "cost", or both if left blank)
   b. Return a table with results for factor

Use Cases

Ex: **A user wants to optimize their PCR reaction for cost**

1. User will install pcr_optimizer
2. Define a new pcr object containing gene, forward primer, reverse primer, and template type.
3. Run **check** function:
   a. Ensure all bases are standard (A/T/G/C) and primers anneal in correct locations on the target gene sequence
   b. Statements are returned to the user specifying errors. They will manually fix them and redefine the pcr object
4. Run **recommend** function
   a. User can designate time or cost as the optimization factor. Function will return a table for **(1) Cost:** enzyme amount and cost per reaction or **(2) Time:** annealing temperature/time, extension time, and total PCR reaction time.

Demo

[Demo Google Colab Notebook](#)

Project Structure

Lessons Learned and Future Work

Software engineering lessons:
- Functional and component specifications are important!
- Small details ➜ big errors
- Creating a package and uploading to PyPI is cool and useful

Future work:
- Add more enzymes/protocols
- Use Dataframes instead of text output
- Have function evaluate multiple factors and choose one overall "winner"
- Model PCR reagents over time
- Include cost of reagents