What bacterial species do we host in our mouths?

April 26, 2013

0.1 Prelab: Molecular Biology Toolbox

This experiment will introduce you to an array of commonly practiced molecular biology methods. After this lab and with your previous exposure to DNA digestion, electrophoresis, culturing microorganisms, you will already have in your toolbox everything you need to design state-of-the-art synthetic biology, evolutionary biology or microbiology experiments! Even more exciting is that by the end of the next couple of labs you will know what bacterial species are in your mouth and how they are evolutionarily related. So that the following descriptions are crystal clear, you will need to look up the following topics. (It's not mandatory to write down in your lab notebook about these topics but it's in your best interest to spend some time if you aim to fully understand this lab and the upcoming bioinformatic exercises.)

1. Central Dogma, transcription and translation machinery, bacterial genomes (plasmids) versus eukaryotic genomes

2. Polymerase Chain Reaction (PCR)

3. Phylogenetic trees and 16S universal Primers

4. Sanger sequencing

0.2 Identifying our mouth bacteria using the 16S barcode

In the previous lab, you extracted DNA from the pool of microbes in your mouth. To identify which species of bacteria comprise the gene pool, we resort to a universal barcode that every bacterium carries, namely the ribosomal RNA gene. This gene is highly conserved across the bacterial world, and variations in its sequence give away the identity of the bacterial species that harbors them. In building phylogenetic trees, the evolutionary relatedness of different bacteria is judged by how closely their 16S ribosomal RNA sequences match up (See Figure 1).

To amplify the DNA sequence coding for 16S, molecular biologists have designed generic primers that amplify the DNA sequence associated with the ribosomal subunit, 16S. To amplify DNA we will use a premixed enzyme solution containing a polymerase (machinery involved in DNA replication from a double stranded template), dNTPs (DNA nucleotides that are input for amplification of DNA from a template sequence), and $MgCl_2$. We will also add our template DNA, which is the bacteria DNA you extracted previously, and add water. Your TAs will instruct you on what temperature cycles to impose on these PCR reactions, and the exact recipe for making the PCR reaction mix in order to achieve the highest amplification efficiency. Briefly, the thermocycling process can be described in three steps: 1) denaturation of double-stranded template DNA into two single stranded template DNA molecules, 2) Annealing of 16S primers to the single stranded template and the start of polymerase activity at double-stranded sites, and 3) elongation. The thermocycler ensures that this three-step process is repeated approximately 30 times so that the template is exponentially amplified (See Figure 2). After PCR amplification of

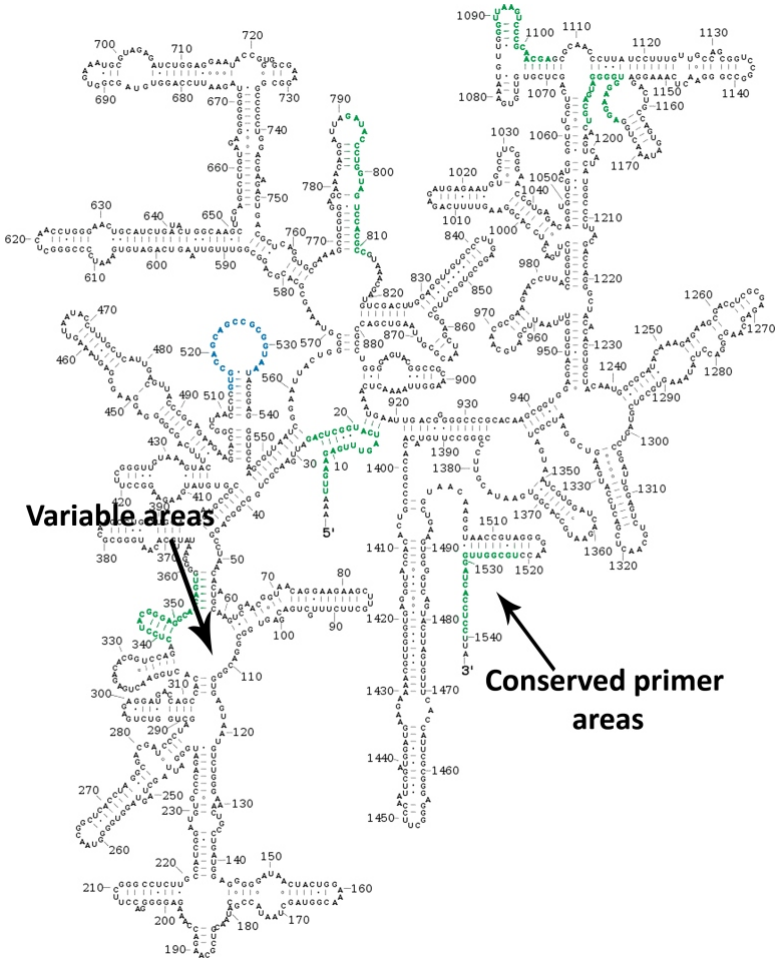

Figure 1: *E. coli* 16S RNA: conserved and variable regions. The variable regions are used for determining the evolutionary relationships of organisms, whereas the conserved regions are used for primer design. (1)

the 16S DNA sequence, we will clean up the reaction from all enzymes using a commercial kit (Qiagen), leaving just amplified DNA.

We will then use a reaction called TOPO cloning (See Figure 3). The purpose of this step is to separate out each 16S gene from the mixed gene pool of amplified 16S sequences, and to individually send these identifiers for sequencing. We cannot simply send a heterogeneous mixture of DNA sequences for Sanger sequencing. The TOPO cloning reaction will be performed using a commercial kit that contains a vector, which is a circular double-stranded DNA (See Figure 4), salt solution, and water. The purpose is to integrate each 16S sequence into one of the TOPO vectors. The integration of a 16S gene into a TOPO vector happens at room temperature during just 5 minutes. The vector has been optimized for this purpose.

From there, we will mix this reaction with electrocompetent cells (cells that after electrical shock will suffer pores in their membrane and consequently uptake foreign DNA, such as the TOPO vector). After cells have been electrically shocked, we will add growth media in which they can repair the damage and continue to replicate. After an overnight incubation of cells at 37C, we will plate the cells, and colonies will form. Note that TOPO vector contains both Kanamycin (Kan) and Ampicillin (Amp) resistance genes. Hence,

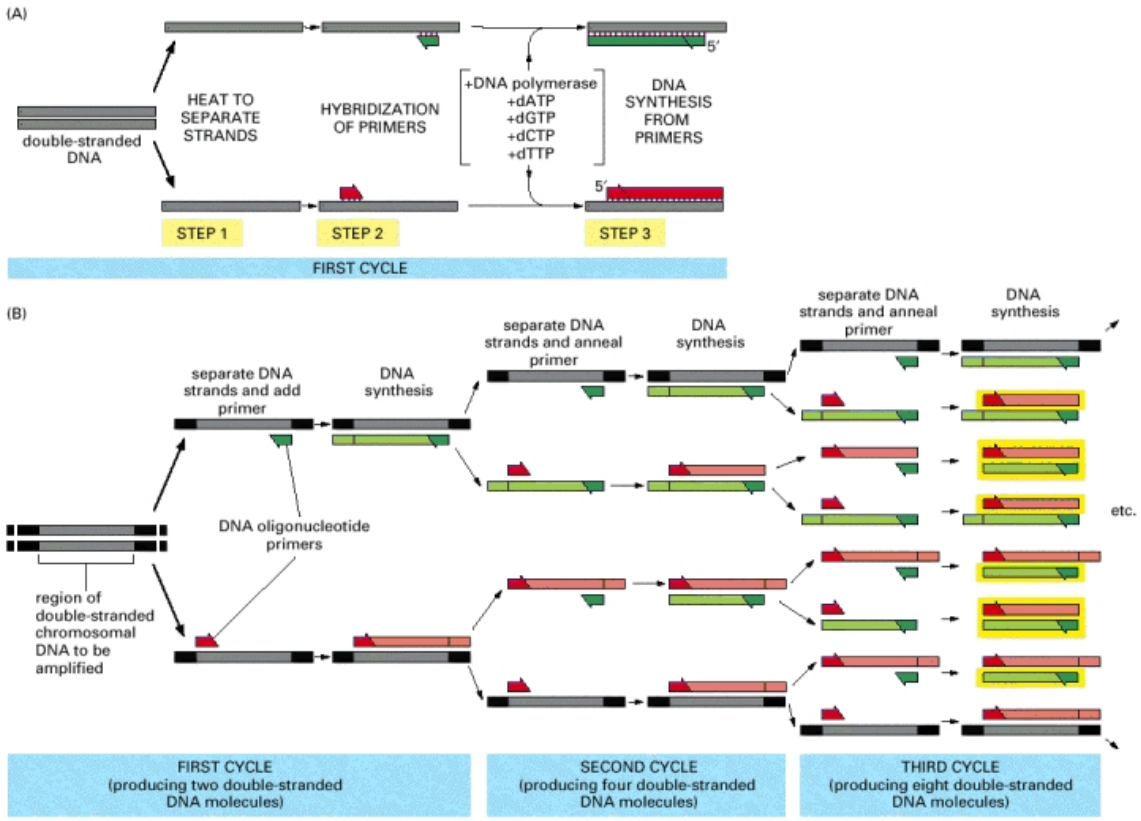

Figure 2: DNA amplification using PCR. First, double-stranded templates dissociate into single-stranded DNA at temperatures close to 95C. Then, the temperature is lowered to 55- 65C and primers bind to templates, forming double stranded DNA. The polymerases bind to double stranded sites and elongate the sequence. These steps are then iterated 30 to 40 cycles and the template DNA is copied exponentially. (2)

we will plate cells on plates containing either one of those antibiotics so as to ensure that only those bacteria that have taken up the TOPO vector can form colonies. Each colony will have formed from one cell, and therefore will only carry one 16S gene. We will pick several colonies, run a PCR with TOPO primers to amplify the 16S signal from each colony, and send the purified PCR product for sequencing.

After sequencing results are sent back to us from Laragen (a sequencing company), we will use the NCBI databases to compare the sequences against those reported by other researchers. We will pick the bacteria whose 16S sequence best matches the 16S sequence that we mined out from our teeth. Later, you will use these sequences to build a phylogenetic tree, and see the evolutionary relationships between the various microorganisms in the human mouth. We will provide with more detailed description of databases, sequence alignments, and other bioinformatical tools in the following weeks.

Sources

1. http://www.accugenix.com/microbial-identification-bacteria-fungus-knowledge-center/micro-id-basics/16s-rna-bacterial/

2. *Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition.*

3. http://tools.invitrogen.com/content/sfs/manuals/topotaseqman.pdf

Reagent*	Volume
Fresh PCR product	0.5–4 µL
Salt Solution	1 µL
Water	add to a total volume of 5 µL
TOPO® vector	1 µL
Final Volume	**6 µL**

Figure 3: Topo cloning reaction recipe. (3)

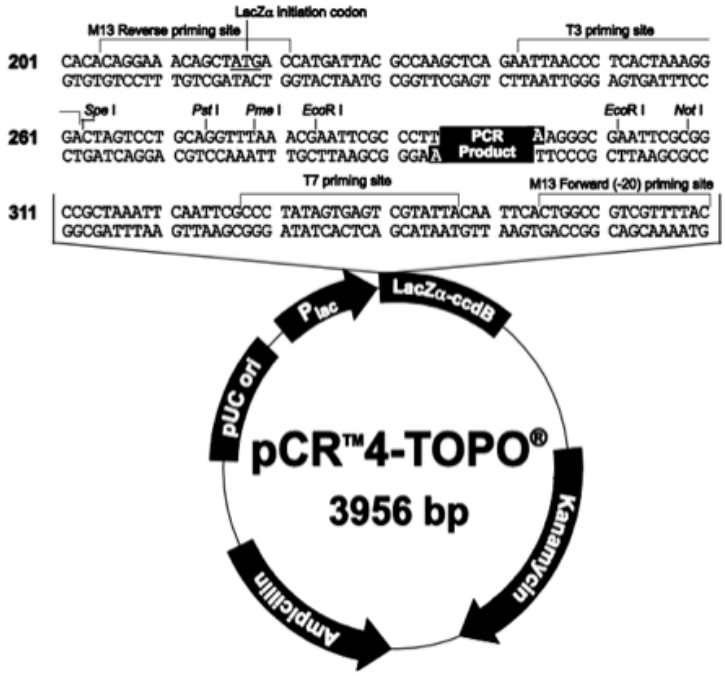

Figure 4: Topo vector. The insert will be incorporated just before the start of LacZ gene. The vector has two antibiotic genes so that bacteria that contain this vector will be able to grow in the presence of both antibiotics. (3)