Module 4: Bioinformatic and Molecular Biological Identification of Mutations in DNA

# Background

Geneticists have identified numerous mutations that cause or contribute to disease. We now know the precise genetic mutations responsible for several forms of muscular dystrophy, anemia, cystic fibrosis, Huntington’s chorea, and other inherited genetic disorders. For some diseases, we even know how the mutation affects the structure and normal function of the encoded proteins. But how were these mutations identified? Investigators use multiple methods to identify general regions in the genome that may differ between normal and disease-prone individuals. Once a candidate region has been identified, it is isolated and sequenced using high-speed automated instruments. Next the nucleotide sequence of the DNA region in disease prone individuals is compared to an equivalent region in normal individuals. The sequences for normal individuals are downloaded from a massive computer database containing complete genome sequences for dozens of organisms, including humans. Then the two sequences are aligned using bioinformatics software, allowing the investigator to identify changes in the DNA that may be involved in the disease process.

In this unit we will use bioinformatics software to find differences between normal and mu­tant DNA sequences, then use DNA restriction mapping to confirm the differences you find. So that the lab procedures can be performed in a reasonable amount of time, we will work with bacterial rather than human DNA. However, the meth­ods and thinking process are much the same as for characterizing genes linked to human disease.

## The Lac operon in *E. coli* as a model system

Free-living bacteria like *Escherichia coli* use many carbon sources for energy and raw materials. Their preferred sources are glucose, lactose, maltose, or other sug­ars. To catabolize sugars, bacteria must transcribe genomic DNA into mRNA, and then translate it to produce the necessary enzymes. *E. coli* does not continuously syn­thesize every enzyme and transporter protein needed to metabolize every possible sugar. Instead, it transcribes mRNAs for the required enzymes only when that par­ticular sugar is present in the local environment. This process can be demonstrated in lab using the*lac* operon.



Figure 1: General features of the Lac operon. The Lac operon consists of the CAP (cAMP binding protein) binding site, the -35 and -10 sites where transcription factors bind, the Operator region where an inhibitory LacI protein can bind, the start site where RNA polymerase begins transcription, and the protein coding region that starts with the initial methionine of the β-galactosidase (βGal) protein.

Most bacterial promoters contain -35 and -10 sites where general transcription factors bind then recruit RNA polymerase for transcription. Upstream of the -35 site in the lac operon is a binding site for an activator protein called cAMP activa­tor protein (CAP). Immediately downstream of the -10 site is an operator region, where a repressor protein called LacI can bind and block transcription. Down­stream of the operator is where RNA transcription starts. Just downstream of that is the “ATG” codon for the first methionine in the LacZ gene that codes for β-galactosidase (βGal), an enzyme that breaks down lactose (a disaccharide) into galactose and glucose.

Figure 2: Nucleotide sequence of the promoter region of the lac operon. The sequence shown begins just upstream of the CAP binding site, and ends several codons into the coding sequence for the LacZ gene (β-galactosidase).

## Regulation of the Lac operon

There are two major regulatory elements—one repressive and one activating—that control the rate of transcription of the coding region of the Lac operon. The Lac I gene within the operon codes for a repressor protein. This protein has a binding site for lactose. The first regulatory element is the ***LacI repressor protein*** which has a binding site for lactose. When lactose is absent, the repressor protein will bind to the Operator region of the lac operon, preventing RNA polymerase from binding to the Lac operon and thus preventing transcription. When lactose is present, lactose binds to repressor protein and induces dissociation of repressor protein from the Operator region, resulting in increased rates of transcription. The second regulatory element is the ***CAP*** which has a binding site for cAMP. When cAMP levels are low, the CAP lacks bound cAMP and CAP fails to bind to its binding site at the beginning of the lac operon. When cAMP levels become elevated, cAMP binds to CAP, and then CAP can bind tightly to its DNA recognition site, activating transcription.

Figure 3 shows four different scenarios in which *E. coli* are growing in an environment with or without lactose and glucose.



Figure 3: Control of Lac operon depends on carbon source. Rapid production of mRNA coding for the βGal enzyme (and other downstream gene products) only occurs in the absence of glucose (Glu) and the presence of lactose (Lac).

* Scenario #1—lactose and glucose both present: lactose will bind the LacI repressor protein, relieving its inhibition of transcription. However, glucose is the preferred source of sugar, and because it utilized to make energy, the level of cAMP (a “hunger” signal in bacteria) is low, and CAP lacks bound cAMP. Without CAP binding to the lac operon, transcription is not activated. End result: transcription of mRNA prevented.
* Scenario #2—glucose present but no lactose: repressor protein will bind to the Operator, repressing transcription. High glucose means low cAMP, and CAP fails to bind to the Lac operon. End result: transcription of mRNA prevented.
* Scenario #3—no sugars present: repressor protein binds Operator, repressing transcription. cAMP levels are high in absence of glucose, so CAP binds to DNA but cannot overcome LacI repression. End result: transcription of mRNA prevented.
* Scenario #4—lactose present, glucose absent: repressor protein binds lactose, dissociates from Operator site, and de-represses transcription. In absence of glucose, cAMP levels are elevated, CAP binds cAMP and can associate with DNA binding site on Lac operon, activating transcription. End result: combination of de-repression and activation lead to mRNA transcription of the LacZ gene (coding for the βGal enzyme) as well as other downstream gene products of the Lac operon.

As you can see, the presence of lactose is necessary—but not sufficient—to turn on the Lac operon. *E. coli* must not only have lactose (hence requiring the enzymes needed for its breakdown), but also must be lacking glucose (the preferred carbon source) so that CAP (with cAMP bound) can activate transcription upon binding to the promoter region of the Lac operon. For an animated presentation of the regulation of the lac operon by the repressor protein, see the video from the NDSU Virtual Cell Project at <http://www.youtube.com/watch?v=oBwtxdI1zvk> (note that this video does not include mention of the positive regulation by CAP, but does describe the regulated expression of lactose permease, a transporter that enhances lactose uptake into the bacterial cell).

## Mutations can affect the normal regulation of the Lac operon and the ability to express protein

There are many different ways in which mutations in the DNA sequence can occur. Some mutations are “silent” and cause no change in the expressed protein’s structure and/or function. Other mutations causing a substitution of one nucleotide for another nucleotide can have profoundly deleterious consequences. Deletion and insertion mutations of one or more base-pairs can change the reading frame, create early stop codons, replace one amino acid for another, prevent binding of a transcription factor, etc. Movement of larger segments of DNA within the genome (“transposition”) can disrupt expression and regulation of gene products.

Consider what would happen if mutations were to occur in each of these elements of the Lac operon:

* CAP binding site
* Promoter (-35 or -10 sites)
* Operator (binding site for the LacI protein)
* Transcriptional start site
* Coding region for the βGAL protein itself

## Mutational analysis is often carried out using bacterial plasmids

Isolating and manipulating genomic DNA directly from any organism (even bac­teria) is very time-consuming. For that reason, molecular biologists routinely take DNAs of interest and insert them into small carriers called **plasmids** (Fig. 4). Plasmids are naturally occurring double-stranded, closed circular DNA molecules found in many bacteria that have been exploited by molecular biologists as a tool for genetic analyses and protein expression.

Most plasmids contain a gene coding for an enzyme that makes the bacteria carrying that plasmid resistant to a particular antibiotic (e.g., ampicillin; Amp.). While each bacterial cell has only one copy of its genomic DNA, it may contain many identical copies of a small plasmid. Plasmid replication starts at a DNA sequence known as the origin of replication (Ori.).

Plasmids are essential laboratory tools as well. Under controlled conditions, plasmids can be used to incorporate fragments of genomic DNA into their multiple cloning sites (MCS). The plasmids carrying individual genes can be multiplied, purified, ma­nipulated, and analyzed in a variety of ways.

Figure 4: pBluescript plasmid showing the ampicillin resistance gene (Amp), the origin of replication of the plasmid (Ori), and the multiple cloning site (MCS) into which genes are inserted into the plasmid.

## Using bioinformatic and molecular biological tools to identify mutations in the Lac operon.

For this module, you will be working with a strain of *E. coli* whose genomic DNA lacks the native Lac operon. We have inserted into this *E. coli* strain a plasmid containing either the wildtype (WT) Lac operon DNA or a plasmid containing a mutated version of the Lac operon DNA. We have grown up cells containing each plasmid encoding a WT or mutant Lac operon, purified the plasmid DNA for you, and performed DNA sequencing of the WT and mutant plasmid DNA sequence in the region of the Lac operon.

The pBluescript plasmid (Fig. 4) is the backbone into which the Lac operon DNA was inserted. We designate the plasmid containing WT DNA sequence as “pLac/WT” (Fig. 5). The mutated Lac operons are numbered “pLac/m1,” “pLac/m2,” etc.

Figure 5: pLac/WT plasmid. The MCS region of pBS was used to insert the wild-type lac operon sequence into the plasmid backbone to create the pLac/WT plasmid.

The first approach you will take is to identify changes in the DNA sequence of the mutant lac operon compared to the WT lac operon. This is accomplished by aligning the sequences of the two DNAs using Muscle within a terminal (a BASH environment). Once you identify the mutation, you can identify its location within the Lac operon to characterize what consequences the DNA mutation might have. Mutations could have occurred in any of the following parts of the operon: CAP binding site, the -35 or -10 promoter regions, the Operator region, the transcriptional start site, or somewhere in the coding sequence for the LacZ gene itself. To simplify your analysis, we will provide only the DNA sequence for the Lac operon, not the entire plasmid DNA sequence.

The second phase of the bioinformatic work is to try to confirm the sequence and location of the mutation using restriction enzymes that recognize unique DNA sequences. You will first have to create a plasmid map showing all restriction enzyme cutting sites and choose an enzyme that will be diagnostic of your suspected mutation. Then in the following week you will digest a sample of plasmid DNA with this enzyme and analyze the size of the DNA fragments that you observe following agarose gel electrophoresis to attempt to confirm the mutation you identified.

## Restriction endonucleases are enzymes that recognize and cut very specific DNA sequences, and are important tools in molecular genetics research

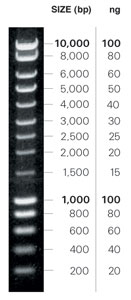
Restriction endonucleases are enzymes that produce cleavage of both strands of double-stranded DNA. These enzymes catalyze hydrolysis of the phosphodiester bond in the middle of a specific sequence of nucleotides. The utility of these “molecular scissors” lies in the specificity of their cutting action. Many restriction enzymes are absolutely specific for a unique sequence of nucleotide bases. For example, BamHI is a restriction endonuclease that cleaves between the first and second guanine (G) of a DNA molecule having the following sequence: 5’-GGATCC-3’. The same cleavage is made on the complementary DNA strand: 3’-CCTAGG-5’.

## DNA molecules can be separated based on size using agarose gel electrophoresis

Agarose is a polysaccharide that is widely used for the separation of DNA molecules. The mobility of DNA on agarose gels depends on the size and shape of the nucleic acid. For linear DNA molecules, the velocity of migration in an electric field is such that larger, linear DNA molecules move more slowly than smaller, linear DNA molecules. [It is standard practice in molecular biology to refer to the size of a DNA molecule in terms of the number of base-pairs (bp), rather than its molecular weight.]

DNA can exist in other conformations besides the linear double-stranded conformation. As mentioned earlier, plasmid DNA normally exists in the form of closed circular DNA which is tightly supercoiled. Because supercoiled plasmid DNA assumes a more compact conformation than linear DNA, it migrates faster on agarose gels than a linear DNA molecule with the same number of base-pairs.

In the second week of this lab module, agarose gel electrophoresis will be used to separate plasmid DNA that has been digested with various restriction endonucleases to determine whether your digestion of plasmid DNA was successful and whether wild-type and mutant DNA molecules can be distinguished. By running a standard containing DNA fragments of known sizes (“DNA ladder”), you can also estimate the size of your digested DNAs. The DNA in the gel is visualized by staining the gel with a dye that fluoresces when it intercalates into the DNA double helix.



DNA ladder

**Bioinformatics**

Bioinformatics is the computer-based analysis of biological information. In the age of “Big Data”, these skills are foundational and highly advantageous for advancing in your careers in research, teaching or industry. Bioinformatics is practiced in a terse environment called BASH (**B**ourne **A**gain **SH**ell) using server based open-source software. This environment requires some training and practice but is worth the effort. The skills you will learn during this module will give you access to the vast library of bioinformatics software. These skills are durable and apply to disciplines outside biology.

You will be using programs to analyze DNA sequences. One of the most important concepts in bioinformatics (or any computer-based analysis) is data structures. To use a computer to analyze information the information needs to be in a predictable format. For DNA sequences, the most important data structure is the FASTA file. These files communicate DNA sequences as one strand on the DNA written in the 5’ to 3’ orientation.

Often bioinformatics analysis involves computer programs that mimic molecular processes. In your exercise, you will use a program that works like a restriction enzyme to recognize a string of specific nucleotides and identify all of the cut sites in a DNA molecule.

# Week 1, Exercise #1: Identification of mutated DNA within the Lac operon using Bioinformatics Software

## Helpful notes about Bash

## Bash is case sensitive

* If the command is not working, you may not have appropriate spaces. There is always a space between the command and filenames, etc. In order to better visualize these spaces, spaces are denoted by an orange space ( ) in this protocol.
* Once you have entered a filename or command once, the subsequent entries can be auto filled by hitting the TAB key. This will ensure correct spelling and that the file exists.
* Commands that need to be typed (or copy/pasted) into the terminal are **bolded.**

## Start a Bash session

* Open your terminal window: for windows this is command prompt, for macs this is terminal.
* Type in ssh -l <YOUR UNH ID> ron.sr.unh.edu
  + Your UNH ID example: abc1234
  + -l is a lower case L
* Type in the password UNH ID + wMS7Jy
  + For example: abc1234 wMS7Jy
  + It is typing, even though nothing appears to be happening. This is a security feature in the terminal you are using
  + If you enter an incorrect password several times in a short time period, you will be locked out of the server. If this happens, let your instructor know as you will need your username and password reset.
* You will need to reset your password, if this is your first session.
  + It will ask you first to enter your current password again and hit enter.
  + Enter a new password that is at least 12 characters and hit enter.
  + Enter the new password a second time and hit enter. This closes the session. Reopen command prompt for windows, or terminal for Mac, if you wish to continue using it at this time.

**Create a new directory and copy files into a folder for use**

* Create a new directory called ‘exercise 1’ in your home directory (~) by entering the following command. You will only need to do this the first time you enter Bash.

**mkdir** **~/exercise1**

* Open the directory, using the following command.

**cd ~/exercise1/**

* Print your working directory (pwd) to ensure you’re in the correct location. You should see /home/UNHBIOL411/YourUserName/exercise1

**pwd**

* Copy the files within exercise 1, using the following two commands.

*\*Note – for # you will use your given mutant number*

**scp /home/share/bio411/Exercise-1/pLac-WT.fasta ./**

**scp /home/share/bio411/Exercise-1/pLac-m#.fasta ./**

* List (ls) the contents of your current directory to ensure that the files copied over correctly. You should see two FASTA files.

**ls**

* Concatenate the wildtype and mutant files into one file.

**cat pLac-WT.fasta pLac-m#.fasta > wt\_and\_mutant.fasta**

## Align the two sequences using muscle

* The first two alignment commands will not give you immediate displayed results, but the alignment will be used in an upcoming command. Don’t panic, if nothing is displayed.

**muscle3 -in wt\_and\_mutant.fasta -out wt\_mutant\_aligned.fasta**

**muscle3 -in wt\_and\_mutant.fasta -out wt\_mutant\_aligned.clw -clw**

* List the current directory to ensure the two alignment files were created.

**ls**

* You will view these alignments in two ways. The first shows you the nucleotide sequences in 60 base lines. Any places that the nucelotides are unchanged will be marked with a \*. You can visualize the mutations by looking for nucleotides with no \*. Record the location of the mutation and the mutation type.

**less wt\_mutant\_aligned.clw**

**\***Note – The sequence is longer than what you can see on the screen. In order to scroll down, use your down arrow key. When you are done viewing your sequence, press ‘**q**’ to quit and return to the command lines.

* + **TAKE A SCREENSHOT HERE OF YOUR ALIGNMENT BEFORE EXITING THIS SCREEN**
* Identify where in the Lac operon your mutation is located:
  + 105 – 116: CAP binding site
  + 140 – 145: Promoter (-35 site)
  + 164 – 169: Promoter (-10 site)
  + 172 – 192: Operator (binding site for the LacI protein)
  + 214-216 Initial methionine (ATG) of the coding region for the βGAL protein itself
  + 217+ : Protein coding region

## Summary Questions for Exercise 1

1. Is the promoter identical in both pLac/WT and your pLac/m# mutant? If not, which elements differ between them? If you found differences, how do you expect them to affect gene expression? What is your reasoning for these predictions?

No difference.

1. Look at the rest of the aligned DNA sequences for pLac/WT and your pLac/m# mutant. What other differences did you find? How do you predict these mutations will affect expression, function, or regulation of the LacZ protein? What is your reasoning for these predictions?   
   There is a deletion in the protein coding region. This removes about 200 amino acids and I would expect this to prevent the enzymes from functioning

# Week 1, Exercise #2: Create a restriction enzyme map of your wild-type and mutant DNA sequences to identify one or more restriction enzymes that could distinguish wild-type and mutant DNAs

In the previous exercise, you were provided with DNA sequences that consisted of the Lac operon portion of the plasmid DNA. For the purpose of creating a map of all of the sites where RE can cut the DNA, we need to work with the entire plasmid DNA sequence which includes what is called the plasmid “backbone” along with the inserted Lac operon DNA sequence.

* If you are not there already, open your terminal and relogin, connecting to the Host: **ron.sr.unh.edu**

**Create a new directory and copy files into a folder for use**

* Create a new directory by entering the following command. You will only need to do this the first time you enter Bash.

**mkdir ~/exercise2**

* Open the directory and copy the files within, using the following three commands.

**cd ~/exercise2/**

**scp /home/share/bio411/Exercise-2/pLac-WT\_FL.fasta ./**

**scp /home/share/bio411/Exercise-2/pLac-m#\_FL.fasta ./**

\*Note – for # you will use your given mutant number

* Concatenate the wildtype and mutant files into one file.

**cat pLac-WT\_FL.fasta pLac-m#\_FL.fasta > wt\_and\_mutant.fasta**

**Identify restriction enzyme cut locations and resulting fragment sizes**

Because you have already concatenated the files, you can run the command for restrictionFinder once for each restriction enzyme, and the output will give you cut sites for both your mutant and wildtype. Record the number of cuts, along with the position and size of the fragments, in the chart on page 11.

The program requires a FASTA file containing one or more sequences, and an enzyme. You can specify the enzyme by name. Another option you will want to set is “--circular True” this tells the program you are dealing with a circular plasmid. You will need to run this one-line command multiple times, specifying different enzymes each iteration.

\*Note: After running the program once you can use the up arrow on your keyboard to repeat the command. Move your cursor to the enzyme by using the left and right arrows, and type in the name of the next enzyme.

**restrictionFinder.py --fasta wt\_and\_mutant.fasta --enzyme BamHI --circular True**

\*Note – Enzyme names are followed by Roman numerals.

## Create a Restriction Site Mini-Map

After you have used restrictionFinder to explore where palindrome cut sites for specific restriction enzymes can be found in your DNA, you must decide which enzymes will be useful and which ones you can ignore. When pLac/WT and pLac/m# DNA are cut with the enzyme, the differences between the fragments released can be seen on an agarose gel. The gels you will use can accurately separate fragments from ~200 bp up to ~4000 bp long, while the best range for resolving multiple bands is from 600 bp to 3000 bp. Smaller fragments may run completely off a gel or be too faint to see when stained, while larger fragments may not separate into distinct bands. Generally, you can see differences between two DNA bands if one is at least 10% smaller than the other.

In order to visualize the cut sites, it is helpful to generate a map of your plasmid. Start by drawing your circular plasmid. Add the appropriate number of cuts generated, based on the results of the restrictionFinder. Space these cuts according to the length of the fragments created. Record the maps for relevant enzymes (those that result in different cuts on your mutant vs. the wildtype).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Restriction  Enzyme | Wild-type: position of  site(s) & size of fragments | # cuts | Map | mutant: Position of site(s)  & size of fragments | # cuts | Map |
| AseI | Position: 742, 1977, 2036 | 3 |  | Position: 742, 1977, 2036 | 3 |  |
| Size: 4757, 1235, 59 | Size: 4757, 1235, 59 |
| BamHI | Position: 2270 | 1 |  | Position: 2270 | 1 |  |
| Size: 6051 | Size: 5427 |
| BsrGI | Position: NA | 0 |  | Position: NA | 0 |  |
| Size: 6051 | Size: 5427 |
| ClaI | Position: 2234, 3144 | 2 |  | Position: 2234 | 1 |  |
| Size: 5141, 910 | Size: 5427 |
| EcoRV | Position: 2246, 3432 | 2 |  | Position: 2246, 2808 | 2 |  |
| Size: 4865, 1186 | Size: 4865, 562 |
| HindIII | Position: 2240 | 1 |  | Position: 2240 | 1 |  |
| Size: 6051 | Size: 5427 |
| NcoI | Position: NA | 0 |  | Position: NA | 0 |  |
| Size: 6051 | Size: 5427 |
| HpaI | Position: 2745, 3369 | 2 |  | Position: 2745 | 1 |  |
| Size: 5427, 624 | Size: 5427 |

* Create a ‘map’ of your plasmid for enzymes that differ between the wildtype and mutant. This is a drawing of the plasmid with all of your cut sites and fragment sizes. Remember:
  + A single cut will linearize the circular DNA and generate a linear DNA whose size equals the total base-pairs of the full-length plasmid.
  + If an enzyme cuts in two locations on the plasmid, the first cut will linearize the DNA, and the second cut will generate two fragments.

## Choosing restriction enzymes that are diagnostic for your lac operon mutation

* Plasmid DNA is double-stranded closed circular DNA that exists in a supercoiled state. If the plasmid is not cut by the restriction enzymes you use, it will migrate on an agarose gel faster than a linear, non-coiled DNA with the same number of total base pairs. Whenever possible, it is a good strategy to make sure each plasmid you test will be cut (and thus linearized) by the RE you use.
* The agarose gels you will be using have a size range of ~200 to ~6000 bp, and can separate fragments whose size differs by ~10% from its nearest neighbor.
  + The best range for resolving multiple bands is from 600 bp to 3000 bp.

## Summary Questions for Exercise 2

In addition to the questions below, you need to turn in your maps show­ing restriction enzyme sites for each of your two plasmids.

1. Based on the maps you generated, which enzymes can you use to demonstrate a difference in the sequences of pLac/WT and your pLac/m# mutant? For each enzyme, what will be the differences in sizes and numbers of fragments produced by digesting pLac/WT versus pLac/m#?

ClaI because it cuts the mutated dna a different number of times and in large enough fragments.

1. Where are the restriction enzyme cut sites you have chosen located in the operon fragment? In other words, are they in the promoter region, coding region for β-galactosidase, or the plasmid backbone? How do you know?

# Week #2, Exercise #3: Restriction endonuclease digestion and agarose gel separation of DNA fragments to confirm the presence of mutations in the lac operon.

In the previous week, you first compared the DNA sequence of the “wildtype” lac operon with that of a DNA sequence that contained a certain type of mutation in the lac operon. You then compared the restriction enzyme map of the WT and mutant plasmids to identify restriction enzymes that could potentially verify the mutation that you discovered. This week you will actually expose your WT and mutant plasmid DNA to these restriction enzymes, and then run agarose gel electrophoresis to separate the DNA that you digested. Due to time constraints, the final step of this procedure, photographing the gel, will be done for you after the end of this lab session, so that you can analyze the results of your diagnostic test.

## General guidelines for molecular biology experiments

### Pipetting

* The ingredients for your restriction digestion will be added in very small volumes (total reaction volume = 20 μl).
* Practice pipetting the smallest volume you will need for your reaction using a P20 pipette. After drawing up the desired volume, visually inspect the pipette tip and be sure that you can see the liquid inside. A 1.0 μl volume is easily visible.
* Transfer these small volumes to the very bottom of the microtube. Once one solution has been added, directly pipette subsequent ingredients into the drop of liquid at the bottom. Be sure none of the liquid remains inside the pipette tip after dispensing.
* NEVER re-use a pipette tip when doing molecular biology work. You will contaminate your solutions and your results will be compromised. Pipette tips are single-use only.

## Restriction enzyme digestion of WT and mutant plasmid DNA

### Reagents needed for setting up your digestion reactions

Each lab group will be given a set of tubes containing the following components:

* **W:** Sterile water
* **WT:** WT plasmid DNA(concentration = 20 ng/μl)
* **M?**: mutant plasmid DNA (concentration = 20 ng/μl)
* **B:** Reaction buffer (5x concentration)
* Note that the restriction enzymes will be kept in a central location in the lab, and should be left in their cold block except at the moment of pipetting.

*Also included but to be used later…*

* **GLD:** Gel Loading Dye (for preparing each of your DNA samples for electrophoresis)
* **H:** HyperLadder I (DNA ladder to estimate size and amount of DNA on gel)

### Controls and standards needed for your experiment

Having the proper controls is essential to being able to interpret the outcome of your experiment. For example, it is a good idea to prepare samples of WT and mutant DNA that have not been exposed to restriction enzymes (negative control), as well as samples that have been incubated with a restriction enzyme that will generate a single cut in the plasmid DNA (thereby converting it from closed circular to linear) regardless of whether you use WT or your mutant DNA (positive control). Finally, you will want to have one lane of your gel contain a DNA ladder so that you can estimate the size of your linearized DNA samples compared to the bands of the ladder.

### Restriction enzyme digestion reactions

* Set up your reactions in sterile 0.5 ml microcentrifuge tubes and label each tube A, B, C, etc. with a Sharpie.
* Add the ingredients in the following order (refer to chart below for volumes):
  + Sterile water
  + Plasmid DNA
  + Reaction Buffer
  + Restriction enzyme
* ***Change tips for each addition, even if you are pipetting the same solution. Be sure that you can see the volumes you are pipetting. Visual verification of having transferred the liquid from one tube to another is the single most important factor in obtaining good results in this lab. Your TA will be checking the volume of your reactions, so make sure to pipette carefully.***
* Once the additions are complete, mix the contents by drawing the entire volume up and down into a P20 pipette several times.
* Recover the entire volume with a quick, low-speed spin to force all the liquid into the bottom of the tube.
* **Incubate the tubes at 37 C for 45-60 min**.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Reagent** | **Tube A** | **Tube B** | **Tube C** | **Tube D** | **Tube E** | **Tube F** |
| **Water** | 6 uL | 6 uL | 4 uL | 4 uL | 4 uL | 4 uL |
| **Plasmid**  **DNA** | 10 uL  **WT** | 10uL  **Mutant** | 10 uL  **WT** | 10 uL  **Mutant** | 10uL  **WT** | 10 uL  **Mutant** |
| **Rxn Buffer** | 4 uL | 4 uL | 4 uL | 4 uL | 4 uL | 4 uL |
| **Res. Enzyme** | 0 | 0 | 2 uL  Hind III | 2 uL  Hind III | 2 uL  ? enzyme | 2 uL  ? enzyme |
|  | *negative control* | *negative control* | *positive control* | *positive control* | *Experimental* | *Experimental* |

* Once you have incubated your samples, add 5 μl of 5x gel loading dye to each tube to stop the reaction, and to prepare the samples for electrophoresis. The gel loading dye contains glycerol which makes the sample heavy so it sinks into the well and also contains dye to track the progress of electrophoresis.
* Mix well with a micropipette, and then briefly spin tubes to collect liquid at bottom.

## Agarose gel electrophoresis

Each group will have a gel box that has 10 lanes per agarose gel. Therefore, you have a maximum of 10 samples that can be run (actually, 9 after using one lane for the DNA ladder).

### Practice loading agarose gels

* A sample gel is available to practice loading 5 μl of a colored solution into the wells of the horizontal gel.
* Use a P20 pipette for loading samples.
* Practice your loading technique on this gel before you load your digested samples.

### Loading your samples on agarose gel

* Lane #1: 5 µl Hyperladder
* Lane #2: 20μl of Tube A
* Lane #3: 20 μl of Tube B
* Lane #4: 20 μl of Tube C
* etc….
* Record which samples are loaded into which lanes.
* This DNA ladder not only allows estimation of size (base-pairs), but the intensity of the bands provides a semi-quantitative estimate of the amount of DNA loaded (ng).

### Agarose gel electrophoresis

* Sample wells should be closest to the negative (black) electrode.
* After attaching electrodes, ask the TA to check your unit.
* Run at 120 V until the fastest-running dye band has reached the 4-5 cm mark.

### Visualize the results of your agarose gel (to be performed after the lab session ends by your TA)

* Remove the gel from the electrophoresis unit
* Properly dispose of the gel running buffer.
* Place the gel on a UV transilluminator.
  + ***Caution***: a special UV-blocking face shield must be worn to look at UV light. Standard safety glasses are *not* adequate protection. Avert your eyes from this light source if you are in its vicinity. Also protect your exposed skin from exposure.
* Take a picture of the gel with a digital camera. Upload the image to Blackboard.
* Properly dispose of the gel.

DNA Hyperladder

# Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Lab Section:\_\_\_\_\_\_\_\_\_\_\_\_**

# Predicted pattern of DNA fragments for your gel samples

* Sketch your predicted results for your agarose gel using the following guidelines
  + At the top of the sketch, label the sample that was loaded in each of the lanes (e.g., Tube A might be called “WT, no RE”).
  + For each lane, draw where you predict the position on the gel of each band you predict will be visible.
  + Repeat this for each of the sample lanes that you loaded.
  + On the other end of your sketch, indicate where you predict uncut WT plasmid DNA to be found.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Tube A | Tube B | Tube C | Tube D | Tube E | Tube F |  |  |
| Sample | |  |  |  |  |  |  |  |  |
| Description | |
| Hyperladder | 10kb |  |  |  |  |  |  |  |  |
| 8kb |  |  |  |  |  |  |  |  |
| 6kb |  |  |  |  |  |  |  |  |
| 5kb |  |  |  |  |  |  |  |  |
| 4kb |  |  |  |  |  |  |  |  |
| 3kb |  |  |  |  |  |  |  |  |
| 2.5kb |  |  |  |  |  |  |  |  |
| 2kb |  |  |  |  |  |  |  |  |
| 1.5kb |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| 1kb |  |  |  |  |  |  |  |  |
| 0.8kb |  |  |  |  |  |  |  |  |
| 0.6kb |  |  |  |  |  |  |  |  |
| 0.4kb |  |  |  |  |  |  |  |  |
| 0.2kb |  |  |  |  |  |  |  |  |

# Summary questions for Exercise #3

* Either print out a copy of your gel image or use an image editing program to label the gel image with the sample run in each lane.
* For the negative controls, what is the apparent size of the uncut WT and mutant plasmid? Does it agree with the number of base-pairs that you analyzed through bioinformatics? If not, why not? Explain if there are multiple bands or any unexpected aspects of the negative controls.
* For the positive controls, did the HindIII restriction enzyme cut both the WT and mutant plasmids? How can you know that the plasmid DNA was cut? Is the observed size (in bp) of the plasmid what you expected based on the sequence analysis? Why or why not? If multiple bands are observed or the sizes of the bands are unexpected, please explain.
* Did the enzyme that you chose as a diagnostic for the presence of a mutation in the DNA sequence of the mutant plasmid confirm that your mutant does indeed have the mutation you identified from the sequence alignment? Compare the bands observed in your experimental samples for WT and mutant plasmid, and identify and explain differences between the two different DNAs.