Appendix C. Phillips *et al*.
Primer Design Exercise

Currently Monsanto owns the patent on glyphosate, which is commonly known as Roundup®. It is the most popular herbicide used today because it kills a broad spectrum of weeds and is easily broken down into non-toxic compounds. The catch is that Monsanto also owns the patent on the gene that confers resistance to glyphosate, which they have transformed into several crops such as corn and soybean to make them "Round-up Ready", or resistant to glyphosate. Many researchers are trying to find novel genes that will also confer resistance to glyphosate for both evolutionary and economic reasons. Recently, a Chinese group found a bacterium, *Pseudomonas putida* strain 4G-1, which is naturally resistant to glyphosate[1]. They have cloned the novel gene, *aroA*, that is significantly different in sequence from the previous *AroA* gene, and are hopeful that it will be another source of glyphosate resistance.

The *AroA* gene encodes the enzyme 3-phosphoshikimate 1-carboxyvinyltransferase, which plays a key role in the biosynthesis of aromatic amino acids. Glyphosate works by disrupting this enzyme and thus the biosynthesis of aromatic amino acids. Resistance is found by mutating the *AroA* gene such that glyphosate cannot bind to the resulting protein.

You have just been hired to identify strains of *Pseudomonas putida* with the novel *aroA* gene to provide a new glyphosate resistant cultivar from a glyphosate-rich soil sample. Your first challenge will be to create primer pairs (forward and reverse) that will amplify a portion of the *aroA* gene that contains the underlined region (see sequence below).

1. Given the sequence on the following page, choose forward and reverse primers that will amplify the underlined portion of the *aroA* gene.

a. Underline the primer sequence on the following page. Then write out your primers below and indicate the 5' and 3' ends. Remember that the 3' or reverse primer is the reverse complement of the template (think about which direction DNA extends).

b. What is the size of your target DNA? (Note: each line contains 70 nucleotide bases)

[1] Sun, Y. et al. 2005. Novel AroA with high tolerance to glyphosate, encoded by a gene of *Pseudomonas putida* 4G-1 isolated from an extremely polluted environment in China. *Applied and Environmental Microbiology*. 71 (8): 4771-4776

>gi|51587624|emb|AJ812018.1| Pseudomonas putida aroA gene for 3-phosphoshikimate 1-carboxyvinyltransferase

5′-

GATCATAAAACATGCTTGTATAAAGGATGCTGCCATGTTCCGTGAACTGGAAGCGAACAATCTTGCGGTA
TATCAGAAAAAGCCAAAGCTGATTGCAGTGCTTCTTCAGCGTAATGCTCAGTTAAAAGCGAAGGTTGTTC
AGGAGGATGAGTTCGAAAAGTCGGTAAGGCGTTTGTTGAACTTTGGTCATACATTGGGGCATGCCATCGA
AAATGAATATGCGTTGATGCATGGCCATGCGGTTGCTATAGGAATGACATACGCGTGTCATATTTCTGAG
CAATTGTCTGGATTCAAACAAACAAATCGCGTGGTAGAAGTGTTGGAACAATATGGGTTACCGACTTATA
TGGCATTCGATAGGGAAAAGGCTTTTAATCTGTTGAAAATGGACAAGAAGCGTGAAAAAAAGGAAATGAA
CTATGTGTTGCTGGAAAAA**GTAGGGAAGGGAGTGGTGAAGAG**<u>**TATTCCACTGGTTCAATTAGAAAAAATC**
ATTCAAGCATTACCAAAGTGAAAGTAACAATACAGCCCGGAGATCTGACTGGAATTATCCAGTCACCCGC
TTCAAAAAGTTCGATGCAGCGAGCTTGTGCTGCTGCACTGGTTGCAAAAGGAATAAGTGAGATCATTAAT
CCCGGTCATAGCAATGATGATAAAGCTGCCAGGGATATTGTAAGCCGGCTTGGTGCCAGGCTTGAAGATC
AGCCTGATGGTTCTTTGCAGATAA</u>CAAGTGAAGGCGTAAAACCTGTCGCTCCTTTTATTGACTGCGGTGA
ATCTGGTTTAAGTATCCGGATGTTTACTCCGATTGTTGCGTTGAGTAAAGAAGAGGTGACGATCAAAGGA
TCTGGAAGCCTTGTTACAAGACCAATGGATTTCTTTGATGAAATTCTTCCGCATCTCGGTGTAAAAGTTA
AATCTAACCAGGGTAAATTGCCTCTCGTTATACAGGGGCCATTGAAACCAGCAGACGTTACGGTTGATGG
GTCCTTAAGCTCTCAGTTCCTTACAGGTTTGTTGCTTGCATATGCGGCCGCAGATGCAAGCGATGTTGCG
ATAAAAGTAACGAATCTCAAAAGCCGTCCGTATATCGATCTTACACTGGATGTGATGAAGCGGTTTGGTT
TGAAGACTCCCGAGAATCGAAACTATGAAGAGTTTTATTTCAAAGCCGGGAATGTATATGATGAAACGAA
AATGCAACGATACACCGTAGAAGGCGACTGGAGCGGTGGTGCTTTTTTACTGGTAGCGGGGGCTATTGCC
GGGCCGATCACGGTAAGAGGTTTGGATATAGCTTCGACGCAGGCTGATAAAGCGATCGTTCAGGCTTTGA
TGAGTGCGAACGCAGGTATTGCGATTGATGCAAAAGAGATCAAACTTCATCCTGCTGATCTCAATGCATT
TGAATTTGATGCTACTGATTGCCCGGATCTTTTTCCGCCATTGGTTGCTTTGGCGTCTTATTGCAAAGGA
GAAACAAAGATCAAAGGCGTAAGCAGGCTGGCGCATAAAGAAAGTGACAGAGGATTGACGCTGCAGGACG
AGTTCGGGAAAATGGGTGTTGAAATCCACCTTGAGGGAGATCTGATGCGCGTGATCGGAGGGAAAGGCGT
AAAAGGAGCTGAAGTTAGTTCAAGGCACGATCATCGCATTGCGATGGCTTGCGCGGTGGCTGCTTTAAAA
GCTGTGGGTGAAACAACCATCGAACATGCAGAAGCGGTGAATAAATCCTACCCGGATTTTTACAGCGATC
TTAAACAACTTGGCGGTGTTGTATCTTTAAACCATCAATTTAATTTCTCATGAATAGCTTCGGCCGCATC
TTCAGGGTGCATATTTTTGGCGAATCACATGGTGAATCAGTAGGCATCGTTATTGATGGTTGTCCTGCTG
GTCTGTCATTGTCCGAAGAAGATC-3′

2. For each primer you designed, use the website and the guidelines on the instruction sheet to determine whether they meet the basic primer requirements. Record the Tm, length, molecular weight, and possibility of secondary structures or primer dimers and use the statistics to qualify your decision to use or not use the primers for a PCR reaction.

3. To determine the specificity of your primer pair to the *aroA* sequence, run a nucleotide BLAST by following the directions on the instruction page.

a. What does the E-value indicate?

b. Write down the organism and E-value score from the two highest matches for each primer sequence. Did you get back the sequence you put in?

c. How many nucleotides aligned between your sequences and the first match for each?

4. Now that you have some experience designing primers and an idea about what makes a good primer, check the CamV35S primers used to determine the presence of a transgene in your food products.
Sense (forward) primer: 5’GCT CCT ACA AAT GCC ATC A3’

Antisense (reverse) primer: 5’GAT AGT GGG ATT GTG GGT CA3’

a. Provide a short summary about the primer pairs including characteristics such as melting temperature, secondary structures, and specificity using this worksheet as a guide.

b. Do you think these primers are a good choice to amplify the 35S promoter from your food product? Consider the specificity of your primer pairs to transgenic ingredients as well as other ingredients that may be in your products.

Appendix C Continued Phillips *et al*.
Primer Design Instructions

The advent of the Polymerase Chain Reaction (PCR) brought about the ability to rapidly make many copies of a segment of DNA. The PCR reaction depends on short pieces of DNA, called primers, to bind to the denatured DNA strands and act as a starting point for replication. Primers are usually 17-24 nucleotides in length and must be specific to the target DNA. Listed below are some rules and helpful tools that scientists use in designing primers. Finally, there are instructions on how to perform a nucleotide BLAST to check the specificity of each primer. Read through the information and then use these instructions to complete the Primer Design Exercise.

Rules for primer sequence design (adapted from Sambrook and Russell, 2001[2])

1. *To increase specificity, primers should range from 17-28 nucleotides in length and the forward and reverse primers should not differ by more than 3 base pairs in length from each other.* If you consider that there is a ¼ (4^{-1}) probability of finding an A, T, C, or G in any given DNA sequence then there is a 1/16 or 4^{-2} chance of finding any two nucleotide sequence such as AC. Therefore, a specific 16 base pair sequence will statistically occur once in every 4^{16} or approximately 4 billion bases. Thus complementary binding of a sequence that is greater than 16 base pairs is an extremely sequence-specific process.
2. *The base composition should have 40-60% (G+C) content*. GC content is used to determine the annealing temperature of the reaction. A GC content of 40-60% will ensure that the melting temperature is above 50ºC.
3. *Melting temperatures (Tm) between 52-62º C are preferred and it is best that the Tm, of the two primers, is within 5º.* Melting temperature indicates the stability of the double helix. Tm above 65ºC may form secondary annealing structures and should therefore be avoided.
4. *The 3'-ends of primers should not be complementary to any other part of the other primer, otherwise the ends will pair and create a primer dimer preferentially to any other product.* If primer dimers are formed early in the PCR, they can reduce the amplification of the target DNA by competing for the DNA polymerase, nucleotides, and primers.
   i.e. 5'-ACGTCCAAGAGG**AAGC**G-3'
   3'-ACTTACGACT**TTCG**AATGTAA-5'
5. *Primer self-complementarity should be avoided so that secondary structures are not formed or are very weak.* The example below shows how the underlined portions of the primer could pair with itself or another primer molecule.
   i.e. 5'-AGTCCTA**GTC**CGAATGTTCT**GAC**

Calculating Melting Temperatures of primers and target DNA
(adapted from Sambrook and Russell, 2001[1])

1. "The Wallace Rule" T_m (in ºC) = 2(A+T) + 4(C+G)
   (A+T) = sum of A and T residues in the primer

[2] Sambrook, S. and Russell, D. W. 2001. Molecular Cloning: A Laboratory Manual. New York: Cold Spring Harbor Laboratory Press.

(C+G) = sum of C and G residues in the primer

2. DNA calculator
   This website will allow you to fill in your primer sequence to determine the length, Tm, and molecular weight for your primers and if there will be any secondary structures or possibilities of primer dimers.
   http://www.sigma-genosys.com/calc/DNACalc.asp

Determining the uniqueness of the primer sequence:

Specificity to the target gene is one of the most important properties of a primer. One way to determine the uniqueness of the primer sequence is to compare the primer to all other sequence data using a bioinformatics tool that is a Basic Local Alignment Search Tool, otherwise known as BLAST. The BLAST tool is maintained by the National Center for Biotechnology Information (NCBI) and can be found at the following link: www.ncbi.nlm.nih.gov/BLAST/ . There are many types of BLASTs, but we will only be dealing with the nucleotide BLAST (blastn) to determine the specificity of the primers.

1. Go to the above website and click on the link in the Nucleotide box titled "Nucleotide-nucleotide BLAST (blastn)".
2. Type the primer sequence into the text box.
3. If you would like to limit the genomes that are searched, scroll down to the options box and using the drop down box, select your organism of choice. For our purposes we will choose Bacteria as there is not an individual listing for *P. putida*.
4. Click on the BLAST! icon.
5. A new page will appear entitled, "formatting BLAST". Click on the icon "Format!" and wait for the results of the BLAST.
6. The first figure on the results page provides an overview of the database sequences that are aligned to the query sequence, which is your primer sequence in this case. Following the figure is a list of sequences that are significantly similar to the query sequence. Each matched sequence contains a link to view the entire sequence, the name of the sequence, a score and an E value. The Expect value (E) is a number that describes the number of matches a particular sequence would get by random chance. The lower the E value, the less likely the match would happen by chance and thus making the similarity of the query sequence to another sequence more significant.

   Scientists will choose their own levels of significance, but often only values smaller than e^{-100} are considered a significant match. As long as there are not any significant matches in the species you are working in, the primers will amplify only your gene of interest. Conversely, if you are using a primer that has significant alignment in multiple genomes, remember that E-values can only be compared between matches within a single genome.
7. Scroll down and see how the base pairs align with the sequence matches of the same species. If more than 5 matches are found, the primer is not specific enough.

After designing the primers and checking the quality, the primer sequence is sent to a DNA synthesis library such as Sigma-Genosys or Integrated DNA Technologies. Primers

are synthesized, desiccated, and returned to you. Once you have received the primers, you need to re-hydrate the primers in sterile water or Tris- EDTA buffer.