Lab 4: PCR Primer Design

Today we'll design primer pairs for two applications: first, a primer pair for amplification of the ribosomal internal transcribed spacer (ITS) of a diverse fungi community and second, a primer pair which could be used in a PCR based test for identification a poisonous mushroom variety.

DUE BEFORE NEXT LAB: The lab write-up (submit via Blackboard) for this lab will consist of answers to the questions (and 2D structure images) in Part I, and a detailed search strategy and initial results from Part II.

Introduction

The polymerase chain reaction (PCR), invented by Kary B. Mullis in 1985, revolutionized molecular genetic research. It allowed the production of millions of copies of a specific segment of DNA sequence from a sample where the sequence was present in very low, often previously unanalyzable quantities. There are now many specialized adaptations of PCR, many of which rely on specially designed primers.

The copies are generated by PCR are synthesized by a DNA polymerase that builds a complementary strand of DNA along a template strand. The DNA polymerase requires an open 3' OH from which to begin synthesizing the new strand. The 3' OH is generally provided by a "primer" - a short fragment of DNA complementary to the template strand. To amplify (make copies of) a region of double stranded DNA there must be one priming location on each side of the region to be copied. The polymerase builds the new strand in the 5' to 3' direction, the polymerase is "reading" the template strand from the template's 3' to 5 end'. The synthesis of the complementary strand generates a double stranded molecule. Each new strand synthesized generates a new priming site, and itself becomes a template strand for synthesis of a new strand. The double stranded molecule must disassociate so that the priming sites are accessible to primers for the next round of synthesis.

The two strand of double-stranded DNA dissociate at high temperatures. However, many enzymes, including most DNA polymerases, are permanently denatured at temperatures sufficient to dissociate double stranded DNA. The solution to this problem (PCR) won Dr. Mullis the 1993 Nobel Prize in chemistry. Dr. Mullis' invention relies on thermostable DNA polymerases - polymerases which are not denatured by high temperatures. The most widely used thermostable polymerase (commonly referred to as "Taq" - pronounced Tac) was isolated from a heat loving bacterium (Thermus aquaticus) first discovered in the hot springs at Yellowstone National Park. Though Mullis initially practiced his invention by manually cycling the temperature of the reaction, machines known as "thermocyclers" have automated the temperature cycling.

The key considerations in primer design become apparent when considering what primers are doing at each step of a PCR cycle. (See Figure 1)

PCR starts with a “cocktail” containing the purified DNA sample, primers, DNA polymerases and deoxyribonucleotide triphosphates (dNTPs, the individuals nucleotide “building blocks” of DNA) in a buffer. The concentration of template DNA and primers and the total volume of the reaction also affect the specificity of PCR - but further discussion of these aspects is beyond the scope of this lab.

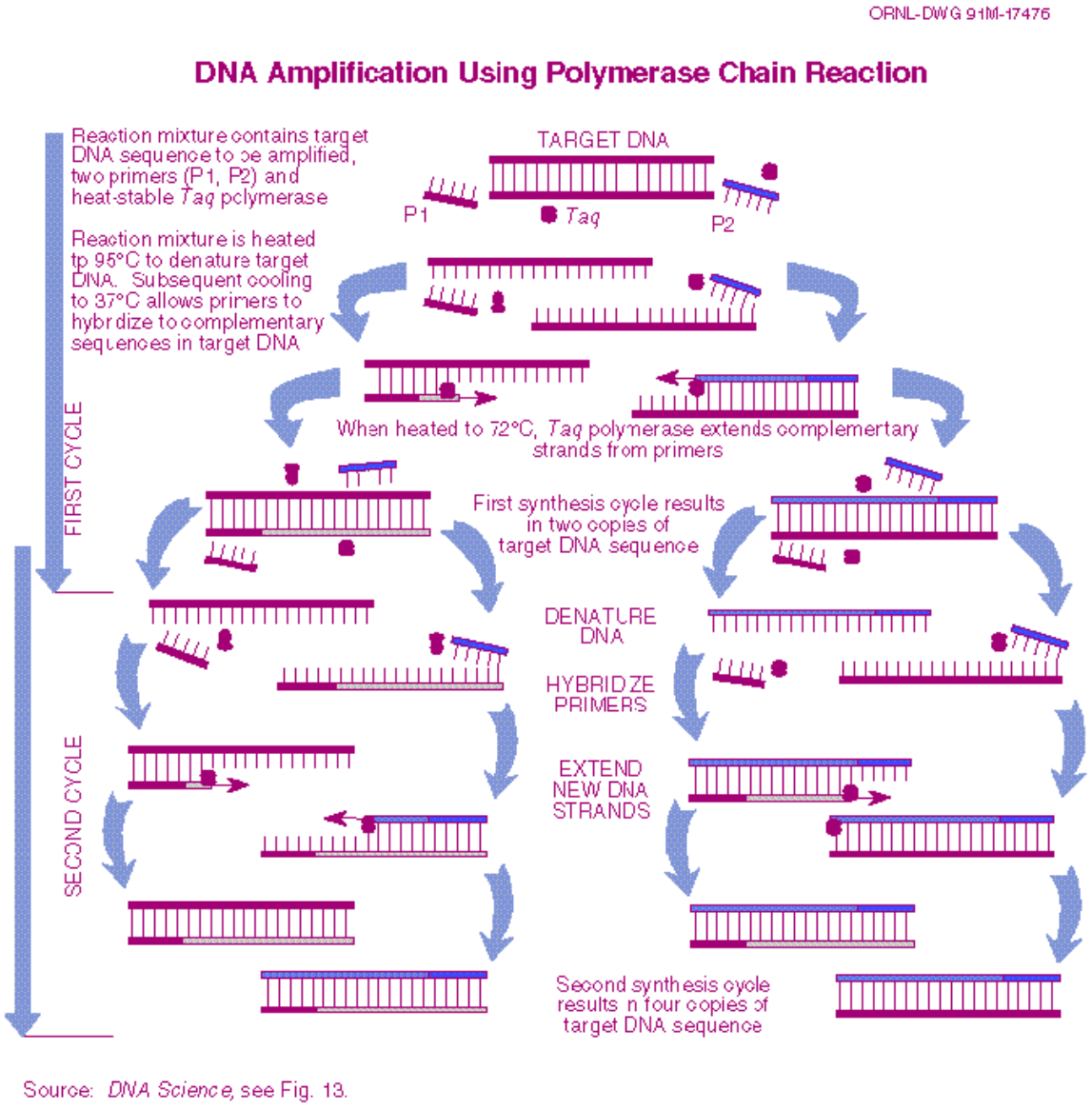

Figure 1: PCR overview. From http://www.ucl.ac.uk/~ucbhjow/b200/pcr.htm.

Step 1: Denature the double stranded DNA molecules **Temperature ~ 95 C**

This step breaks apart secondary structures formed by base-pair interactions.

Step 2: Annealing **Temperature ~ 45-65 C**

At the cooler temperature it is more energetically favorable for the single stranded DNA molecules form double-stranded structures. Primers compete with the strands of the sample

DNA to anneal (form base-pairs) with complementary sequences. (See Figure 2) Primers should not be self-complementary or complementary to each other as this results in “primer-dimers”. Primer-dimers decrease primer concentration and lead to a lower yield of the target product.

Mis-priming occurs when the primers don’t bind to their exact sequence or bind to other regions within a DNA sample. The specific annealing of the primer with sequences flanking the target region is what allows PCR to generate a specific product. For specificity, it is suggested that primers be at least 20 bp long.

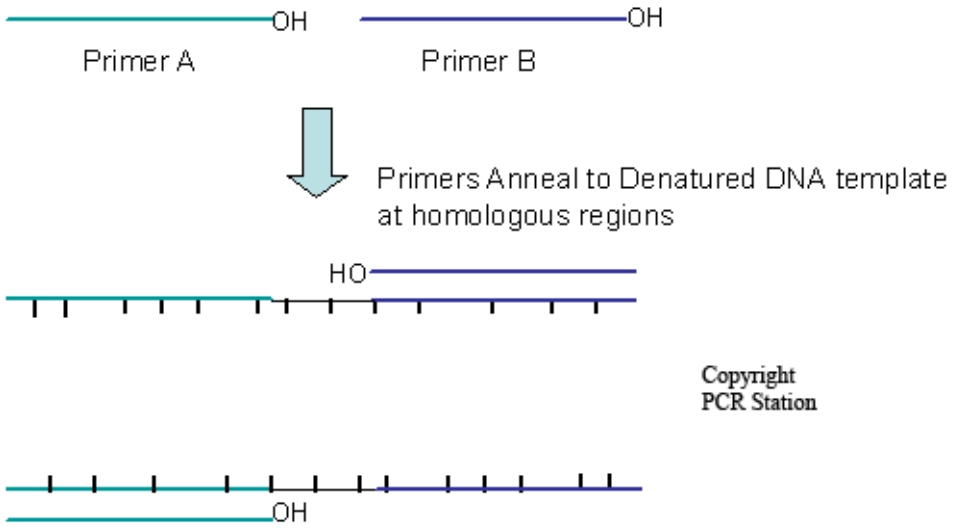

Figure 2: Primer annealing step. From http://www.pcrstation.com/pcr-primer/.

Q1) Without considering repetitive elements or regulatory motifs, what is the probability of finding, by chance, a particular 20 base-pair long sequence within a genome?

Annealing temperatures which are either too high or too low can result in non-optimal annealing (thus non-optimal amplification). When the annealing temperature is too high, primers may not form stable relationships with their complementary sequences. When annealing temperature is too low, primers can form stable relationships with sequences which are not perfectly complementary resulting in additional unwanted amplification products.
In general the annealing temperature is 5 degrees lower than the melting temperature of the primer-template hybrid. The melting temperature is the temperature at which the double-stranded molecules (or secondary structure) dissociate.

Q2) The melting temperature of double-stranded DNA is highly dependent on its nucleotide composition. Why?

Extension **Temperature ~ 72 C**

The DNA polymerase associates with the 3’ OH of the primer-template hybrid and extends the double-stranded region by incorporating nucleotides complementary to the template strand.

Q3) It is generally not advisable to have 3 or more G or C nucleotides at the 3' end of a primer. Why?

The time allotted to the extension phase of the cycle is determined by the length of the desired PCR product. A polymerase can generally synthesize a new strand at a rate of 50-100 nucleotides per second.

At this point the cycle repeats. In a typical PCR the "denature, anneal, extend" cycle is repeated 20-30 times.

Q4) If one copy of your target sequence is initially present in your sample, how many copies of the target would you have after 20 cycles of PCR?

Part I - Fungi Diversity

Generally PCR with single primer pair results in a population of DNA molecules all having exactly the same nucleotide sequence. However, the unique characteristics of the regions encoding ribosomal RNA (rRNA) allow us to sample genetic diversity within populations by amplification of many template DNA sequences with a single primer pair ("multi-template PCR"). Ribosomes (and their rRNA components) are essential to one of the most basic cellular processes, translation, so not surprisingly there are many copies of rRNA sequences within a single genome and these sequences are highly conserved. The rRNA molecules are often found in clusters and transcribed as a single unit separated by internal transcribed spacers (ITS) that are later cleaved from the transcript and degraded. Because the ITS regions are not functional they are not under the same stringent selective evolutionary pressure as the flanking rRNA and thus the ITS sequences accumulate significant sequence heterogeneity. The highly conserved rRNA sequences flanking an ITS are convenient regions in which to "anchor" primers.

We're going to use a variant of BLAST (Primer-BLAST) to help us pick our priming sequences. Primer-BLAST performs a modified BLAST search (optimized for short query sequences) to find other places in the genome where mis-priming might occur.

Step 1. Here are the accession numbers for rRNA clusters in different Amanita mushrooms. Each team member should chose a different accession number to work with.

Species	Accession Number
Amanita abrupta	AB015685.1
Amanita caesarea	AY486237.1
Amanita caesareoides	AB759116.1
Amanita esculenta	AB721453.1
Amanita hemibapha	AB759083.1
Amanita phalloides	HQ641120.1
Amanita rubescens	JF313652.1

Amanita similis	AB750727.1
Amanita vaginata	AB015691.1
Amanita virosa	AB015676.1
Agaricus bisporus	JX684007.1

Step 2. Locate the corresponding GeneBank record.
Take a minute to look over the fields of the GeneBank record. In what year was this sequence submitted to GeneBank? Is the record associated with a journal publication? Is there information on the source this DNA sequence was isolated from? If the answer to any of these questions is no, ask if your teammates' GeneBank record has this information.

Find the annotation for the internal transcribed spacer 1 (ITS1). What is the rRNA on the 5' side ("upstream") of ITS1? What is the rRNA on the 3' side ("downstream") of ITS1? Write down the beginning and ending positions of each of these regions.

The two members of the primer pair are referred to as the forward and reverse primers. The forward primer is also known as the left-handed primer because it will be on the left-hand side of the PCR product, while the reverse or right-handed primer will be on the right-hand side of the PCR product.

Remembering that DNA sequences are by convention written in the 5' to 3' direction, and recalling that DNA polymerases build the new strand in the 5' to 3' direction . . .

Q5) Will the forward primer contain sequence identical or complementary to the FASTA sequence? Will the reverse primer contain sequence identical or complementary to the FASTA sequence?

Step 3. From the GeneBank record, click "Pick Primers" (best to open this in a new tab, you may want to refer back to the GeneBank record) - this will direct you to NCBI's primer designing tool, Primer-BLAST.

Step 4. Enter a range of positions for the forward and reverse primers. Briefly discuss your thinking with your team. Without changing the default settings, click "Get Primers". Note any error messages you receive, you might need to change the database used in the specificity check.

Some of the accession numbers listed only have partial sequences for the rRNA gene upstream of ITS1. If you can't find a good forward primer within the partial sequence and you may need more sequence in which to anchor your forward primer. Use the species and gene name to search NCBI Nucleotide for the complete sequence, use bl2seq to see how the partial and complete sequences lineup. Carefully copy and paste the sequences together so that you have a complete upstream rRNA, ITS1 and downstream rRNA. Then try again.

Write down the parameters and the sequences, location and characteristics of the two best two primer pairs from this query ("Query 1") on the chart provided.

Q6) What do the numbers for self complementarity mean?

Step 5. Adjust the Primer Pair Specificity parameters and perform two different queries with more stringent parameters. Write down the parameters and the sequences, location and characteristics of the two best two primer pairs from each query ("Query 2" and "Query 3") on the chart provided.

Step 6. Adjust the Primer Pair Specificity parameters and perform two different queries with parameters less stringent than default. Write down the parameters and the sequences, location and characteristics of the two best two primer pairs from each query ("Query 4" and "Query 5") on the chart provided.

Step 7. From the all the primer pairs returned in queries 1-5, choose your best two primer pairs (four sequences total). We're going to use IDT's OligoAnalyzer - take a minute to review the instruction page. (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/)
One at a time: copy and paste a sequence into IDT's OligoAnalyzer and Analyze!

Is the melting temperature the same as the one listed by Primer-BLAST?

Q7) Why might the melting temperatures differ?

Does the sequence form a hairpin? Save an the image of the secondary structure in a word processor.

Q8) What is the difference between a homo- and hetero-dimer? Does the sequence form a homo-dimer? Do the sequences form a heterodimer?

Repeat melting temperature, homo-dimer, hetero-dimer analysis with the other member of your primer pair.

Repeat with the other primer pair.

Step 8. Decide which is your best primer pair.

Q9) List the sequences you chose as you best primer pair. Describe why you chose this pair, what criteria did you consider?

Q10) Compare your primer pairs with those found by your teammates. Do they differ in their position relative to the beginning and end of ITS1 or in their sequence? How many nucleotides differ?

Step 9. Test the specificity of your best primer pair, use Primer-BLAST without a template and enter your forward and reverse primer sequences in the "Use my own [forward/reverse] primer" boxes. Be sure the specificity check-box is checked, select a database, and don't limit the query to a specific organism, then click "Get Primers".

Q11) How many of the results were in the genus Amanita? Were your teammates' species in the results? How many results were not mushrooms? [see NCBI Taxonomy Browser, "Agaricales" for a list of gill mushroom genera]

Part II. Poison primers

PCR based assays are commonly used to diagnose viral infections and in forensic DNA analysis because they can detect the presence of specific sequences present in very low quantities.

Many of the members of the Amanita family are extremely poisonous to humans, but some members are edible. Your job is to propose detailed primer design strategy for a PCR based assay to distinguish a extremely poisonous mushroom, Amanita phalloides, mushroom from an edible look-alike, Amanita caesarea. You'll run initial queries and propose targets for further research, but we don't have the time or resources to fully develop and test this assay - so no one should test out the accuracy of their primers at home!!!

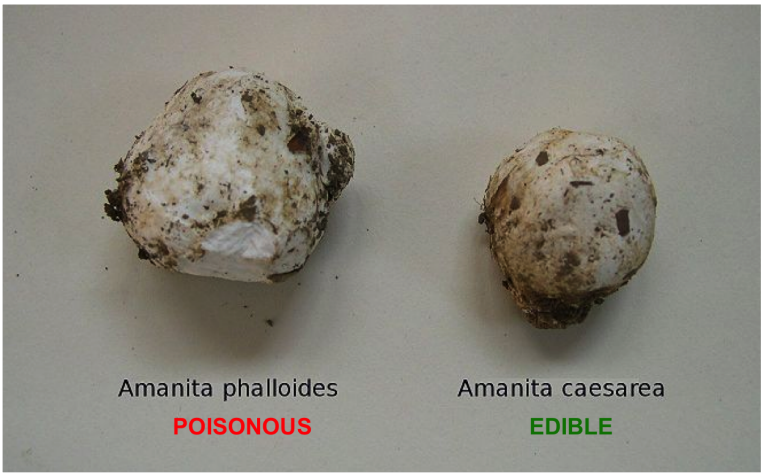

Photo: http://en.wikipedia.org/wiki/Mushroom_poisoning

Step 1. Using the accession numbers listed above for the rRNA genes of Amanita phalloides and Amanita caesarea and the tools from Part I, propose a primer search strategy.

Items to consider include (but are not limited too!):
- Description of the potential results of the hypothetical PCR assay in terms of true and false positives and true and false negatives.
- What are the benefits of targeting a sequence with a rRNA cluster? Drawbacks?
- Within which portion of the rRNA cluster are your primers located? Explain your reasoning.

- Which species will your primers detect?
- Which species or taxa will you include in your specificity checking? Which database will you use?
- What settings did you chose for the primer specificity stringency? Explain your reasoning.

Step 2. Run some test queries, adjusting parameters as necessary (make note of the queries, parameters and results on the tables provided) and from the results pick the best primer pair.

Step 3. Test the specificity of your best primer pair, use Primer-BLAST without a template and enter your forward and reverse primer sequences in the “Use my own [forward/reverse] primer” boxes. Be sure the specificity check-box is checked, select a database, and don’t limit the query to a specific organism, then click “Get Primers”.

Include in your write up:
- The sequences of your best primer pair(s) and their melting temperatures and secondary structure characteristics.
- Comment on the specificity of your best primer pairs.
- What is the size of the PCR product produced by your best primer pair?
- Based on your best primer pair, include the temperatures for each of the stages of a PCR cycle. Consider the consequences of false positives and false negatives while making your calculations. Be sure to explain your reasoning.