Day 3, PCR Session: Designing PCR primers

A walk-along with the first lecture

Find a genome sequence to use while following the Primer3Plus demo.

1. Go to ncbi Genome: <http://www.ncbi.nlm.nih.gov/genome/>
2. Browse by Organism
3. Pick the Viruses Tab
4. Enter SARS in the Search bar
5. Click on the organism
6. Click on the RefSeq ID (NC\_004718.3)
7. In the right corner, click on the arrow next to Change region Shown, when the selection boxes appear enter 26938 in begin and 27063 in end.
8. Click on FASTA
9. Copy and paste the sequence to a Word file, save as text on your desktop.

Using the Primer3 Plus PCR design software

1. Bring up your browser and direct it to
2. <http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>
3. Can anyone tell me what this input sequence is on my presentation, and how you would find out?
   1. Hint: the Sequence ID has a gene identifier, used by ncbi – what is the original sequence?
   2. How do I figure out what subset of the sequence was input?
4. Enter gi|30271926:26938-27063 in the Sequence Id box.
5. Provide some excluded regions – for example you could avoid the nucleotides from 371,5 because there are 5 Gs in a row and it is hard to synthesize primers with a log of Gs.
6. Include a region that is essential to your experiment – for a bar code region this would be the core sequence that is needed for a good statistical score. Maybe for this product you are going to clone it, express it in bacteria and see if the protein makes a good epitope for a vaccine – pick a starting position and length that you want (protein epitopes are usually 6-10 aminoacids, this would be 18-30 nucleotides)
7. Check the boxes for having the software pick the left primer and right primer.
   1. A hybridization probe is used in a different kind of assy, ignore this one.
8. Declare the product sizes (must be less than your complete sequence, of course (which is 666nt in this case). You can highlight a region with the mouse and click on the button, or write the numbers in the box.
9. Now go to the General Setting Tab
   1. The primer size is what is easy to synthesize – this can be very short (6 nt) to quite long (35-40nt) but usually is between 15-25.
   2. The melting temperature can vary from room temperature to about 75C – this depends on where the primers stay bound to the template so the polymerase can bind to a double-stranded region, so usually 50-60C is what you see.
   3. Balance the GC content against the genome/region of interest. Human DNA is about 50% GC, most plants are about 30% GC and bacteria are all over the map.
   4. Monovalent cation is the salt in the buffer you use to keep the polymerase functioning, it is usually Na+, often K+, and occasionally NH4+. It can range from 20-100mM, but is usually 30mM.
   5. Annealing Oligo – this is in nM and has to do with the size of the genome and its concentration. In practice 20-50nM works well.
   6. All polymerases need Mg2+ in order to work, this is usually 1-5mM.
   7. You need building blocks (free dNTPs) – usually you use 0.2mM of each (but it can range from 0.1-0.4) and since there are 4 of them you multiply by 4.
   8. There are lists of known simple sequence repeats for the most-studied genomes – in this case we re looking at a human pathogen, so probably there will be some human DNA around when I purify my sample. I don’t want to amplify a human gene, so I am going to check my primers against the human genome as well as the viral genome.
10. Advanced setting
    1. Poly-X means the longest homopolymer run you will allow in your primer (for example we have a run of 5 G’s in our sequence, we might want to set 4 instead of 5 in order to avoid this).
    2. Maximum Ns – this avoids ambiguous base calls that you cannot design around.
    3. The various mispriming parameters have to do with repeat sequences in the library – I usually leave these on the default settings.
    4. Product Tm – I usually don’t set this – you might want to if you have an assay that is going to use heat and you either want the PCR product to melt into 2 single strands below that temperature, or you want it to be stable at the temperature.
    5. Set the PCR product size range you will allow (obviously in our example it won’t be above 666).
    6. Liberal Base – maybe you have a SNP at one position where the primer could land – you can design in mixed bases in the synthesis process so you might want to allow this.
    7. The Sequencing parameters have to do with Sanger-type sequencing and sequence walking, you can ignore these for NGS sequencing.
11. Pick Primers!
    1. Red is the excluded region
    2. Green is the included region
    3. The top primer pair are yellow and lavender highlighted.
    4. I can look at each set and decide which ones I like the best.