A Simulation of DNA Sequencing Utilizing 3M Post-it® Notes.

Doug Christensen

Wayne State College, 223 Carhart Science, 1111 N. Main
Wayne NE, 68787

Email: dochris1@wsc.edu

Abstract: An inexpensive and equipment free approach to teaching the technical aspects of DNA sequencing. The activity described requires an instructor with a familiarity of DNA sequencing technology but provides a straight forward method of teaching the technical aspects of sequencing in the absence of expensive sequencing equipment. The final sequence analysis can be used as a springboard to a number of activities including literature reviews and writing projects.

Keywords: DNA sequencing, low technology

Introduction

As a teaching professional in the area of molecular biology, I have seen a substantial increase of student interest in learning how bands on X-ray film or within a gel can be interpreted as irrefutable evidence. Many small colleges and universities have little or no molecular equipment which they can utilize to foster the interest their students show in learning more about the topic.

This article provides informed college professors with a simulation of DNA sequencing that can be easily depicted with four colors of 3M Post-it® notes. The article will describe how DNA sequencing is carried out in a wet lab and will provide a dry lab simulation of the critical steps of the process. Upon completing the "dry lab", students will be able to determine the nucleotide sequence of genes, confirm the names of those genes and, if desired, use this as a starting point for generating a literature review based paper on the role these genes play in diverse organisms. This novel technique was presented to secondary science teachers (during a 2 hour session) from around the state of Nebraska in a workshop for Rural Academic Secondary School Science Partnership (RASSSP). It was found very useful by all participants and comments by these teachers are included at the end of the article.

Materials

(listed per student or group)

- Approximately 200 sheets (73 mm x 73 mm) each of four Post-it® note colors. In this example yellow, pink, orange and green sheets were utilized.

INSTRUCTOR PREPARATION (Preparation time per student or group is approximately 20 minutes).

STEP 1 – Simulation of an "unknown" single-stranded gene sequence.

Obtain a sequence from The National Center for Biotechnology Information (NCBI) website (see instructions under IDENTIFICATION OF YOUR GENE FRAGMENT) or utilize one or more of the gene sequence fragments provided (see appendix A). It is recommended that if choosing a gene sequence from NCBI on your own, that you utilize bacterial or viral genes to eliminate confusion that may result from introns present within eukaryotic genes.

Representing the sequence with Post-it® notes is carried out by assigning a colored Post-it® note to each of the four types of nucleotides. The following code is suggested:

Yellow sheets represent the base adenine.
Pink sheets represent the base guanine.
Green sheets represent the base cytosine.
Orange sheets represent the base thymine.

For example, a gene fragment with a known sequence of GATCGTCAC would be represented with Post-it® notes that were stuck together in the following sequence: pink-yellow-orange-green-pink-orange-green-yellow-green. The sequence of Post-it® notes will serve as "template DNA" in our sequencing reaction (described below) and should be about 25 sheets (nucleotides) in length. Shorter sequences may result in less specificity when gene identification is carried out.

STEP 2: Generation of primers.

For each gene sequence generated in step 1 (above) you will need to generate approximately ten "primers" that are each the same four nucleotides in length. A primer is a short segment of nucleic acid

that will hybridize (form hydrogen bonds between bases of nucleotides) to complementary DNA at a designated location and provide a chemical structure (3'-OH) needed for further extension of a growing DNA strand from that specific location on the DNA template strand. Using our sample sequence above of GATCGTCAC this primer set of Post-it® notes would complement the first four bases of GATC. A single strand of DNA can only hybridize to a second strand of DNA in a guanine/cytosine or adenine/thymine manner (Weaver, 2008). So the complementary strand of GATC would be CTAG and would be represented with approximately ten Post-it® note sets of a green-orange-yellow-pink sequence, all to be given a black border (Figure 1) to designate these short sequences as primers that provide the necessary starting point for extending DNA from that specific "primed" location.

Figure 1. Sequence elongation. Shown are both the template DNA stand (bottom) and a newly generated DNA strand (top) that could be produced in reaction Tube A. Note that the primer (the first four nucleotides of the top strand) are outlined in black and a newly incorporated deoxynucleotide (fifth sheet shown in green) and terminal dideoxynucleotide (yellow crossed-out sheet) are added in a fashion that complements the template strand in a green-pink (cytosine-guanine) and orange-yellow (thymine-adenine) manner.

STEP 3: Simulation of components needed for Sanger Sequencing.

Wet lab sequencing requires a polymerase enzyme, primers, template DNA and both deoxynucleotides and dideoxynucleotides. A description of the role of each of these components can be obtained from various sources (Sanger 1977, Dolan). Each of these major components of the "wet" sequencing reaction needs to be represented in the "dry" Post-it® note reaction. The role of the polymerase will be carried out by the students. Template DNA and primers were prepared in steps 1 and 2 above. Individual deoxynucleotides will be represented by Post-it® notes of various colors and dideoxynucleotides should also be represented with Post-it® note sheets. However, the sheets representing dideoxynucleotides should have an X drawn on them. About $1/5^{th}$ of your Post-it® notes should be crossed-out to represent dideoxynucleotides. Sequencing in a "wet" lab also requires a buffer which can be represented by the air in the room. The level of detail describing each of these components will, of course, be dictated by the individuals you are teaching.

Assemble four table tops for each sequencing reaction. Each table will represent one of four "reaction tubes" for DNA sequencing which we will call Tube A, Tube C, Tube G and Tube T. Each "reaction tube" (table top) includes all the following reagents:

1. A simulated Post-it® sequence (DNA template approximately 25 nucleotides in length) derived from appendix A or other source. Note that this sequence must be the exact same sequence on each of the four tables.

2. Approximately ten complementary Post-it® note primer sequences (four nucleotides in length). Again, these should be exactly the same primer sequence on each of the four tables.

3. Deoxynucleotides represented by all four Post-it® note colors. Since each of the four deoxynucleotides will be utilized multiple times, approximately 200 sheets of each color of Post-it® note should be made available on all four tables.

4. Dideoxynucleotides represented by crossed-out Post-it® notes for one of the four dideoxynucleotides, EITHER A-adenine (yellow), G-guanine (pink), C-cytosine (green) or T-thymine (orange). Each of the four tables will have a different color of crossed-out Post-it® notes.

A student or group will work with one of the four tables or "reaction tubes" to identify an unknown sequence. If more sequences are to be deciphered, any number of groups can be set up as long as each group utilizes a set of four table tops.

Instructor preparation should have resulted in a simulation of sequencing reaction tubes. Each of the four tables will represent 1 of 4 possible reactions in the sequencing process. The primer (with the help of a student playing the role of polymerase) will start matching the DNA template with appropriate deoxynucleotides and 1 of the 4 dideoxynucleotides. The first table represents the reaction in Tube A which includes the following "dry" reagents:

1. Template DNA (Post-it® note sequence prepared in step 1 of Instructor Preparation).

2. Deoxynucleotides of all four bases (Approximately 200 sheets of each of the four colored Post-it® notes).
3. Dideoxynucleotides of adenine only (Approximately 40 yellow Post-it® notes with an X drawn on them).
4. Primers (approximately ten).
5. Polymerase (A student will play the role of this enzyme).

A second table will act as "Tube C" and should contain the exact same components as "Tube A" with one exception: dideoxynucleotides of cytosine (green crossed-out Post-it® notes) should replace the adenine sheets. "Tube G" and "Tube T" should also be represented at individual table tops, in each case replacing Post-it® notes that represent the dideoxynucleotide adenine with guanine and thymine, respectively.

STUDENT ACTIVITY

Simulation of the sequencing reaction (using Tube A as an example):

DNA sequencing relies on the generation of many single-stranded DNA products that are the result of copying a single-stranded DNA template. Each step of sequencing will be described below, followed by a simulation procedure to be used in our dry lab. For a more detailed account of sequencing, see the Dolan animation provided in the reference section. In each reaction tube, the entire sequencing process is actually a series of reactions described below:

Reaction 1 (Denaturation). Wet lab sequencing is initiated by heating the components in each of the four reaction tubes (A, C, G and T) to 94°C. This will denature our separate double-stranded DNA into single-stranded template. In our "dry" reaction, the DNA is already "denatured" and represented as a single-stranded template. The dry lab could be easily modified to include a double-stranded DNA molecule by attaching complement Post-it® notes to the single-stranded template assembled by the instructor during simulation preparation. Simulation of heating is achieved by breaking the hydrogen bonds between guanine/cytosine and adenine/thymine bases. This would be simulated by separating the linear Post-it® note sequences from one another (separating pink from green and yellow from orange). The end result is a single-stranded DNA template.

Reaction 2 (Annealing). In a wet lab, the reaction tubes are cooled (approximately 55 °C) to facilitate the binding of primers to the single-stranded template DNA. The role of a primer (normally a 10 to 20 nucleotide single-stranded sequence) is to anneal (form complementary double stranded DNA by hydrogen bonding between guanine/cytosine and adenine/thymine bases) to a specific complementary region of DNA you wish to sequence. The primer provides a necessary chemical structure (known as a 3'-OH group) needed for the addition of nucleotides to a growing DNA strand (Stefan, 1998).

The "dry" lab simulation of this step is demonstrated with primers that are merely 4 nucleotides in length. The students should anneal a primer to the region of DNA that is complementary at all four bases. This should, if set up properly in the instructor preparation portion of the lab; result in priming at a terminal end of your single-stranded DNA template. This primer will provide a necessary starting point for the addition of more nucleotides that complement the template DNA strand during the elongation phase (immediately below).

Reaction 3 (Elongation). Wet lab elongation steps generally occur during a slight elevation of temperature (approximately 60-72°C). This allows a polymerase enzyme to attach an additional nucleotide to the primer that is complementary to the template strand. Hydrogen bonds are formed only between guanine and cytosine or adenine and thymine. As a result, if a template strand contains a thymine, only a nucleotide containing the base adenine can form the second strand of this DNA molecule. If the template strand consists of a guanine at the terminal end of the primer, a cytosine will be added to the primer. This will continue until a dideoxynucleotide is incorporated into the growing DNA strand. Since dideoxynucleotides do not contain a necessary 3'-OH group, no new nucleotides can be added to this chain and elongation stops.

In the dry lab simulation of the extension step, students play the role of the polymerase enzyme and Post-it® notes (standard and crossed-out, respectively) play the role of deoxynucleotides and dideoxynucleotides. The student, reading the primed template strand, progresses by attaching randomly chosen Post-it® notes in only a yellow to orange and green to pink fashion (see figure 1) from the priming site. When the student randomly places a crossed-out Post-it® note (representing a dideoxynucleotide) on the extending strand, that particular fragment of DNA is terminal and cannot be elongated further as dideoxynucleotides lack the necessary 3'-OH chemical structure needed for continued addition of DNA nucleotides.

Repeat reactions 1-3 (Repeat of Denaturation, Annealing and Elongation). At this point, the wet lab would repeat starting with the denaturation step that would separate the newly formed DNA strand from the template strand. Primers would again anneal to the template and a second elongation step would take place. This cycle

would repeat several times, generating a large population of newly formed DNA strands which terminate at various locations due to the random addition of a dideoxynucleotide. Simulation of this process for the dry lab consists of removing the newly generated Post-it® DNA strand, priming again and carrying out elongation as described above. This cycle should continue until all possible termination sites are represented in that particular tube.

Because the only terminators available in "Tube" A are adenine dideoxynucleotides, all termination sites in this particular tube will contain a terminal nucleotide of adenine. By choosing adenine (yellow Post-it® notes) randomly, it should become obvious that many of the sequences you are generating will be repeats, just as they would be in a "wet" sequencing reaction. However, eventually all possible sizes of fragments ending in adenine will be generated.

The same process is to be carried out in "Tubes" C, G and T. Note that each of these tubes, due to the presence of their respective dideoxynucleotides, will terminate with a cytosine, guanine or thymine, respectively. Again, all possible lengths of DNA with the proper termination nucleotide in each tube will eventually be generated. The number of nucleotides in each sequence of Post-it® notes should be counted and written on the terminal Post-it® note for use in the gel electrophoresis portion of the activity.

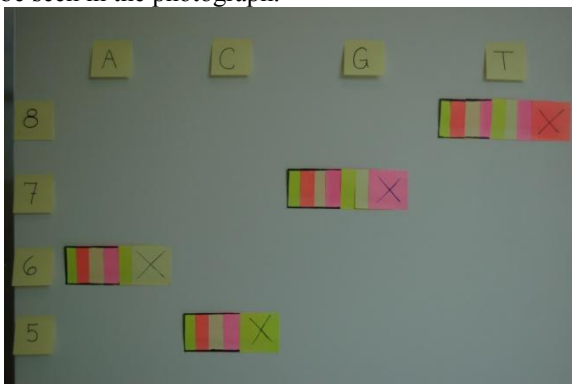
The result of the reactions in Tubes A, C, G and T are a population of newly synthesized DNA strands in the form of Post-it® notes that represent every possible length and terminal nucleotide of the original template strand. However, all the products of tubes A, C, G and T are still in separate reactions and need to be resolved through a simulation of acrylamide gel electrophoresis.

SIMULATION OF ACRAYLAMIDE ELECTROPHORESIS:

In a wet lab, the newly synthesized DNA fragments generated in Tube A, C, G and T are each loaded into a separate lane of an acrylamide gel. Once a current is applied, the newly formed DNA fragments migrate through the gel and are separated based on size, with the smallest fragments traveling through faster then larger fragments (Dolan).

To simulate electrophoresis, utilize a section of floor or wall space. Linear tile patterns on a floor work particularly well. Label a column of space numerically 25 (at the top) to 1 (see figure 2). Next to the numerically labeled column, label a new column as Tube A. Add columns for Tubes C, G and T as well. These will act as your acrylamide gel lanes. The samples need to be loaded in the gel and separated based on size. This is simulated simply by a student holding generated Post-it® note DNA fragments from Tube A. As the student walks down the gel (s)he looks for matching sizes of fragments with the numbers labeled on the gel. As the student in Column A moves down the gel from 25 nucleotides to 1 nucleotide, (s)he should lay corresponding length fragments at the appropriate location in the gel. For example, if (s)he does not posses any Post-it® notes fragments of 25 nucleotides in length, (s)he steps forward to position length 24. If (s)he does possess Post-it® note DNA fragments at this length, (s)he lays them on the floor at this site and continues moving down column A through length 23, 22, 21……..1. This same thing should be done in column C, utilizing only Tube C products. Additionally columns G and T should follow with Tube G and T products, respectively.

Figure 2: Simulation of Gel Electrophoresis. The numbers on the left represent the number of nucleotides (including primers) in the newly generated Post-it® note DNA fragments. Columns from reaction Tubes A, C, G and T are labeled at the top of the photo. Post-it® note fragments are pooled from Tubes A, C, G and T reactions and run in a single gel (albeit separate lanes) to generate the pattern revealed. By reading nucleotide lanes from the bottom to the top, a sequence can be obtained. In this case the sequence is C-A-G-T. Note: Only four nucleotide fragments are shown to ensure detail can be seen in the photograph.



READING THE SEQUENCE

In a wet sequencing lab, DNA strands separated on the acrylamide gel are not visible unless they are tagged with either radioactive or fluorescent markers. For more details on methods of detection please see the following reference (Dolan). In our example, the fragments are visible. Students should have generated a band of DNA five nucleotides long (four nucleotides representing the primer, plus an additional dideoxynuclotide) in only a single reaction tube (A, C, G or T). Similarly, each reaction tube

should have only specific fragment lengths because of the deoxynucleotide randomly complemented by the specific dideoxy variant of that reaction tube. As you look at the Post-it® notes on the floor/wall (in your gel), you will simply be able to see which lane (column A, C, G or T) contains the smallest newly generated DNA fragments (five Post-it® notes in size). If this happens to be in column G for your particular sample, then your first base in the unknown sequence is marked as a G. Look for the next largest and identify it (A, C, G or T) based on the column in which it is present (See figure 2). Continue the process until you reach the top of your gel. You have just generated a sequence of your unknown gene. The next step will be to figure out what protein results from the DNA sequence you just deciphered.

IDENTIFICATION OF YOUR GENE FRAGMENT.

This part of the procedure is the same whether you determined your DNA sequence in a "wet" or "dry" lab. The sequence you have deciphered from your "dry" lab can be identified utilizing the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov/). On the home page click on BLAST and then nucleotide blast. Type your sequence (nucleotides 1-25) into the prompt area and submit your query with your "search set" marked as "other" (as opposed to either the human or mouse genome). The best matches for your sequence will be displayed. By clicking on link sites, the name of the gene and organism in which it is housed can be readily determined. The instructor can then ask for detailed reports based on researching the role this gene plays within the organism in which they are housed. The depth and quality of these reports can be modified to fit your particular audience.

RASSSP Review Comments:

"The hands on activities were great. I can take most of the activities and implement them directly into my curriculum."

"The lab Dr. Christensen did with us was great. I wish we could have that lab written down for us because I'd like to do it but I'm not sure I could repeat it right from memory."

"The activity that Dr. Christensen provided was great!!"

"Great stuff! Just what I wanted, a great, hands-on low technology lab that I can do w/ students. We want more of this type of low technology."

"Doug did a very good job of using a do-able activity to teach a difficult concept!"

"I can tell my students about the technology and answer "How do they do that" questions."

"Excellent prep and decent activities."

"Wonderful presentations by Christensen. I appreciated how we were able to do the activities. … provided good background information for the content/activities they delivered."

"Very good for biology teachers. Nice activity! Extremely well done."

"Well presented and usable at our level."

"This gave me a much better understanding of DNA sequencing."

"This was one of my favorite sessions. It was new and very usable information."

Acknowledgement

Thanks to Drs. Barb Engebretsen and Shawn Pearcy for review aid during the writing of this article. Partial salary funding during write-up was provided by The National Institute of Health grant #P20 RR016469.

References

DOLAN DNA LEARNING CENTER. ©Copyright, Cold Spring Harbor Laboratory. Available online at: http://www.dnalc.org/ddnalc/resources/sangerseq.html

SANGER, F., NICKLEN, S, AND COULSON, A.R, 1977. DNA sequencing with chain-terminating inhibitors. Proceedings from the National Academy of Science, 74(12), 5463-5467.

STEFAN, H., VINGRON, M., POUSTKA, A., AND WIEMANN, S. 1998. Primer design for large scale sequencing. Nucleic Acids Research, 26(12), 3006-3012.

WEAVER, R. 2008. *Molecular Biology* (4th). New York, NY: McGraw-Hill.

APPENDIX A:

Bacterial/Viral Sequences

1. gcgaaagacg atgcggcagg tcaggcgatt (fliC Esherichia)
2. tatatagtga gaaaaaacga tatgtatggt (Internalin A Listeria)
3. aaagtgagaa atacgaacaa agcagaccta (actA Listeria)
4. caatgggaac tctgccggga ttcccactgc (invA Salmonella)
5. agtgccggag tttgacatcg acgatgaggt (rpoS Pseudomonas)
6. aatagcatgt aagcaaaatg ttagcagcct (OspA Borrelia)
7. aaaaaagcag aagaaaacaa acaaaaaggc (mip Legionella)
8. gttggacccg aagaccaggt ccacgcggct (katG Mycobacterium)
9. aaagcaggtc agaaaacgga tgatatgctt (slo Streptococcus)
10. tcacatcaat gacagtaatt tctgcatctg (yop Yersinia)
11. gagcgggagg tgcgacatat acatatagaa (gag HIV)