Modeling DNA Sequencing with Beads

Students will engage their analytical skills as they sequence a gene and interpret their data, using only simple, inexpensive materials.

Part One – Introduction to Dideoxy DNA sequencing

Materials

- Pony beads --- white (approximately 250) and four colors (about 10 of each red, blue, green and yellow) per group[
- Paper lunch bags or other opaque container – four per group
- Chenille stems (formerly known as pipe cleaners) – 2 or more per student
- "DNA Template" with instructions (at end of handout); one per student

Preparation

- Label the bags. Each group requires four bags, one labeled "G", one labeled "A", one labeled "T", one labeled "C".
- Add beads to the bag so that around 15% are of one color, and the rest are white. For example, we used about 60 white beads and 10 colored beads per bag. Color codes for the beads are:

white = regular nucleotides
G = yellow (in "G bag only); represents ddGTP
A = green (in "A" bag only); represents ddATP
T = red (in "T" bag only); represents ddTTP
C = blue (in "C" bag only); represents ddCTP

To Do and Notice

Before beginning this activity, *review DNA base pairing as appropriate.*

- What do you notice about the DNA template? (it is single stranded)
- Fill in the spaces for the *complementary strand* of the DNA template. Has everyone in the class filled it I correctly?
- The original template strand serves as the "unknown" DNA and will be the template for DNA sequencing:

- String beads onto the chenille stem, using the order of the nucleotide sequence of the complementary strand of DNA. Do not look into the bag! ;
- When a colored bead is selected, add it to the growing strand, but then STOP – do not continue adding. Put the strand aside.
- Repeat the process with a second chenille stem, and a third if available.
- Each student determines the number of nucleotides in their strand, and which nucleotide ends the strand (if it is a colored bead.)
- Collect the class data on the board, beginning with the shortest strand and ending with the longest. Who has a strand that is one base (bead) long? What color is the last bead, and what nucleotide does it represent? Who has two

bases? Three? Etc. Record the results on the board. Is every base of the DNA strand represented? If not, why do you think it is so?

Part Two – Reading Dideoxy Sequences

Materials

• Pony beads, white, red, blue, green, and yellow (hundreds for whole class)
Key for the nucleotides: G = yellow; A = green; T = red; and C = blue.
• Chenille stems or yarn of 3 different colors
• Plastic sealable bags for completed sets of "DNA fragments"

Preparation

• Create sets of dideoxy sequence fragment by stringing the appropriate beads onto yarn or chenille stems as per the chart at end of handout. There will be 15 fragments of increasing size per set, and each of the three different sets are strung on a different color of yarn or chenille stem. Store individual sets in sealable plastic bags.
• Provide each student group with one set of 15 fragments, being sure that each of the three possibilities are represented at least once.

To Do and Notice

• Remove the "dideoxy sequence fragments" from the bag. Place the fragments in order by length.
• Determine the DNA sequence by reading the last base (bead) using the key G = yellow; A = green; T = red; and C = blue.
• Record each of the three different sequences on the board, and have the students copy them onto a sheet of paper.
• Do any stretches of the different sequences overlap? How? Can they form one long strand of DNA?

What's Going On?

Dideoxy DNA sequencing (also known as the "chain termination method" or "the Sanger method" is based on terminating a replicating chain of DNA at a known spot. Nucleotide analogs (called "dideoxy" nucleotides") lack an oxygen atom at the site where the next nucleotide in a growing chain would attach. This alteration prevents the addition of the next nucleotide, hence terminating the growing chain (see Figure 1).

The essential process of dideoxy sequencing is as follows: Cloned and purified DNA is denatured by heat. A primer to one of the strands is added (required for DNA replication from a template strand), along with DNA polymerase, nucleotides (G, A, T, and C) and dideoxynucleotides (ddNTP's – ddG, ddA, ddT, and ddC). The reactions proceed, (sometimes in four separate reactions with ddG, ddA, ddT, and ddC , or everything in one reaction tube depending on the specific method. The replication begins, and the chain elongation is terminated

when a ddNTP is incorporated. Because this reaction takes place thousands of times, the result is a collection of newly synthesized DNA strands that can be separated based on length, each one base longer than the last. Because we can identify the ddNTP that terminated the strand, we can "read" the entire replicated sequence.

In *Part One – Introduction to Dideoxy DNA Sequencing,* the template DNA strand given to the students is the "unknown" strand that will be replicated in the sequencing process. Students reproduce the complementary strand by choosing beads from the appropriate bag and placing them on the chenille stem, modeling chain elongation. Each of the four bags of beads represents a different sequencing reaction, since each contains a small percentage of colored beads that represent ddNTP's. When a ddNTP is selected and added to the growing chain, the elongation of that DNA chain stops. Students repeat the process many times, and the strands from the entire class represent the results of the sequencing reactions.

Just as in the genuine sequencing reactions, we can only read the final nucleotide in each strand, in our case based on its color. We don't know what any of the others bases are by looking at our chenille stem. Recording class data data of the shortest sequence to the longest sequence reinforces that the replication of the template is begun and ended thousands of times, and that repetition increases the probability that terminated strands will increase in length by one base, and that all possibilities will be represented. The sequence will read TACGGATACCTG.

Notes:
• If one or more of the strands aren't represented, you may have the students repeat the activity a few more times, and/or add extra ddNTP's to the bags for the next activity.
• In this activity, we did not initiate DNA replication with a "primer" sequence on the chenille stem, but it is easily added if desired.

Part Two – Reading Dideoxy Sequences illustrates the sequencing process in automated systems, such as those used in the human genome project. A different colored fluorescent molecule is attached to each of the ddNTP's. All of the sequencing reactions occur in the same tube, and the products are separated by length, by a technique called capillary electrophoresis. The shortest fragments migrate the fastest through the electrophoresis matrix, the longest fragments migrate the most slowly. A laser excites the fluorescent molecule attached to the ddNTP's, and a detector distinguishes the fours types. The time of migration indicates the length of each fragment. A computer) records the length and the final base. Compiling all the data, the computer can reconstruct the sequence of DNA, and look for areas of overlaps from other clones.

In this activity, each group was given one set of fragments from one of three separate reactions. Students performed electrophoresis and separated the strands by length and recorded the sequence based on the final colored bead, the ddNTP. Comparing the sequenced from the three different reactions, overlaps were identified, and a longer sequence was compiled, just as is done by computer analysis of read data.

So What?

This powerful technique has revolutionized biology. Over 180 genomes have been sequenced, ranging from viruses and bacteria to human beings. The information gathered from the sequences has provided us insights into disease and health, evolution, and human origins.

Etc.

• The technique modeled in this activity was developed by Frederick Sanger and called the *chain termination method, dideoxy termination method, or Sanger method*. It's development earned him the Nobel Prize in chemistry in 1980 (his second chemistry Nobel!)

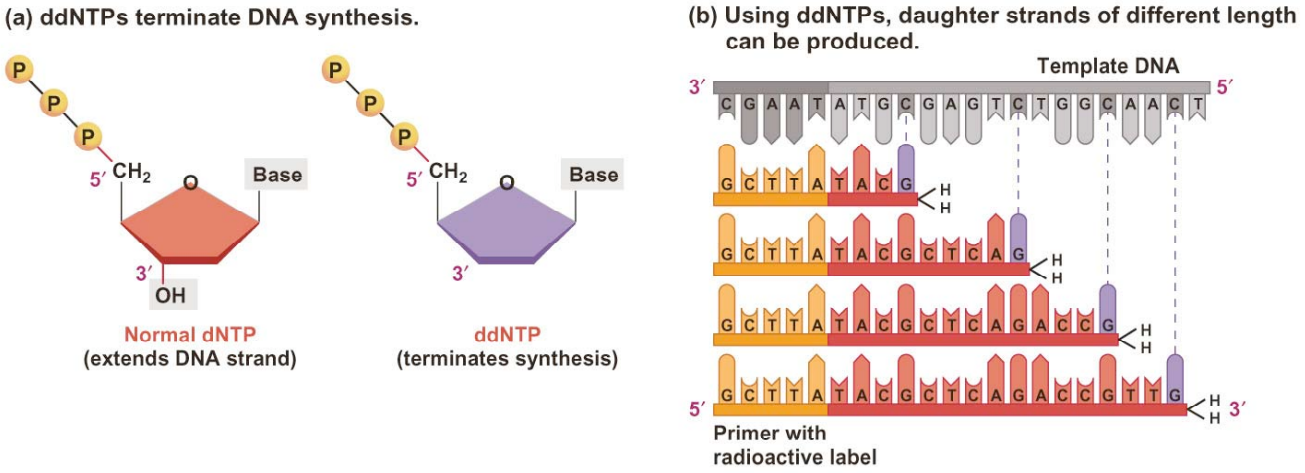

Figure 1. From: Queen's University of Charlotte
http://campus.queens.edu/faculty/jannr/Genetics/images/dnatech/

Template – Part One: Introduction to Dideoxy Sequencing (Cut in Half)

3’ end 5’ end
A T G C C T A T G G A C

__ __ __ __ __ __ __ __ __ __ __ __

5’ end 3’ end

Fill in the blanks above with the sequence of the complementary strand of DNA

Key to the Beads (Parts One and Two)

G = Yellow T = Red
A = Green C = Blue

3’ end 5’ end
A T G C C T A T G G A C

__ __ __ __ __ __ __ __ __ __ __ __

5’ end 3’ end

Fill in the blanks above with the sequence of the complementary strand of DNA

Key to the Beads

G = Yellow T = Red
A = Green C = Blue

Part Two: Making the DiDeoxy DNA Strands/Key to Activityy

Note: this DNA sequence is part of the Cytochrome P450, 2D6. The first four white beads on the strand represent the primer sequence, which is added to the reaction and is necessary for DNA polymerase to function.

Key for Nucleotides:
G = yellow; A = green; T = red; C = blue (actual colors of the fluorophores used in automated dideoxy sequencing)

Set One #White; Letter		Set Two #White; Letter		Set Three #White; Letter	
5	C	5	G	5	A
6	G	6	C	6	T
7	G	7	G	7	C
8	A	8	T	8	T
9	T	9	G	9	G
10	G	10	A	10	G
11	T	11	G	11	G
12	G	12	C	12	A
13	C	13	C	13	A
14	A	14	C	14	C
15	G	15	A	15	A
16	C	16	T	16	G
17	G	17	C	17	T
18	T	18	T	18	G
19	G	19	G	19	C

Complete sequence and areas of overlap

5' end **3' end**

CGGATGTGCAGCGTGAGCCCATCTGGGAAACAGTGC full sequence

CGGATGTGCAGCGTG first set

GCGTGAGCCCATCTG second set

ATCTGGGAACAGTGC third set

Etc.
It is a convention that DNA sequences determined by dideoxy DNA sequencing are presented in the 5' to 3' orientation (left to right), since DNA replication occurs in that direction, and we are, essentially, replicating an "unknown" strand.

A Second Set of DiDeoxy DNA Strands

Note: this DNA sequence is part of the ShigaToxin 2c that confers disease-causing properties to *E. coli 0157:H7* strains that are often the cause of serious disease outbreaks from eating contaminated meat, or recently, from eating spinach grown in California

Key for Nucleotides: G = yellow; A = green; T = red; C = blue

Set One #White; Letter		Set Two #White; Letter		Set Three #White; Letter	
18	T	5	C	5	C
19	C	6	C	6	T
20	T	7	G	7	G
21	G	8	G	8	A
22	A	9	C	9	G
23	G	10	T	10	A
24	C	11	G	11	T
25	T	12	C	12	A
26	C	13	C	13	G
27	A	14	T	14	C
28	C	15	C	15	G
29	C	16	T	16	T
30	G	17	G	17	T
18	G	18	A	18	G
19	C	19	G	19	A

Complete sequence and areas of overlap

5' end **3' end**

TCTGAGCTCACCGGCTGCCTCTGAGATAGCGTTGA full sequence

TCTGAGCTCACCGGC first set

CCGGCTGCCTCTGAG second set

CTGAGATAGCGTTGA third set

Etc.

For those of you keeping track of directionality, in this case, the *complementary strand* of our sequence would have been used as the template for dideoxy synthesis, since DNA replication occurs in the 5' to 3' direction.

Overlap of the Three Sets of DiDeoxy Fragments