Answers to Challenge #3

Question #1. Consider the following four individuals:

```
          1 2 3 4 5 6
Calvin:   AATGTA
Hobbes: A - T - TA
Snoopy: ACCATG
Tom:      A - T - TA
```

a. How would you find polymorphisms (SNPs, indels) and what are they?

Any column with more than two nucleotides is a SNP. So column 2, 3, 4 and 6 in the alignment

Any column with a gap character is an indel. So columns 2 and 4 above

b. How many alleles are there?

As many of you said, # of unique strings is # of alleles. So three total.

Question #2. Generate a random DNA string of size 8-10 nucleotides.

```
          0 1 2 3 4 5 6 7
String:  AAAATTTT
```

a. Using the string->binary conversion in Lecture 7 (slide 6), what is the lookup table using a dinucleotide index.

0000 (AA) index 0 – 0, 1, 2
0011 (AT) index 3 – 3
1111 (TT) index 15 – 4, 5, 6

b. In your own words, what is the advantage of a lookup table?

All instances of AA (or the 15 others) can be located by looking at index 0 in linear time

Question #3: Using the (zoom) whiteboard, draw a picture of the puppy/grill markov model