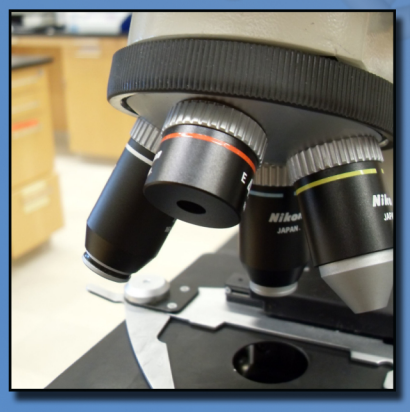
Investigations in Molecular Cell Biology

Author: Clare O'Connor

2014



Investigations in

Molecular Cell Biology

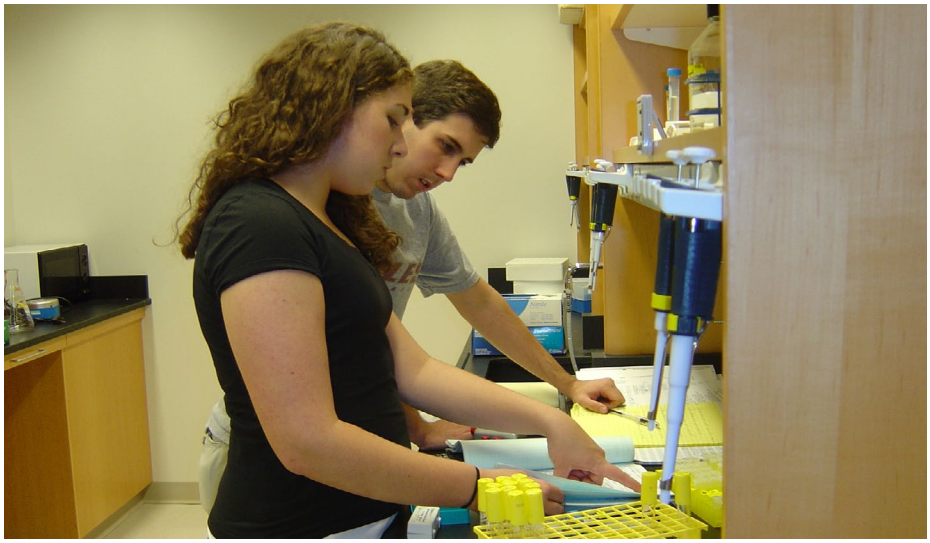
Boston College | Clare O'Connor

2014-2015

Contents

Chapter 1

Introduction



We live in the "post-genomic" era, in which the availability of complete genome sequences from a host of organisms offers exciting opportunities for undergraduate research. In this course, we will use some of the strain and clone resources generated by the *Saccharomyces* genome project to investigate the evolution of genes involved in the synthesis of methionine and cysteine, essential sulfur-containing amino acids.

Investigations in Molecular Cell Biology

- Course design and learning goals
- Pathways over Time: Our research project
- Course overview
- References
- Acknowledgments

Chapter 1

Welcome to Investigations in Molecular Cell Biology. a new kind of introductory lab course that incorporates an authentic research project in functional genomics. It has been said that we live in a "post-genomic" era. Large-scale genome projects have generated tremendous amounts of sequence data, and complete genome sequences are available for thousands of organisms. In a typical genome project, genes are identified largely by their sequence similarity to known genes from other organisms (Goffeau *et al.*, 1996), with the assumption that the proteins encoded by the genes perform similar functions. These "loose ends" connecting sequence and function provide exciting opportunities for undergraduate students to participate in functional genomics research. In this course, students will study the functional conservation of the genes involved in methionine (Met) and cysteine (Cys) biosynthesis. Met and Cys are essential amino acids in all living cells. These two amino acids contain sulfur in their side chains, which allows Met and Cys to play unique roles in proteins.

We expect that students will make novel findings in their projects each semester and that students will be able to build upon the results obtained in preceding semesters. We hope that you enjoy the research experience and we look forward to your experimental results!

Course design and learning goals

Biology education at the undergraduate level is undergoing a transformation. For decades, many have viewed biology as an encyclopedic subject, because of the vast amount of content matter included in the undergraduate curriculum. A recent reevaluation of undergraduate biology education, however, is guiding biology curricula in a new direction, stressing the importance of involving students in the process of scientific investigation in their coursework (Bauerle *et al.*, 2011). This reevaluation process has also challenged educators to sort through the vast amount of content in introductory biology to identify the core concepts that students should learn and the key competencies that students should acquire during their undergraduate education. This course has been designed in line with these recommendations.

Our course research project is designed to illustrate the core concepts of biology:

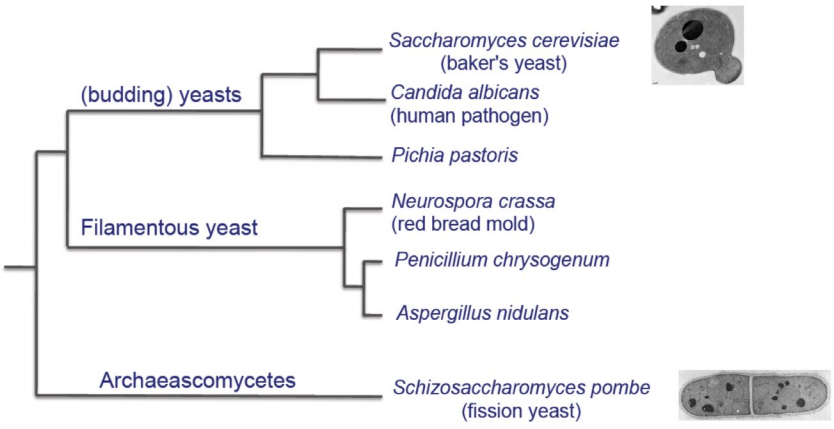
- ***Evolution:*** The proteins involved in Met and Cys synthesis show varied patterns of conservation during evolution.
- ***Structure and function:*** The structures of the proteins involved in Met and Cys synthesis are adapted to their catalytic roles.
- ***Information transfer:*** Met and Cys synthesis requires enzymes encoded by multiple genes.
- ***Pathways and energy transformation:*** The enzymes involved in Met and Cys synthesis are parts of intersecting energy-consuming pathways.
- ***Systems biology:*** The reactions involved in sulfur amino acid synthesis intersect with many other metabolic pathways in cells.

During the semester, students will develop proficiency in some key competencies of professional biologists. Working in teams, students will:

- propose hypotheses and design experiments to test their hypotheses.
- learn basic skills of molecular cell biology.
- collect, organize and interpret experimental data
- find and use information from the primary scientific literature and online databases.
- communicate scientific results in a series of short oral presentations and written reports.
- use feedback from their peers and the teaching staff to compile data from the short interim reports into a final poster and a final report written in the format of a scientific publication.

Pathways over Time: our research project

In the 2014-2015 academic year, we will explore the conservation of Met and Cys biosynthetic enzymes between the budding yeast, *Saccharomyces cerevisiae*, and the fission yeast, *Schizosaccharomyces pombe*. As their names imply, *S. cerevisiae* and *S. pombe* are sugar-loving fungi that were originally isolated from beer. *S. pombe* and *S. cerevisiae* are members of the phylum Ascomycota that can be propagated in both diploid and haploid forms. In response to various stresses, haploid strains of opposite mating types are induced to mate and undergo meiosis. The four spores generated from meiosis are contained within a resistant structure known as the ascus, from which the phylum derives its name. The two species are thought to have diverged from a common ancestor about 1 billion years ago (Hedges, 2002). Since their divergence, the *S. cerevisiae* lineage has undergone a whole genome duplication, followed by rounds of gene elimination and diversification (Mortimer, 2000). Today, the size of the *S. cerevisiae* genome (Kellis *et al.*,2004), ~12.5 Mbp, is similar to that of *S. pombe*. Because it has undergone less genome diversification, *S. pombe* is considered to be much closer to ancestral members of the p[illegible]



Diversification of selected yeast species within the Phylum Ascomycota
Different mechanisms of cell replication in *S. cerevisiae* and *S. pombe* are apparent in electron micrographs. (*S. cerevisiae* image reproduced with permission of Christopher Buser. *S. pombe* image from Hochstenbach *et al.,* Copyright National Academy of Sciences, U.S.A (1998), is reproduced with permission.)

Chapter 1

Of the two yeasts, *S. cerevisiae* is far and away the more thoroughly studied. Scientists have worked with genetically pure strains of *S. cerevisiae* for over a century, and it is widely used as a model organism (Botstein and Fink, 2011). *S. cerevisiae* has many of the same biochemical pathways as higher eukaryotes, but its genome is significantly smaller than vertebrate genomes and powerful genetic techniques are available for manipulating gene expression. For these reasons, the *S. cerevisiae* genome was the first eukaryotic genome to be sequenced in its entirety. Completion of the yeast genome sequence (Goffeau *et al.*, 1996) allowed researchers to prepare genome-wide collections of mutant strains (Winzeler *et al.*, 1999) and plasmids (Gelperin *et al.*, 2005) that are available to the yeast community. This semester, we will use *S. cerevisiae* strains with defined defects in Met and Cys biosynthesis as the hosts for homologous genes from *S. pombe*. If the *S. pombe* gene restores the ability of the *S. cerevisiae* mutant to synthesize Met, in a process known as **complementation**, we will know that gene function has been conserved over the evolutionary time that separates the two species.

Course overview

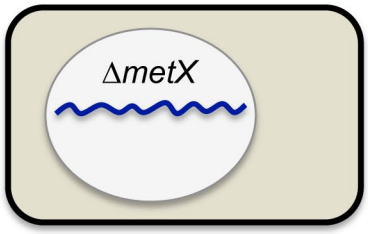
The course can be viewed as a series of six related units (see opposite page).

1. **Boot camp**: Students will become acquainted with basic laboratory equipment and techniques for handling and viewing yeast. Students will also be introduced to some of the many online databases, which are important sources of gene and protein information.
2. **Yeast deletion strains**: Students will use selective growth media and the polymerase chain reaction (PCR) to identify three *S. cerevisiae* mutants, each of which is missing a single gene involved in Met or Cys synthesis (Winzeler *et al.*, 1999). *Students will use one of the deletion strains for the remaining experiments of the semester.*
3. **Plasmid identification:** Students will isolate yeast overexpression plasmids from bacterial stocks and use restriction endonucleases to identify the plasmids. One plasmid contains the *S. cerevisiae MET* gene that is missing in their yeast strain (Gelperin *et al.*, 2005). A second plasmid carries the *S. pombe* homolog for the *S. cerevisiae MET* gene, and the third plasmid carries an unrelated bacterial gene.
4. **Transformation and complementation:** Students will transform their yeast deficiency strain with the three plasmids. Students will use selective plates to determine if the plasmid-encoded genes complement the *met* deletions in transformed strains.
5. **Analysis of protein expression:** Students will use western blots to analyze expression of plasmid-encoded genes. The plasmid-encoded proteins are fusion proteins with epitopes at their C-termini that can be recognized by antibodies.
6. **Team-designed experiment:** Teams will design and conduct their own experiments, based on questions that have arisen during the previous experiments.

Overview of the semester's experiments

1. Introduction to the lab (boot camp)
Micropipettes and measurement (Ch. 2)
Meet the microbes (Ch. 3)
Yeast culture techniques (Ch. 4)
Databases (Ch. 5)

2. Identify yeast deletion strains



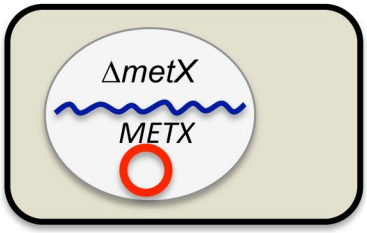
Identify missing genes in 3 Δ*met* strains by: their ability to grow on sulfur sources (Ch. 6) and yeast colony PCR (Chs. 7-8)

3. Characterize *MET* expression plasmids

Identify *S. pombe* homolog (Ch. 9)
Isolate plasmids (Ch. 10)
Restriction mapping (Ch. 11)

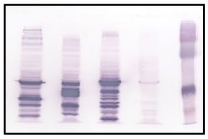
focus on one met strain

4. Transform strain with expression plasmids



Has complementation occurred? (Ch. 12)

5. Analyze Metp expression with western blots



Prepare cell extracts (Ch. 13)
Separate proteins by SDS-PAGE (Ch. 14)
Analyze expression on western blots (Ch.15)

6. Team-designed follow-up experiment

Chapter 1

Acknowledgments

It has been a pleasure to work with many Boston College colleagues and students on this fourth edition of the BI204 laboratory manual. In particular, Dr. Douglas Warner has helped to design the experiments and has contributed many revisions to the text. Many thanks to Holli Rowedder for her careful editing and her help with the experiments.

Over the past three years, many teaching assistants have provided able leadership for their sections and offered valuable feedback about the effectiveness of this manual and suggestions for improving both the manual and the course. Hundreds of undergraduate students have also contributed very constructive comments and suggestions. Special thanks are due to David Chou, Class of 2012, who designed the cover and layout of this manual.

Finally, I would like to acknowledge support provided by the National Science Foundation for the Pathways over Time project through grant NSF114028.

References

Bauerle C, DePass A, Lynn D *et al.* (2011) *Vision and Change in Undergraduate Biology Education: A Call to Action.* National Science Foundation/American Association for the Advancement of Science, Washington, D.C.

Botstein D & Fink GR (2011) Yeast: an experimental organism for 21st century biology. *Genetics* **189:** 695-704.

Duina AA, Miller ME & Keeney JB (2014) Budding yeast for budding geneticists: A primer on the *Saccharomyces cerevisiae* model system. *Genetics* **197:** 33-48.

Gelperin DM, White MA, Wilkinson *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

Goffeau A, Barrell BG, Bussey H *et al.* (1996) Life with 6000 genes. *Science* **274:** 563-567.

Hedges SB (2002) The origin and evolution of model organisms. *Nat Rev Genet* **3:** 838-849.

Hochstenbach F, Klis FM, vanden Ende H, van Donselaar E, Peters PJ & Klausner RD (1998) Identification of a putative alpha-glucan synthase essential for cell wall construction and morphogenesis in fission yeast. *Proc Natl Acad Sci USA* **95**: 9161-9166.

Kellis M, Birren BW & Lander ES (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428:** 617-624.

Mortimer RK (2000) Evolution and variation of the yeast (*Saccharomyces*) genome. *Genome Res* **10:** 403-409.

Thomas D & Surdin-Kerjan Y (1997) Metabolism of sulfur amino acids in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **61:** 503-532.

Winzeler EA, Shoemaker DD, Astromoff A *et al.* (1999) Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. *Science* **285:** 901-906.

Chapter 2

Mastering the micropipette

Welcome to the microworld! In this class, you will be working with microorganisms. These include yeast and bacteria, millions of which would fit into a period on this page. You will also be working with costly reagents, such as plasmids and enzymes. Therefore, in every experiment, you will be required to accurately measure volumes as small as a few microliters (μL). Micropipettes will allow you to do this accurately and precisely.

Objectives

At the end of this laboratory, students will be able to:

- select and adjust the most suitable micropipette to transfer a given volume.
- accurately transfer microliter volumes.
- use a spectrophotometer to measure light absorbance.
- explain how experimental errors affect measurements.

Chapter 2

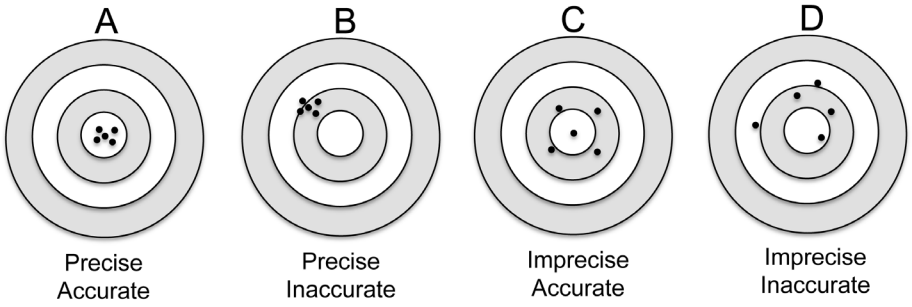
Using micropipettes correctly

Arguably, the most important scientific equipment that you will use in this class are adjustable micropipettes, which you will use in nearly every experiment. Micropipettes are precision instruments that are designed to ***accurately*** and ***precisely*** transfer volumes in the microliter range. You may use microliters or milliliters as the units of volume in your lab notebooks and lab reports, but be careful to always state the volume unit that you are using. Recall the relationships between volume units:

$$1 \text{ microliter (abbreviated } \mu L) = 10^{-3} \text{ milliliter (mL)} = 10^{-6} \text{ liter (L)}$$

Accuracy and precision

Accuracy depends on the micropipette delivering the correct volume. Precise results are reproducible. Let's use a target analogy to demonstrate the difference between accurate and precise results. Imagine that four students try to hit the bulls-eye five times. Students A and B are precise, while students A and C are accurate.

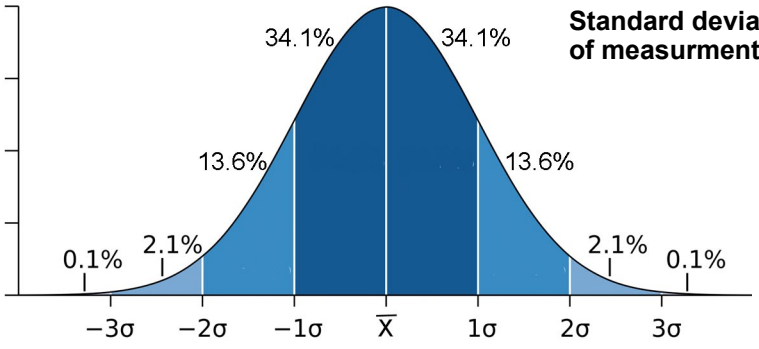


Manufacturers determine the accuracy and precision of micropipettes by using them to transfer defined volumes of distilled water to a drop that is then weighed on an analytical balance. The density of water is 1.0 gram per mL at 25˚C. The process is repeated several times during the calibration process, and the data is used to calculate the accuracy and precision of a micropipette.

Accuracy refers to the performance of the micropipette relative to a standard (the intended) value. Accuracy is computed from the difference between the actual volume dispensed by the micropipette and the intended volume. Note that this can be a negative or positive value. When micropipettes are calibrated, the accuracy is normally expressed as a percent of the selected value. Micropipettes are designed to operate with accuracies within a few percent (generally <3%) of the intended value. The accuracy of a micropipette decreases somewhat when micropipettes are set to deliver volumes close to the lowest values in their range.

Precision provides information about reproducibility, without any reference to a standard. Precision reflects random errors that can never be entirely eliminated from a procedure. Thus, a series of repeated measurements should generate a normal or binomial distribution (opposite). Precision is expressed as the standard deviation (σ) of the set of measurements. In a normal

distribution, ~2/3 of measurements will fall within one standard deviation of the average or mean ($\bar{x}$), and 95% of measurements will fall within two standard deviations of the mean. The standard deviation for a set of *n* measurements is calculated using the formula below.



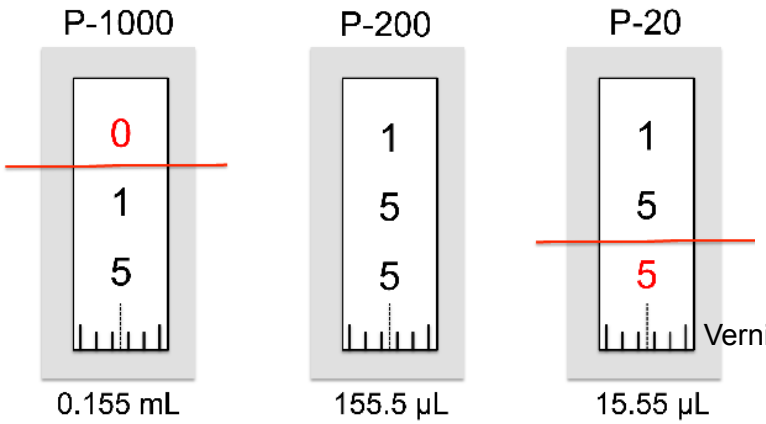
Standard deviation describes the distribution of measurments relative to the mean value

$$\text{Std. Dev.} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

(Adapted from http://en.wikipedia.org/wiki/Normal_distribution#mediaviewer/File:Standard_deviation_diagram.svg)

Choosing the micropipette

We use three different sizes of micropipettes in the laboratory, the P20, P200 and P1000. Our micropipettes have been purchased from several different manufacturers, but the principles of operation are the same. The numbers after the "P" refer to the maximum number of microliters that the micropipette is designed to transfer. Note that there is some overlap in the ranges of the different micropipettes. For example, both the P200 and P20 can be used to transfer 15 µl, but the P20 is more accurate within that range. As a rule of thumb, always select the smallest volume pipette that will transfer the volume.



Volume indicators on the P-1000, P-200 and P-20 micropipettes.

The top digit on the P-1000 (red on some models) indicates milliliters and the bottom digit on the P-20 (red on some models) indicates tenths of microliters. Note the differences in accuracy between the micropipettes.

Specifying the transfer volume

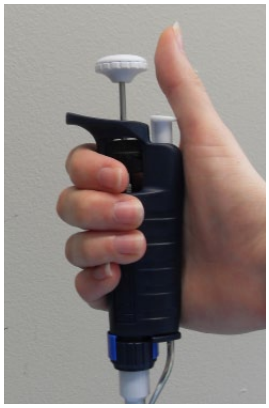
There are three numbers on the volume indicator. With each of the micropipettes, you will specify a volume to three digits by turning the volume adjustment knob. You will also be able to extrapolate between the lowest numbers with the vernier marks on the lower dial. Most of the measurements you will make with the micropipettes will be accurate to four significant figures!

NEVER turn the indicator dial beyond the upper or lower volume limits of the micropipette! This could damage the piston.

Chapter 2

Transferring volumes accurately

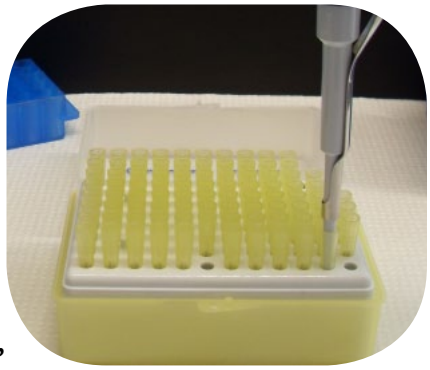
Micropipettes work by air displacement. The operator depresses a plunger that moves an internal piston to one of two different positions. The first stop is used to fill the micropipette tip, and the second stop is used to dispense the contents of the tip. As the operator depresses the plunger to the first stop, an internal piston displaces a volume of air equal to the volume shown on the volume indicator dial. The second stop is used only to dispense the contents of the tip.


Start


First stop


Second stop

Filling the micropipette



- Remove the lid from the box containing the correct size micropipette tips. P-1000 tips may be blue or white, while P-20 and P-200 tips are yellow or white.
- Attach the tip by inserting the shaft of the micropipette into the tip and pressing down firmly (figure on right). This should produce an airtight seal between the tip and the shaft of the micropipette.
- Replace the lid of the tip box to keep the remaining tips sterile. ***Avoid touching the tip (especially the thinner end), because the tips are sterile.***
- Depress the plunger of the micropipette to the FIRST stop.
- Immerse the tip a few millimeters below the surface of the solution being drawn up into the pipette. Pipetting is most accurate when the pipette is held vertically. Keep the angle less than 20˚ from vertical for best results.
- Release the plunger ***S L O W L Y***, allowing the tip to fill smoothly. Pause briefly to ensure that the full volume of sample has entered the tip. Do NOT let the plunger snap up. This is particularly important when transferring larger volumes, because a splash could contaminate the shaft of the micropipette. If you inadvertently contaminate the shaft, clean it immediately with a damp Kimwipe.

NEVER rest a micropipette with fluid in its tip on the bench!

Dispensing the contents of the micropipette

- Place the micropipette tip against the side of the receiving test tube. Surface tension will help to dispense the contents of the micropipette. Do NOT attempt to eject the contents of the micropipette into "thin air."
- Smoothly depress the plunger to the first stop. Pause, then depress the plunger to the second stop. The contents of the pipette should have been largely released at the first stop. The second stop ensures that you've released the "last drop."
- Use the tip ejector to discard the tip.

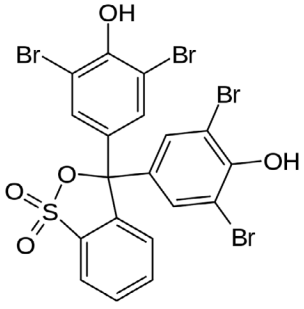
WARNING: The most common - and serious - operator error is depressing the plunger to the second stop before filling the micropipette tip.

DO NOT DO THIS!!!

Using the spectrophotometer to evaluate your pipetting skills

Since you will be using micropipettes for all of your experiments, the quality of your results will depend on proper operation of the micropipette. Today's laboratory will lead you through some exercises that will show you how to use micropipettes correctly and point out some common pitfalls associated with their use. Your results will also provide information about whether the pipettes are functioning properly.

In these exercises, you will be using the spectrophotometer to determine if your pipetting is accurate and precise. You will be using micropipettes to combine various volumes of water and solutions of a blue dye, bromophenol blue (BPB). You will measure the absorbance of the resulting solutions at 590 nm (A_{590}), which is close to the absorbance maximum of bromophenol blue at neutral pH. Measuring errors will be reflected in the spectrophotometer readings.



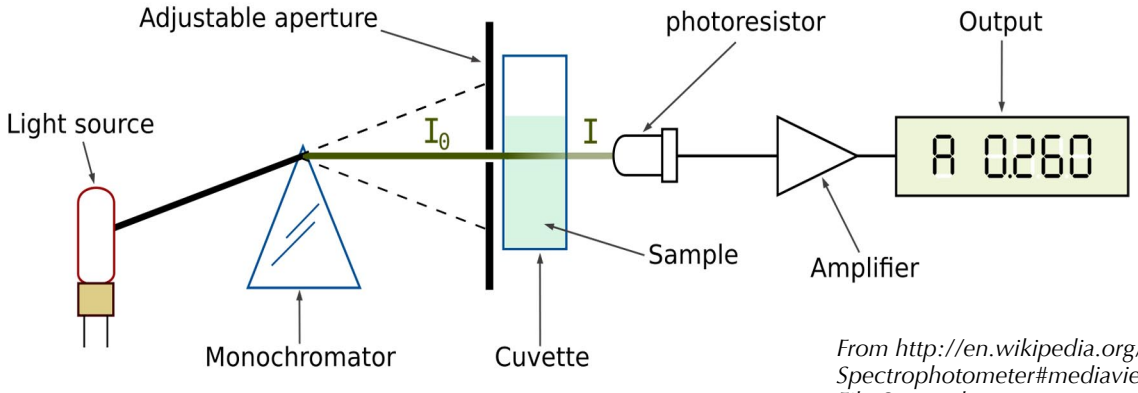
The spectrophotometer readings provide an indirect measurement of pipette performance. The proper way to calibrate the micropipettes would be to weigh out volumes of water, which has a specific gravity of 1.0 g/mL. Unfortunately, we do not have enough balances with sufficient accuracy for the class to perform the measurements. If you suspect inaccuracies in the micropipettes that you are using, refer them to the teaching staff, who will test them properly.

Chapter 2

Light spectroscopy

Spectrophotometers measure the amount of light absorbed by a sample at a particular wavelength. The absorbance of the sample depends on the electronic structures of the molecules present in the sample. Measurements are usually made at a wavelength that is close to the absorbance maximum for the molecule of interest in the sample.

The diagram below shows the elements present in a typical spectrophotometer. The light sources used in most spectrophotometers emit either ultraviolet or visible light. Light (I_o) passes from a source to a monochromator, which can be adjusted to allows only light of a defined wavelength to pass through. The monochromatic (I) light then passes through a cuvette containing the sample to a detector.



From http://en.wikipedia.org/wiki/Spectrophotometer#mediaviewer/File:Spetrophotometer-en.svg

The spectrophotometer compares the fraction of light passing through the monochromator (I_0) to the light reaching the detector (I) and computes the transmittance (T) as I/I_0. Absorbance (A) is a logarithmic function of the transmittance and is calculated as:

$$A = \log_{10}(1/T) = \log_{10}(I_0/I)$$

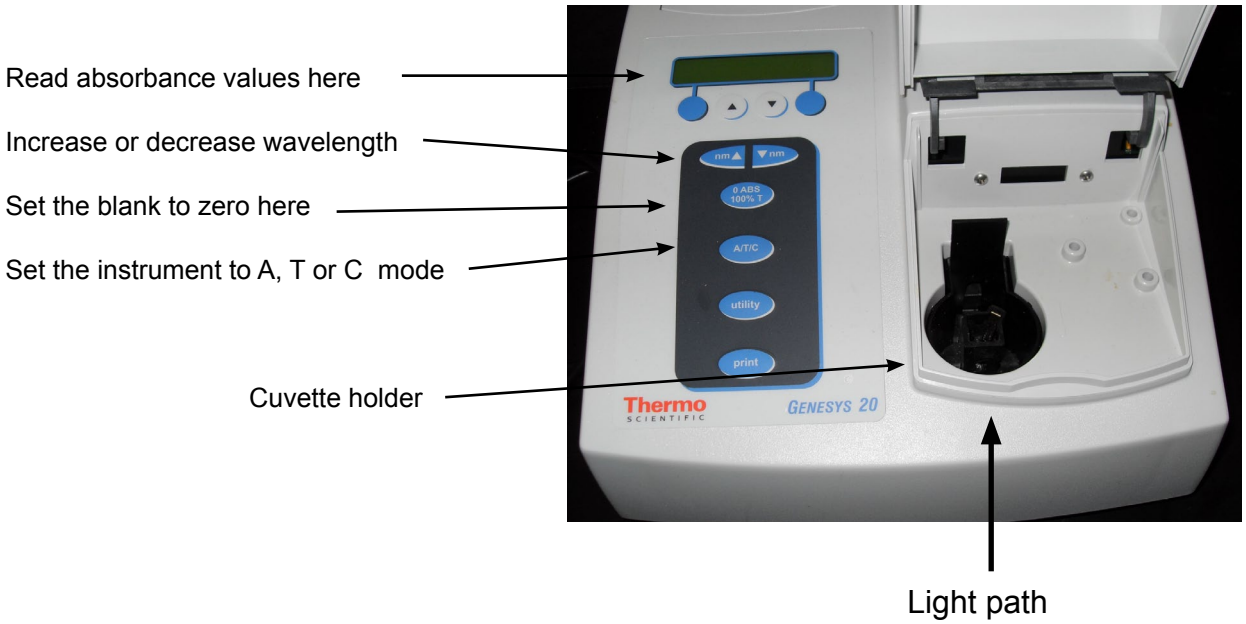
Spectrophotometers can express data as either % transmittance or absorbance. Most investigators prefer to collect absorbance values, because the absorbance of a compound is directly proportional to its concentration. Recall the Lambert-Beer Law, traditionally expressed as:

$$A = \varepsilon\, b\, C$$

where ε is the molar extinction coefficient of a compound, b is the length of the light path through the sample, and C is the molar concentration of the compound. Cuvettes are formulated to have a 1 cm light path, and the molar extinction coefficient is expressed as L/moles-cm. Consequently, absorbance is a unitless value.

Using the GeneSys 20 spectrophotometer

Our labs are equipped with GeneSys 20 spectrophotometers. Programming is very simple and is accomplished with a few buttons, as shown in the figure below. In our labs, we will be using the instruments in absorbance (A) mode, rather than the transmittance (T) or concentration (C) modes.



1. Calibrate the spectrophotometer. The spectrophotometer has a diagnostic program that needs to run before you start your measurements. ***Before turning on the spectrophotometer***, check that the cuvette holder is empty and close the lid. Turn on the power to the spectrophotometer and wait until the diagnostics are complete. Keep the door closed. The progress of the diagnostics can be monitored on the instrument panel.

2. The spectrophotometer should be in absorbance mode, its default setting. Adjust the wavelength to the proper value (590 nm) using the arrow keys.

3. Prepare a cuvette containing deionized water to serve as a blank. Insert the cuvette into the cuvette holder. Be sure the cuvette is oriented correctly in the light path. ***The light path in the GeneSys 20 is from front to the back of the instrument.*** Close the lid.

4. Press the 0 Abs/100% T key. This will zero the instrument. Remove this cuvette. Save this blank for others to use.

5. Place a cuvette containing your sample into the cuvette holder. Read the absorbance values and record them in your notebook.

6. Repeat step 5 with all of your samples, recording the absorbance readings in your notebook.

Chapter 2

Exercise 1 - Getting the feel of micropipettes

Concept: Micropipettes work by an air displacement mechanism

1. Set the P200 to deliver 200 µL. Be careful not to overshoot, which could damage the pipette piston.
2. Grip the pipette by wrapping your fingers around the barrel. Use your thumb to depress the plunger to its first stop.
3. Next press the plunger to the second stop. *Compare the distance that the plunger moved during the first and second strokes.*
4. Set the P200 to deliver 20 µL and depress the plunger to its first stop. *Compare the distance that the plunger moved when the P200 was set to 200 or 20 µL.*
5. Depress the plunger to the second stop. *How does the distance between the first and second stops compare for 200 and 20 µL?*
6. Set the P20 to deliver 20 µL. Depress the plunger to the first stop. *Compare the distance to the first stop when a P20 and P200 are set to deliver 20 µL.*

Concept: The filling and dispensing strokes are different.

1. Place a tip on the shaft of the P200.
2. Set the P200 to deliver 50 µL.
3. Draw up 50 µL of 0.05% BPB solution into the pipet.
4. Dispense the BPB into a microcentrifuge tube down to the first stop, holding the tip against the wall of the tube. Note whether all of the dye has been expelled. Push the plunger down to the second stop to release any remaining BPB.

Exercise 2 - How NOT to pipette!

1. Use the P1000 to add 990 µL of water to two microcentrifuge tubes. Label the tubes A and B. Dispose of used tips in the containers provided.
2. Use a P20 to correctly transfer 10 µL of 0.05% BPB to tube A. Make a mental note of what fraction of the pipet tip is filled with the dye. Use the vortex mixer to disperse the BPB in the water.
3. Use a P20 to ***INCORRECTLY*** transfer 10 µL of 0.05% BPB to tube B. Do this by depressing the plunger to the second stop before you take up the BPB solution. Make a mental note of how well the dye fills the tip this time.
4. Set the wavelength of the spectrophotometer to 590 (A_{590}). Pipette 1 mL of water into a plastic cuvette and blank the spectrophotometer at this wavelength.
5. Read the A_{590} of the solutions in tubes A and B, in the spectrophotometer. *How do the two readings compare? What kind of error results from drawing solution into the pipette incorrectly?*

Exercise 3 - How precise and accurate are your transfers?

Work in groups of three. One person in the group should work with the P-20, another with the P-200 and the third with the P-1000. Each person should prepare three identical samples and then determine the A_{590} of the three samples. From the data, you will be able to determine if the micropipette is measuring volumes correctly.

1. Each person in your group of three will work with a different micropipette and perform the same transfers in triplicate, as detailed below. The final volume (water + BPB) in each tube will be 1.0 mL. Calculate the volume of water that will need to be combined with each of the following to give 1.0 mL, and record your calculations in your lab notebook:

   Group member A: Use the P-20 to transfer 10 µL of 0.05% BPB to ____ µL of water.

   Group member B: Use the P-200 to transfer 100 µL of 0.005% BPB to ____ µL of water.

   Group member C: Use the P-1000 to transfer 300 µL of 0.005% BPB to ____ µL of water.

2. To minimize our plastic waste, strategize how to minimize the number of tips that you use without contaminating the stock solutions. A tip can be used multiple times, but a tip that has been used for BPB cannot be used to subsequently transfer deionized water. Combine the BPB solution and water to give a final volume of 1.0 mL.

3. Measure the A_{590} of the three solutions with the spectrophotometer and record the data in your notebook.
4. Compute the mean and standard deviations for your three measurements, using either a calculator or Excel. The standard deviation reflects the precision of your measurements.
5. Enter these values on the chart that your TA has prepared on the whiteboard. Compare the values that your group obtained each of the three pipettes with the aggregated class measurements. If the averages that your group obtained are significantly different than those that other groups obtained with the same micropipette, your micropipette may not be transferring volumes accurately.

Notify your TA if any of the micropipettes are not performing properly. Your TA will follow up on your observations and test the micropipettes with the gravimetric test described earlier in the chapter.

Test yourself

As part of an interview for a research position, three applicants are asked to transfer 150 μL of distilled water with a P-200 micropipette to a weighing paper and to determine the weight of each drop with an analytical balance. The three measurements made by each of the applicants are listed below. Use the space below to calculate the mean and standard deviation of the measements made by each student.

Applicant A: 0.161 g, 0.147 g, 0.142 g
Applicant B: 0.158 g, 0.156 g, 0.157 g
Applicant C: 0.143 g, 0.153 g, 0.150 g

Which applicant makes the most precise measurements?

Which applicant makes the most accurate measurements?

Which applicant would you hire? Why?

Chapter 3

Meet the microbes

Microorganisms are too small to be seen with the naked eye, so biologists use microscopes to see them. In this lab, you will use the light microscope to observe the budding yeast, *Saccharomyces cerevisiae*, and the fission yeast, *Schizosaccharomyces pombe*. Since the two yeast diverged from a common ancestor, they have evolved distinct morphologies and controls on cell division that you will study in this lab. You will also observe the much smaller bacterium, *Escherichia coli*, which is one of the workhorses of molecular biology.

Objectives

At the end of this laboratory, students will be able to:

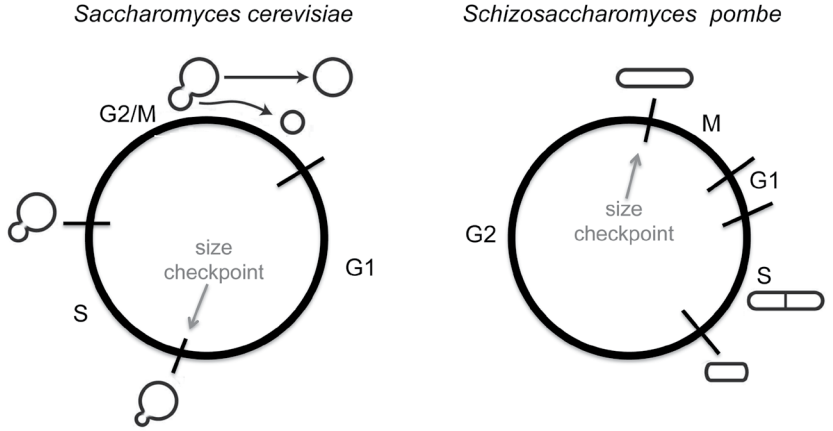
- identify the components of a compound light microscope.
- adjust a light microscope to observe bacterial and yeast cultures.
- stain microbial samples with iodine to improve their optical contrast.
- distinguish *E. coli* and yeast by their sizes.
- distinguish *S. cerevisiae* and *S. pombe* by their morphological characteristics.

Chapter 3

Microscopes are essential for viewing microorganisms. The first person to observe yeast and bacteria was Anton van Leeuwenhoek, who called them *animalcules*. Yeast cells typically have diameters of ~10 μm, while bacteria have diameters of ~1 μm, both of which are far too small to be seen without considerable magnification. In this lab, you will use the compound light microscope to observe two eukaryotic yeast species, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, as well as the much smaller prokaryote, *Escherichia coli*.

Two very different yeast

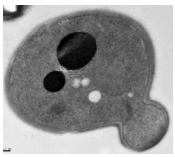
S. cerevisiae and *S. pombe* are both members of the phylum Ascomycota, but they use dramatically different methods of reproduction. Befitting their popular names, *S. cerevisiae* reproduces by budding, while *S. pombe* reproduces by fission along a central plane. During each round of cell reproduction and division, the two yeasts show predictable changes in morphology. *S. cerevisiae* and *S. pombe* mutants that do not display these characteristic changes in morphology contributed to our current molecular understanding of the cell cycle - and the 2001 Nobel Prize.



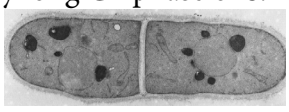
Cell cycle control is different in *S. cerevisiae* and *S. pombe*.

The principal size checkpoint in *S. cerevisiae* occurs at the G1/S boundary. The corresponding checkpoint in *S. pombe,* which spends most of its cell cycle in G2, occurs at the G2/M boundary.

As above, *S. cerevisiae* and *S. pombe* spend different fractions of each cell cycle in the G1, S, G2 and M phases, because the principal size checkpoint occurs at a different place in the cycle for each species (Turner *et al.*, 2012). In *S. cerevisiae* (right), buds begin to form when cells enter S phase. The size of the bud, which will become the daughter cell, continues to grow until the cells divide in M phase. When the cell divides, the daughter cell is still smaller than the mother cell. The daughter cell will need to grow a bit before it enters another round of cell division. *(Electron micrograph by Christopher Buser, reproduced with permission.)*



By contrast, *S. pombe* (right) divides by medial fission. Cells grow in length until they are 12-15 μm, when the cell divides and a septum begins to form. The unusually long G2 phase of *S. pombe* may reflect the fact that it is found primarily as a haploid in nature (unlike *S. cerevisiae*, which is found in both diploid and haploid forms). Because *S. pombe* has two copies of each gene during most of the cell cycle, it is able to survive a lethal mutation in one copy of an essential gene. *(Electron micrograph from Hochstenbach et al., 1998)*



The cells in both yeast cultures are in "log (short for logarithmic) phase," when cells are dividing exponentially (Chapter 4). The cultures are not synchronized and will contain cells at all the points in the cell cycle. Our microscopes do not provide the same kinds of magnification as the electron micrographs shown above, but you should be able to see some subcellular compartments when you view the yeast at 1000X magnification.

Escherichia coli is a small prokaryotic cell

In this course, we will be using the proteobacterium, *E. coli*, to propagate copies of *MET* genes on plasmids, which are small circles of DNA that replicate in the *E. coli* cytoplasm. Unlike the *E. coli* strains that you hear about in food-borne outbreaks, our lab strain has been engineered for use in molecular biology labs and is unable to colonize the human intestinal tract. *E. coli* are particularly useful to molecular biologists because they rapidly grow to very high densities in laboratory culture media, reaching densities of 1-10 billion cells per mL. Although *E. coli* are 10-100 times smaller than yeast cells, their sheer numbers and their distinct motion renders them visible with the light microscope. Nonetheless, you will need use the microscope at its highest magnification to see any detail of the cells.

Observing microorganisms with light microscopy

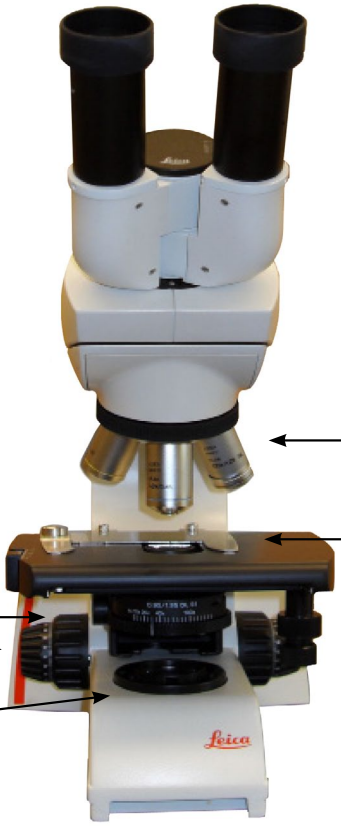
Light microscopes have a maximum resolution of ~0.2 μm, which is sufficient to resolve individual yeast cells and provide rough infomation about their intracellular organization. (More detailed information about subcellular structure requires an electron microscope.) Compound light microscopes use a system of lenses to gather and focus light passing through a specimen and to project the image on the viewer's retina. The specimens used for light microscopy are usually stained to increase their contrast prior to observations. Today, a large number of specialized reagents and protocols for staining cells have been described, and investigators select stains to suit the purposes of their individual experiments. In this lab, we will use an iodine solution that stains glycogen particles to increase the contrast of the cultures.

Our labs are equipped with Leica DM500 light microscopes (see the following page). Light from an LED source at the base of the microscope enters a condenser that focuses the light that will reach the specimen on the microscope stage. Users are able to control the amount of light reaching the specimen by opening or closing an iris diaphragm on the condenser. The microscope has four, interchangeable objective lenses, with magnifications of 4X, 10X, 40X and 100X. Ocular lenses in the eyepieces magnify specimens an additional 10-fold, producing final magnifications of 40X, 100X, 400X and 1000X. The lenses on the DM500 are parfocal, meaning that specimens remain reasonably well-focused when the lenses are changed. When working with the microscope, ***ALWAYS BEGIN WITH THE LOWEST POWER OBJECTIVE***, which is easiest to focus, and work your way up to the higher power objectives.

Chapter 3

Leica DM500 Light microscope

Eyepieces have 10X magnification
Interpupillary distance is adjustable



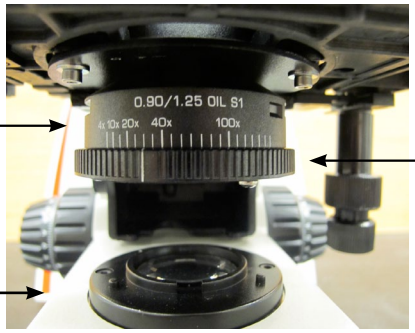
Parfocal **objective lenses**
4X, 10X, 40X and 100X

Slide holder is mounted on microscope stage

XY Controls for moving stage

Coarse focus
Fine focus

Light source



Condenser focuses the light reaching the specimen

Iris diaphragm regulates the amount of light reaching the condenser

Dimmer switch
(not visible here)

We will also be using the 100X lens, which is also known as an oil immersion lens. ***This is the longest, most powerful and most expensive lens on the microscope, requiring extra care when using it.*** As the name implies, the 100X lens is immersed in a a drop of oil on the slide. Immersion oil has the same refractive index as glass, so it prevents light from bending as it enters the lens. The oil should be removed immediately from the 100X lens after use. Oil should NEVER touch the 4X, 10X or 40X lenses, which are destroyed by the oil.

Exercise 1 – using the compound light microscope

NOTE: Lenses are fragile and expensive—treat them with care!

Objectives should NEVER touch the slide!

Clean lenses with LENS PAPER only. Kimwipes™ and other paper may scratch a lens.

1. Identify the parts of the microscope. Note the positions of the objectives, the coarse and fine focus adjustments, the adjustable light switch and the condenser diaphragm. Adjust the positions of the eyepieces to fit the distance between your eyes.
2. Locate the four objective lenses on the microscopes. The magnification of each lens (4x, 10x, 40x, and 100x) is stamped on its casing. Rotate the 4x objective into position. Adjust the position of the iris diaphragm on the condenser to its corresponding 4x position.
3. Turn on the microscope lamp and adjust the dimmer switch until the light is not too intense when you look through the eyepieces. You may need to adjust the distance between the eyepieces to fit your eyes.
4. Place the transparent ruler on the microscope stage. Use the coarse focus knob to bring the ruler into focus. You may also need to adjust the light. Make additional adjustments with the fine focus knob. The ruler is graduated into 1 mm divisions. (Remember that 1 mm is equal to 1000 μm, the unit of distance usually used by microscopists.)

   What is the diameter of the field of view with the 4x objective?

5. Use the stage manipulators to move the ruler to either the right or the left.

   What direction does the image move?

6. Dial the 10x objective into position and adjust the condenser diaphragm.

   How does the distance between the specimen and the objective change?

   Adjust the focus with the coarse and fine focus adjustment knobs.

   What is the diameter of the field of view with the 10x objective?

   Can you find a mathematical relationship between the magnification and the diameter of the field of view?

7. Swing the 40x objective into position and adjust the condenser diaphragm. Adjust the focus using ONLY the fine focus knob.

   What happens to the working distance and the field of view?

Chapter 3

Exercise 2 – observing yeast and bacterial cultures

Students should work in groups of three. Each student should prepare one of the three slides listed below. Students in a group should share their slides. Each group will receive three cultures:

SC - *S. cerevisiae*
SP - *S. pombe*
EC - *E. coli*

1. Label the three slides. Each slide will contain two samples for comparison. The slides are large enough to accommodate two samples—and two coverslips. Number the slides with a Sharpie (Use the frosted area, if the slide has one.) As you work, be sure to record which of the two samples is closer to the labeled end of the slide. Use the spaces provided below to record your data.

Slide 1: compare cultures of *S. pombe* and *S. cerevisiae*.
Slide 2: compare cultures of *S. cerevisiae* and *E. coli*.
Slide 3: cultures of *S. pombe* and *E. coli* with the 100X objective (1000X).

2. Prepare concentrated suspensions of cultured cells. All of the cell cultures appear cloudy, because the cell densities are high. Even so, you will be observing a few microliters of cell suspension under the microscope, and cells may be few and far between. Concentrate the cultures to ensure that there will be enough cells in the microscope's field of view for meaningful observations.

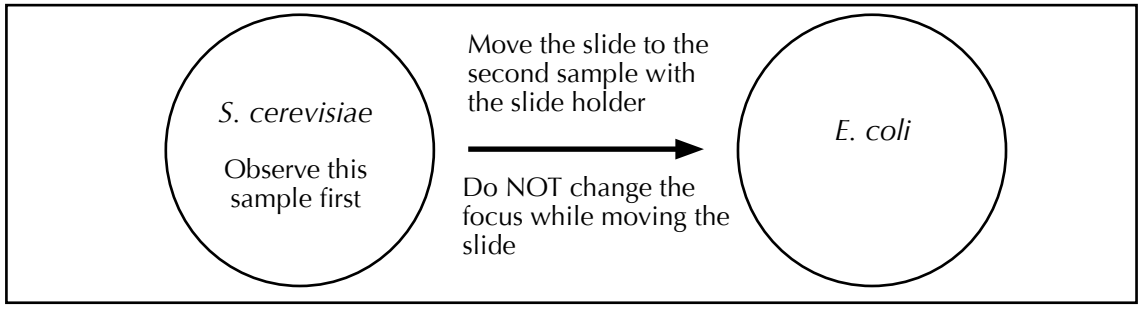
- Concentrate the cells by centrifuging the test tubes containing the cultures for a count of 10 in a microcentrifuge set at top speed. Hold down the Quick button on the Labnet microcentrifuges or the button between the two dials on the Eppendorf microcentrifuges.
- Remove most of the culture medium with a P1000 micropipette, until the medium just covers the cell pellet.
- Resuspend the cells with the vortex mixer.

3. Transfer and stain the cell samples.

- Transfer 2 µL of each cell suspension to the slide, using a P20 micropipette.
- Stain the cells by adding a 6 µL of Gram's Iodine to each cell suspension with a P20 micropipette. Cover each sample with a coverslip.

4. Observe samples of the three species using the 40X and 100X objectives. Each member of your group should observe the two species on the slide that he/she prepared. Work with another member of the group to make observations of the third species. Record your observations and answer the questions in the spaces below.

The oil immersion lens is required for the 100X objective, and this oil can damage the 40X objective. Be careful to follow the instructions below exactly. The steps must be performed in the correct order to protect the lenses!



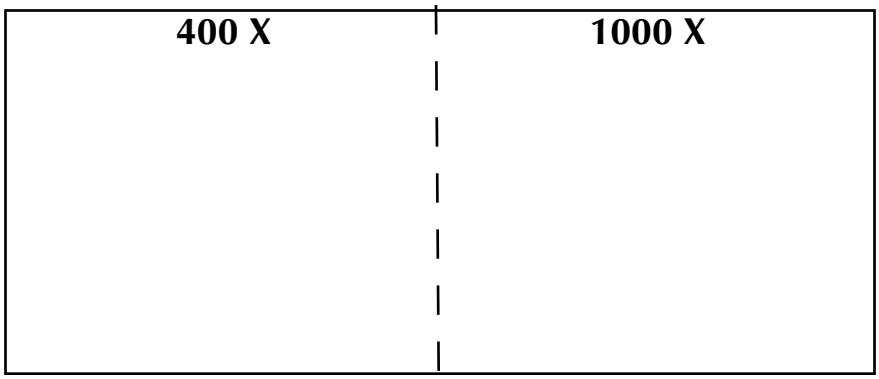
1. Focus the microsope on **EITHER an *S. cerevisiae* or *S. pombe* culture**. Focus first with the 10X objective, then move to the 40X objective and refocus with the fine focus control. Record your observations.
2. You are now ready to move to the 100X oil immersion objective! Rotate the nosepiece halfway between the 40X and 100X objectives. WITHOUT moving the stage, apply a drop of immersion oil on top of your coverslip where the light is shining through the slide. *SLOWLY* rotate the 100X objective into place, submerging it into the drop of oil. Use the FINE focus knob to bring the yeast into focus. Record your observations.
3. Rotate the nosepiece halfway between the 40X and 100X objectives again. ***Do NOT attempt to move the stage with the 100X objective in place.***
4. Use the XY controls to move the stage to the second sample on the slide WITHOUT changing the focus.
5. Rotate the 40X lens into position. Adjust the focus and record your observations.
6. Repeat step 2 above and record your observations at 100X.
7. Rotate the 100X objective out of position. Remove the slide and discard it in the glass waste. Clean the 100X lens with lens paper. Check that no oil is present on any of the other lenses.

S. cerevisiae observations

Sketch the cells that you see with the two different lenses in the boxes below.

What fraction of the cells show buds?

Can you detect any subcellular structures at 1000X magnification?

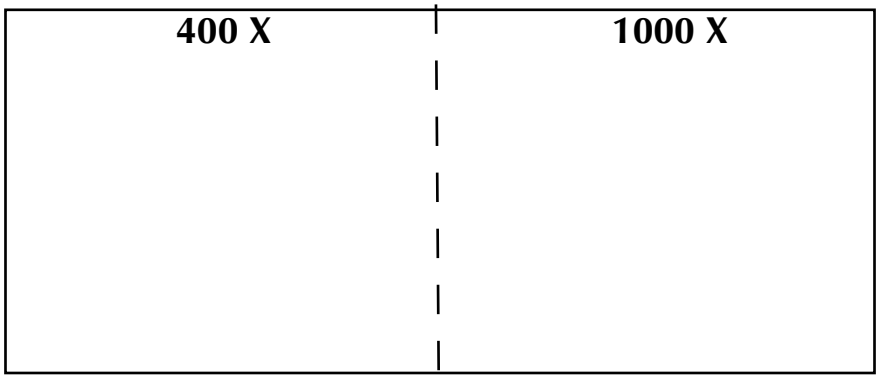


S. pombe observations

Sketch the cells that you see with the two different lenses in the boxes below.

What fraction of the cells show septa?

Can you detect any subcellular structures at 1000X magnification?

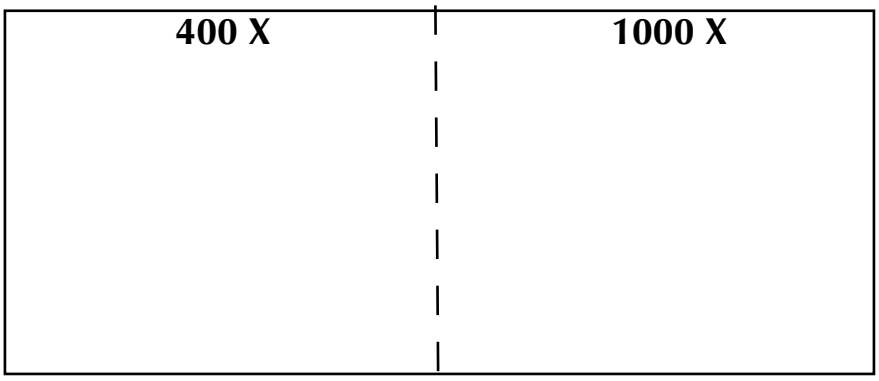


E. coli observations

Sketch the cells that you see with the two different lenses in the boxes below.

What fraction of the cells are motile?

How would you describe the structure of an *E. coli*?



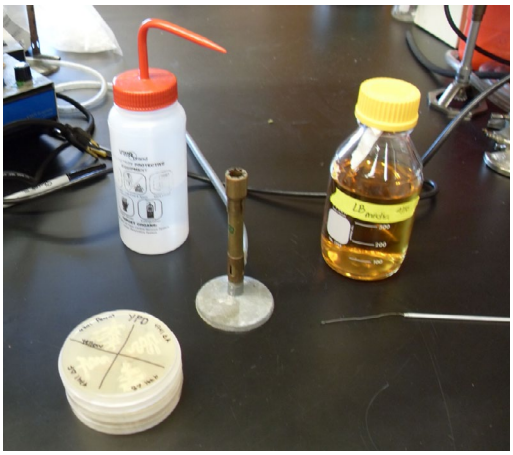
References

Hochstenbach F, Klis FM, vanden Ende H, van Donselaar E, Peters PJ & Klausner RD (1998) Identification of a putative alpha-glucan synthase essential for cell wall construction and morphogenesis in fission yeast. *Proc Natl Acad Sci USA* **95**: 9161-9166.

Turner JJ, Ewald JC & Skotheim JM (2012) Cell size control in yeast. *Curr Biol* **22:** R350-R359.

Chapter 4

Working with Yeast



This lab will introduce you to standard techniques used in microbiology. Very similar techniques are used to culture yeast and bacteria, although the culture conditions are optimized for each organism. In this lab, you will learn sterile techniques required for maintaining the integrity of yeast strains in the lab, as well as methods for culturing cells and estimating cell numbers.

Objectives

At the end of this lab, students will be able to:

- define and describe the relationship between species, strains and genotypes.
- use sterile techniques to culture yeast strains on agar plates.
- describe the phases of microbial growth.
- estimate the density of cells in a liquid culture using spectrophotometry and spot plating.

Chapter 4

In this project, you will be working with multiple strains of yeast. **Strains are microorganims of the same species that are derived from a single cell and are assumed to possess the same genotype.** Laboratory strains have often been derived from other strains by careful planning and experimentation, and their genotypes are at least partially defined. Strains are named according to that laboratory's conventions, but the strain names are not usually informative about the strain's genetic composition. Yeast strains are easily cultured in media that contains a carbon source, a nitrogen source, salts, vitamins and essential minerals.

Your success in this lab will depend on your ability to use sterile culture and transfer techniques that will maintain the genetic isolation of your yeast strains. An equally important element in laboratory success is careful bookkeeping, since yeast strains look alike! In this lab, you will learn basic techniques used to culture and quantify microorganisms. You will prepare stock cultures of *S. cerevisiae met* mutant strains on streak plates. You will also learn how to quantify the number of cells in liquid cultures of *S. cerevisiae* and *S. pombe cells* using spot plates and the spectrophotometer.

Sterile technique

Sterile technique is ESSENTIAL when working with microorganisms! It is important to protect strains from contamination with other strains and from the many undefined microbes in the environment. Large numbers of diverse microorganisms are all around us - in the air, on laboratory surfaces, on your skin and on your clothing. True to their name, microorganisms are too small to be detected by the eye, but they grow rapidly in laboratory culture media. Correct transfer techniques and the use of sterile reagents are usually enough to prevent contamination.

Some simple precautions will reduce the possibility of contamination:

- Wipe down a small working area on the lab bench with 70% ethanol.
- Light a Bunsen burner in your work area while working with strains. The burner produces an updraft that prevents airborne microorganisms from falling into cultures.
- Use sterile reagents, micropipette tips, and test tubes. Tips and microcentrifuge tubes should be kept in covered containers when not in use.
- Minimize contamination from clothing and body surfaces. Pull back and secure long hair. Avoid touching or breathing on sterile surfaces that will contact microorganisms.
- Avoid talking when you are transferring strains.
- Work quickly. Minimize the time that the tops are removed from vessels containing microorganisms or media.
- Keep caps right-side up to prevent contamination from airborne microbes.

The culture media and reagents that we will use have been sterilized by either autoclaving or filtration. An autoclave is a chamber that uses pressurized steam to kill cells on surfaces or in

solutions, using temperatures near 121˚C and pressures from 30-40 psi. (For comparison, atmospheric pressure is ~15 psi.) Ultrafiltration is used in the place of autoclaving when solutions contain temperature-sensitive compounds. The pores in the filters are typically 0.2 or 0.45 μm in diameter. These pores are sufficiently small to prevent the passage of bacteria, but not viruses or macromolecules.

Autoclave

Ultrafiltration unit

Yeast growth media

For routine culture, scientists usually use rich media that supply all the nutrients that cells need to grow. The individual components of rich media are often undefined. For example, yeast are commonly grown in a medium known as **YPD**, which is simple and inexpensive to prepare. The "**Y**" in YPD refers to a yeast extract, which contains the water-soluble compounds generated when yeast are forced to self-digest. (Those of you who have visited Australia may have encountered yeast extract in the popular spread, Marmite.) The "**P**" refers to peptone, a mixture of peptides and amino acids prepared by digesting animal protein with proteases. The "**D**" refers to dextrose, or glucose, which is the favored carbon source of yeast (Sherman, 2002).

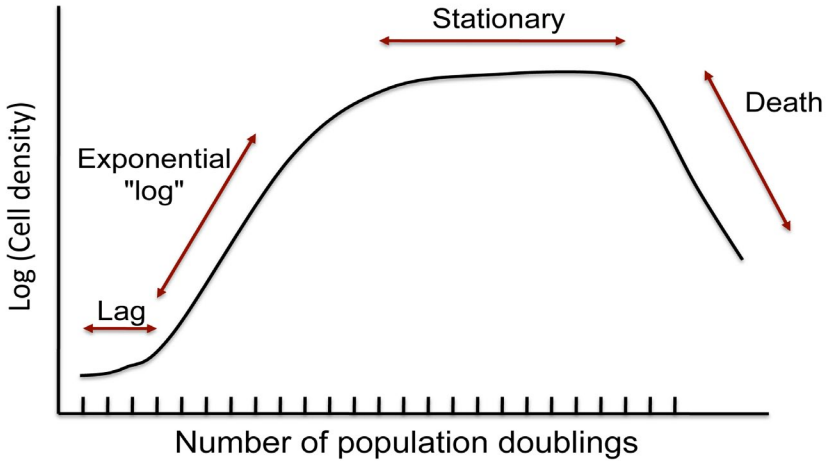
Because YPD is composed largely of crude extracts, its composition may show significant batch-to-batch variation. This variation is rarely a problem, because YPD contains more than enough essential nutrients to satisfy the metabolic requirements of cells. Many experiments, however, require media with a defined composition. To meet this need, the yeast community has developed a variety of synthetic media. Individual components of the synthetic media may be manipulated to suit the needs of an experiment. (Later in the semester, we will use defined media to select for particular genotypes.)

Yeast can be grown in liquid cultures or on the surface of plates containing solid media. Agar is usually used to solidify liquid growth media when preparing plates. Strains are typically maintained on agar plates for routine use. Cells grow in colonies on plates. The cells in a colony are genetically very similar, if not identical, because they are derived from the same progenitor cell. Most yeast strains can be stored on plates in the refrigerator for several months with minimal loss of viability.

Chapter 4

Yeast growth phases

When yeast are grown in liquid medium, the culture follows a well-established pattern for microbial growth. (Bacteria follow this same general pattern, although they divide much more rapidly.) Cultures are usually started by inoculating media with a small number of cells. A lag phase follows the inoculation, during which cells become acclimated to the new environment and begin to condition the media with their own metabolites. ***Lag phase*** is followed by an ***exponential,*** or ***log phase***, when the number of cells increases exponentially.



The exponential growth of yeast can be described by the equation:

$$N = N_0 e^{kt}$$

where N represents the number of cells at any time (t), and N_0 represents the number of cells at the beginning of the interval being analyzed. Scientists often find it convenient to think of the growth constant k in terms of the doubling time of the culture. In this rendering, k = ln2/T (T = the doubling time of the culture). The growth rate of yeast varies with temperature. Yeast grow well at room temperature, but they grow more rapidly at 30˚C. Well-aerated cultures grow more quickly than those that are not, so liquid cultures are usually grown on a rotary shaker or rotating wheel. At 30˚C, wild-type yeast strains have a doubling time of ~90 minutes in YPD.

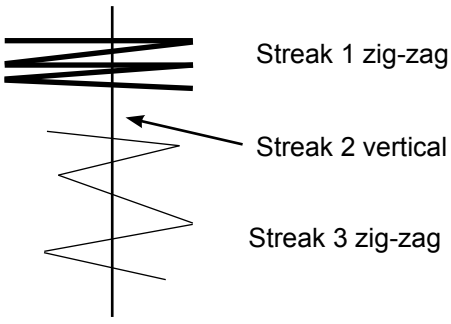
After a few doubling times, cells begin to deplete the nutrients in the culture, their growth rate slows, and the cells enter ***stationary phase.*** Yeast entering stationary phase adjust their metabolism by altering the transcription of hundreds of genes, leading to many physiological changes, including the accumulation of carbohydrate reserves and the assembly of a more resistant cell wall (reviewed in Werner-Wasburne *et al.*, 1993). In stationary phase, the rate of cell division is similar to the rate of cell death, so the number of cells does not change appreciably. Cells can survive in stationary phase for extended periods of time, resuming growth when conditions are favorable. Eventually, cells enter ***death phase*** if conditions do not improve.

Exercise 1 – Streak plates

Microbiologists like to begin their experiments with a single colony, because the cells in a colony are the progeny of a single cell. A concern in all genetic experiments is unknown mutations that arise spontaneously and may affect the phenotype being studied. Spontaneous mutations arise constantly in all cells, with a rate of approximately 10^{-8}/base/generation. For *S. cerevisiae*, with a genome of 12 Mbp, most cells will have accumulated at least one mutation by the time that they have undergone 9-10 divisions. A colony, which has hundreds of millions of cells, is therefore a population of genetically very similar, but not completely identical, organisms.

Researchers commonly use streak plates to isolate single colonies. A streak plate is actually a serial dilution of an existing culture on solid media. Researchers begin a streak by picking up a small sample of yeast or another microorganism with a **sterile** instrument, which could be a platinum loop, a toothpick or micropipette tip. They then spread the culture by making a series of zig-zag strokes across the surface of the plate. The number of cells on the loop or toothpick decreases as the streak progresses. Consequently, streaks appear thickest at their starting points, and the streak thickness decreases until it is possible to detect well-isolated single colonies near the end of the streak. Because it may be difficult to resolve colonies from a single streak, many labs use a series of streaks on the same plate to separate colonies. Each new streak is done with a fresh toothpick or micropipette tip that picks up cells by crossing over the tracks of the previous streak, before beginning a new series of zig-zags. In our experiments, we will use a multi-streak protocol, which allows us to culture multiple strains on a single plate of culture medium. (See the figure below.) The streak plates that you prepare will contain the stocks for future experiments. As you streak your strains, pay careful attention to detail to avoid cross-contamination or confusion about the identities of individual strains.

Streak plate with three sectors.
Plate has been divided into three clearly labeled sectors. Three streaks were used to spread the cells in each sector. The third streak in each sector contains well-separated colonies that can be used for genetics experiments.



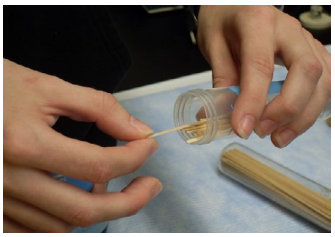
Chapter 4

Preparing a streak plate

1. Your team will be assigned three different *S. cerevisiae met* strains to culture. Gather the YPD stock plates with the parent strains to be propagated, sterile toothpicks, and an agar plate with fresh YPD media. This last plate will serve as your team's stock plate for the semester.

2. Divide your team's stock plate into sectors by marking the bottom of the plate with a magic marker. CLEARLY label each sector with a code for the strain that will be streaked in it. Keep the labels at the rim of the plate and use small letters. ***Note your initials and the date.***

3. Carefully shake the container of sterile toothpicks until you can grab the tip of one toothpick without touching any of the others in the container with your fingers. Alternatively, you can use a plastic micropipette tip. Carefully draw it out of the container. (Note: toothpicks and tips will have been sterilized in the autoclave; they should not be placed in the flame.)



4. Partially remove the lid of the petri dish containing the parent strains with one hand and hold the lid at an angle as you work. With the other hand, lightly touch a colony with the tip of a sterile toothpick or inoculation loop. Cultures should be *barely* visible on the tip. Avoid removing too many cells. Parent cultures are very concentrated, and a barely perceptible drop contains millions of cells. If the starting volume is too high, it may be difficult to separate single colonies on the new plate.

5. Transfer cells to the plate with fresh media. Make several zigzags across the outside edge of one sector with the toothpick or loop. *LIGHTLY* touch the agar surface as you move the toothpick. Think of pushing a hockey puck across an ice rink, rather than digging a ditch. Replace the lid and discard the toothpick.

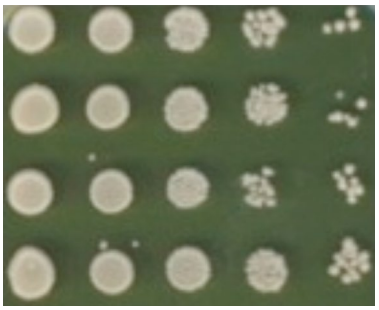
6. Using a new sterile toothpick, make a second vertical streak from the rim of the plate toward the center, staying within the sector. The streak should cross the zigzags in the first streak. Discard the toothpick.

7. Using a new sterile toothpick, make a third series of zigzags that cross back and forth over the straight second streak, beginning at the outer edge of the plate and moving toward the center. Be careful to stay within the sector.

8. Invert the plate and incubate it at 30˚C until individual colonies are visible, which is usually 24-48 hours.

Exercise 2 - Spot plates

Scientists use spot plates both to calculate the number of cells in cultures and to obtain information about the growth properties of strains on different media. The figure below shows an example of a typical spot plate. Each row represents a dilution series from a different yeast culture. The same volume of diluted culture is used for each spot. The dilution series is planned so that the most dilute spots contains a small number of individual colonies that can be distinguished from one another, typically less than ten.



Spot plate.
Each row on the plate represents a series of 1:10 dilutions of a liquid culture of *S. cerevisiae*. Five µL of each dilution was spotted on the plate. The plate was incubated for two days at 30˚C. Individual colonies are apparent at the highest dilution of each extract.

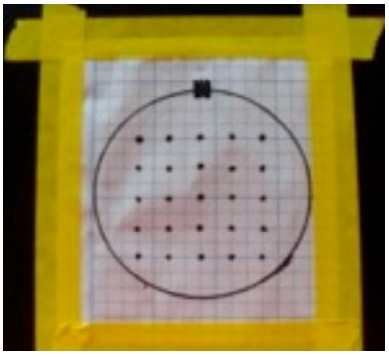
Most commonly, investigators make a series of 1:10 dilutions in **sterile** (**NOT deionized**) water and then spot a few microliters of each dilution in a row. In this experiment, 5 µL aliquots were spotted from the serial dilutions. Note that it is possible to count individual colonies in the most dilute samples. This in turn enables you to calculate the number of viable cells in the original culture. In the top row, you can distinguish 4 colonies in the sample that has been diluted 100,000-fold. The original culture would have contained 400,000 cells in 5 µL, which corresponds to 80 million cells per mL (8×10^7 cells/mL).

In this experiment, you will use spot plates to estimate the cell densities of log phase and stationary phase cultures of *S. cerevisiae* and *S. pombe*. Each group will receive four cultures:

CL - *S. cerevisiae* log phase culture
PL - *S. pombe* log phase culture
CS - *S. cerevisiae* stationary phase culture
PS - *S. pombe* stationary phase culture

Spot the dilution series from each culture on a separate row of the plate. Be sure to LABEL the rows!

Preparing the spot plate



1. Alignment grids are useful for preparing good-looking spot plates! Obtain an alignment grid (right) and mark the target positions for culture dilutions. Place an orientation mark at one point along the circumference.

2. Label the plate with your initials and date with small letters around the BOTTOM rim of the dish. Put a hash mark on the edge of the plate to serve as an alignment marker.

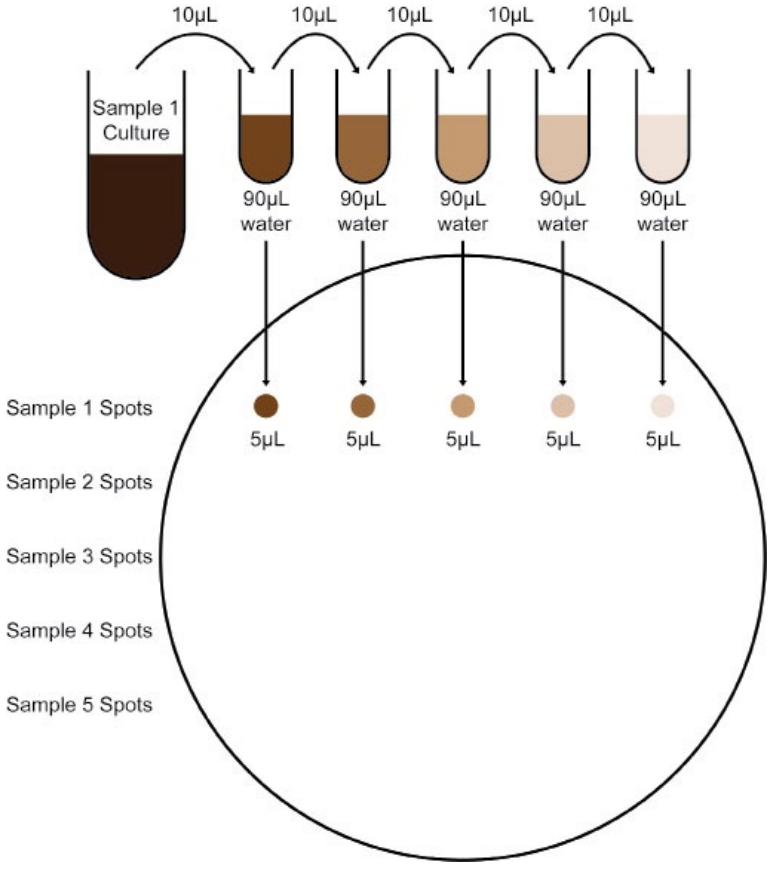
3. Prepare a series of five 1:10 dilutions from each culture using sterile distilled water. (Diagrams in your lab notebook are often helpful in designing dilution series.) To prepare a serial dilution, first pipette 90 µL sterile water into five microcentrifuge tubes. Next, use a P20 to transfer 10 µL from the original culture into the first tube. ***Vortex the cell suspension***, and then transfer 10 µL from this tube to the second tube in the series, and so on. *Use the same pipette tip for the entire dilution series.* Eject the tip into the appropriate waste container.

4. *Beginning with the last dilution in the series*, spot 5 µL spots in a row. Vortex each dilution before spotting it, because cells may have settled. Again, you will be able to use a single pipette tip for a dilution series, since you started with the most dilute sample.

5. Repeat step 3 for each culture that you are analyzing. Be careful to note in your lab notebook which culture has been spotted into each row on the plate!

6. Leave the plate right side up for ~30 minutes, to allow time for the yeast to settle and attach to the medium.

7. When the cells have settled, invert the plates and incubate them at 30°C. Plates are inverted to prevent water droplets that form on the inner surface of the lid from falling on the colonies. Plates can also be kept at room temperature, but cells will grow more slowly. Do NOT incubate the cells above 30°C, which stresses the yeast.



8. When the colonies are large enough to count, the plates will be removed from the incubator by the staff and placed in the refrigerator or cold room for your analysis later.

9. ***At the next class:*** Record your data with the scanner. To do this, remove the top from each plate and invert both the plate and its lid. Place the bottom half of the dish on the scanner and leave the inverted lid on the bench. (The lid is inverted to avoid contamination from spores and microorganisms that may be present in the air.) Place a black piece of cardboard or a folder over the plates before lowering the lid of the scanner.

10. Use spots where you can count individual colonies to calculate the density of cells in the original cell culture, correcting for the dilutions and the volume of the spot (see above).

Exercise 3 - Estimating cell densities with a spectrophotometer

The spectrophotometer provides a "quick and dirty" way to estimate the density of cells in a culture. In contrast to spot plates, which must be incubated for several days before colonies appear, spectrophotometer readings can be instantly converted into cell densities. On the other hand, the method does not discriminate between living and dead cells. The spectrophotometric method is based on light scattering by cells in the culture. Most visible light cannot penetrate a cell. When the light beam in a spectrophotometer hits a cell, the light is deflected from the light path, so some of the light never reaches the detector. The greater the number of cells in a sample, the more light scattering that occurs. The light scattering ability of a cell depends on its size and geometry, so a calibration curve is necessary to extrapolate optical density measurements to cell number. For example, the same number of yeast cells would scatter light more than the same number of bacterial cells, because the bacterial cells are much smaller.

Light scattering is measured with the spectrophotometer set to report absorbance. Because the principles used to measure light scattering and absorbance are different, the amount of light scattered by a solution is referred to as its "optical density" rather than its "absorbance." The optical density of a sample analyzed at 600 nm is abbreviated OD_{600}, with the subscript indicating the wavelength used for the measurement.

Estimating cell densities with the spectrophotometer
Follow the directions in Chapter 3 for operating the GeneSys 20.

1. Turn on the GeneSys 20 spectrophotometer. Adjust the wavelength of the monochromator to 600 nm.

2. Fill a cuvette with 1.0 mL deionized water to serve as the blank. Orient the cuvette in the holder so that the flat side of the cuvette faces the front of the instrument. Close the lid and press the "0 Abs/100%T" button to establish a baseline value for further measurements.

3. Replace the blank with a cuvette containing 1 mL of undiluted culture. Close the lid and read the OD_{600}. Record this value in your lab notebook. If the optical density of the sample is greater than 1.0, dilute the sample 1:10 with deionized water and read the optical density again. (The linear relationship between the OD_{600} and cell density is lost when OD_{600} values exceed 1.0) Record the new value in your lab notebook, noting how you diluted your sample. Dispose of all cell material in the white liquid waste container.

4. Repeat step 3 with each of your cultures.

5. Calculate an approximate cell density for each sample, assuming that an OD_{600} of 1.0 corresponds to approximately 1.3×10^7 cells/mL. Use only data where the OD_{600} is less than 1.0 for these calculations.

Discussion questions

How did cell densities of stationary phase cultures compare with those of log phase cultures? Propose an explanation for any differences that you noted.

An investigator spots two different strains of *S. cerevisiae* on a YPD plate. Both strains have the same number of colonies in each dilution, but the colonies from one strain are much smaller than those of the other strain. Propose an explanation for these results.

A yeast colony on a YPD plate contains millions of cells. Would you expect all of the cells in the colony to be in the same growth phase? Explain your reasoning.

References

Sherman F (2002) Getting started with yeast. *Meth Enzymol* **350:** 3-41.
Werner-Washburne M, Braun E, Johnston GC & Singer RA (1993) Stationary phase in the yeast *Saccharomyces cerevisiae*. *Microbiol Rev* **57:** 383-401.

Chapter 5

Introduction to Databases

The computer belongs on the benchtop in the modern biology lab, along with other essential equipment. A network of online databases provides researchers with quick access to information on genes, proteins, phenotypes, and publications. In this lab, you will collect information on a *MET* gene from a variety of databases.

Objectives

At the end of this lab, students should be able to:

- explain how information is submitted to and processed by biological databases.
- use NCBI databases to obtain information about specific genes and the proteins encoded by those genes.
- describe the computational processes used by genome projects to decode DNA sequences and process the information for molecular databases.
- use the *Saccharomyces* Genome Database to obtain information about role(s) that specific proteins play in yeast metabolism and the phenotypic consequences of disrupting the protein's function.

Chapter 5

Biomedical research has been transformed in the past 20 years by rapid advances in DNA sequencing technologies, robotics and computing capacity. These advances have ushered in an era of high throughput science, which is generating a huge amount of information about biological systems. This information explosion has spurred the development of bioinformatics, an interdisciplinary field that requires skills in mathematics, computer science and biology. Bioinformaticians develop computational tools for collecting, organizing and analyzing a wide variety of biological data that are stored in a variety of searchable databases. Today's biologist needs to be familiar with online databases and to be proficient at searching databases for information.

Your team has been given three *S. cerevisiae* strains, each of which is missing a *MET* gene. In the next few labs, you will determine which *MET* gene is deleted in each strain. To identify the strains, you will need information about the *MET* gene sequences that have been disrupted in the mutant strains. You will also need information about the role that each of the encoded proteins plays in methionine synthesis. In this lab, you will search for that gene-specific information in several online databases. Each member of the group should research a different one of the three *MET* genes. As you progress through this lab, you may feel like you are going in circles at times, because you are! The records in databases are extensively hyperlinked to one another, so you will often find the same record via multiple paths. As you work through this chapter, we recommend that you record your search results directly into this lab manual.

Databases organize information

Databases are organized collections of information. The information is contained in individual records, each of which is assigned a unique accession number. Records in a database contain a number of fields that can be used to search the database. For a simple example, consider a class roster. Students in a class roster are identified by a unique ID number assigned by the college, which serves as the equivalent of an accession number. Class rosters contain a variety of fields, such as the student names, majors, graduation year and email addresses. Thus, class instructors are able to quickly search the rosters for students with a particular graduation year or major. (The class roster is actually a derivative database, because it draws on information from the much larger student information database maintained by the college.)

Information on genes and proteins is organized into multiple databases that vary widely in their size and focus. Many of the largest database collections receive support from governments, because of their importance to biomedical research. By far, the largest collection of databases is housed at the National Center for Biotechnology Information (NCBI) in the United States. NCBI includes literature, nucleotide, protein and structure databases, as well as powerful tools for analyzing data. The NCBI is part of an international network of databases, which includes smaller counterparts in Europe and Japan. Information can enter the network through any of these three portals, which exchange information daily.

It is important to keep in mind that information in databases is not static! Scientists make mistakes and technology continues to improve. It is not uncommon to find changes in a database record. Scientists with an interest in a particular gene are well-advised to check frequently for updates!

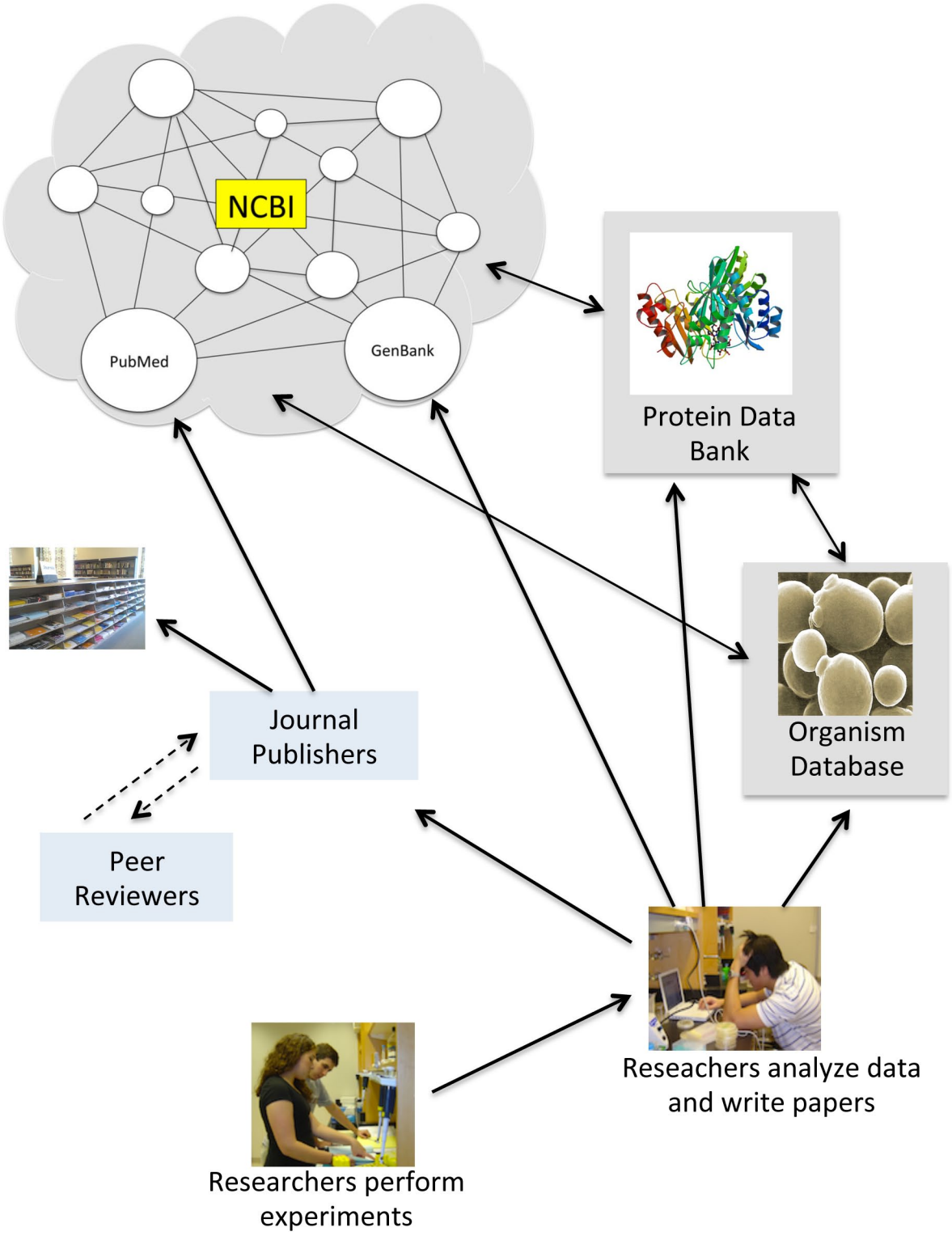
From the research bench to the database

The ultimate source of information in databases is the research community, which submits their experimental data to primary databases. Primary databases ask investigators for basic information about their submission. A record that meets the standards of the database is accepted and assigned a unique accession number that will remain permanently associated with the record. Each database has its own system of accession numbers, making it possible to identify the source of a record from its accession number. Once a record is accepted into a primary database, professional curators take over. Curators are professional scientists who add value to a record by providing links between records in different databases. Curators also organize the information in novel ways to generate derivative databases. Derivative databases, such as organism databases, are often designed to fit the needs of particular research communities. The *Saccharomyces* genome database (www.yeastgenome.org), for example, links information about genes to information about the proteins encoded by the genes and genetic experiments that explore gene function. In this course, we will be using both primary and derivative databases.

The figure on the following page summarizes information flow from the bench to databases. The information in databases originates in experiments. When researchers complete an experiment, they analyze their data and compile the results for communication to the research community. These communications may take several forms.

PubMed indexes publications in the biomedical sciences

Researchers will usually write a paper for publication in a scientific journal. Reviewers at the journal judge whether the results are accurate and represent a novel finding that will advance the field. These peer-reviewed papers are accepted by the journal, which then publishes the results in print and/or online form. As part of the publication process, biomedical journals automatically submit the article citation and abstract to PubMed, a literature database maintained by NCBI. PubMed entries are assigned a PMID accession number. PMID numbers are assigned sequentially and the numbers have grown quite large. PubMed currently contains over 23 million records! PubMed users can restrict their searches to fields such as title, author, journal, publication year, reviews, and more. The usability of PubMed continues to grow. Users are able to paste citations on a clipboard, save their searches, and arrange for RSS feeds when new search results enter PubMed. Students in the biomedical sciences need to become proficient in using PubMed. You can access PubMed at pubmed.gov or through the BC Library's database portal. An advantage of using the library's portal is that you will be able to use the library's powerful "Find It" button to access the actual articles.



Information flow from experiments to databases. Researchers analyze their data and prepare manuscripts for publication. Journal citations are submitted automatically to PubMed. Researchers also submit data to more specialized, interconnected databases.

Investigators submit experimental data to specialized research databases

Depending on the experiment, researchers will submit their data to a number of different databases. Consider the hypothetical example of a researcher who has isolated a novel variant of a *MET* gene from a wild strain of *S. cerevisiae* with a sophisticated genetic screen. The researcher has sequenced the gene, cloned the gene into a bacterial overexpression plasmid, and crystallized the overexpressed protein, which possesses unique regulatory properties. The researcher is preparing a manuscript on the experiments. In preparation for the manuscript submission (reviewers of the manuscript will want to see the accession numbers), the researcher plans to submit data to three different databases: a nucleotide database, a structure database and an organism database.

If our researcher is working at an institution in the U.S., he or she will probably submit the nucleotide sequence to NCBI's GenBank, a subdivision of the larger Nucleotide database. GenBank was founded in 1982, when DNA sequencing methods had just been developed and individual investigators were manually sequencing one gene at a time. The rate of GenBank submissions has increased in pace with advanced in DNA sequencing technologies. Today, GenBank accepts computationally generated submissions from large sequencing projects as well as submissions from individual investigators. GenBank currently contains over 300 million sequence records, including whole genomes, individual genes, transcripts, plasmids, and more. Not surprisingly, there is considerable redundancy in GenBank records. To eliminate this redundancy, NCBI curators constructed the derivative RefSeq database. RefSeq considers the whole genome sequences produced by sequencing projects to be the reference sequences for an organism. RefSeq currently contains nonredundant records for genome, transcripts and protein sequences from over 36,000 organisms.

The researcher in our hypothetical example will also want to submit the atomic coordinates and structural models for the crystallized protein to the Protein Data Bank (PDB). The PDB is part of an international consortium that accepts data for protein and nucleic acids. The vast majority of PDB records have been obtained by X-ray diffraction, although the database also accepts models obtained with nuclear magnetic resonance (NMR), electron microscopy, and other techniques. The number of entries in the PDB databases is orders of magnitude smaller than the number of predicted proteins in GenBank, reflecting the difficulties inherent in determining structures of macromolecules. PDB offers tools for visualizing macromolecules in three dimensions, allowing investigators to probe amino acid interactions that are important to protein function.

Finally, our researcher will want to submit data about the new mutant's phenotype and information about its regulation to the *Saccharomyces* Genome Database (SGD). The SGD serves as a central resource for the *S. cerevisiae* research community - which now includes you. The SGD is only one of many organism-specific databases. Similar databases exist for other model organisms such as the fruit fly *Drosophila*, the plant *Arabidopsis thaliana*, zebrafish and more. In addition to providing information, these specialized databases also facilitate research by providing links to important resources such as strain collections and plasmids.

Chapter 5

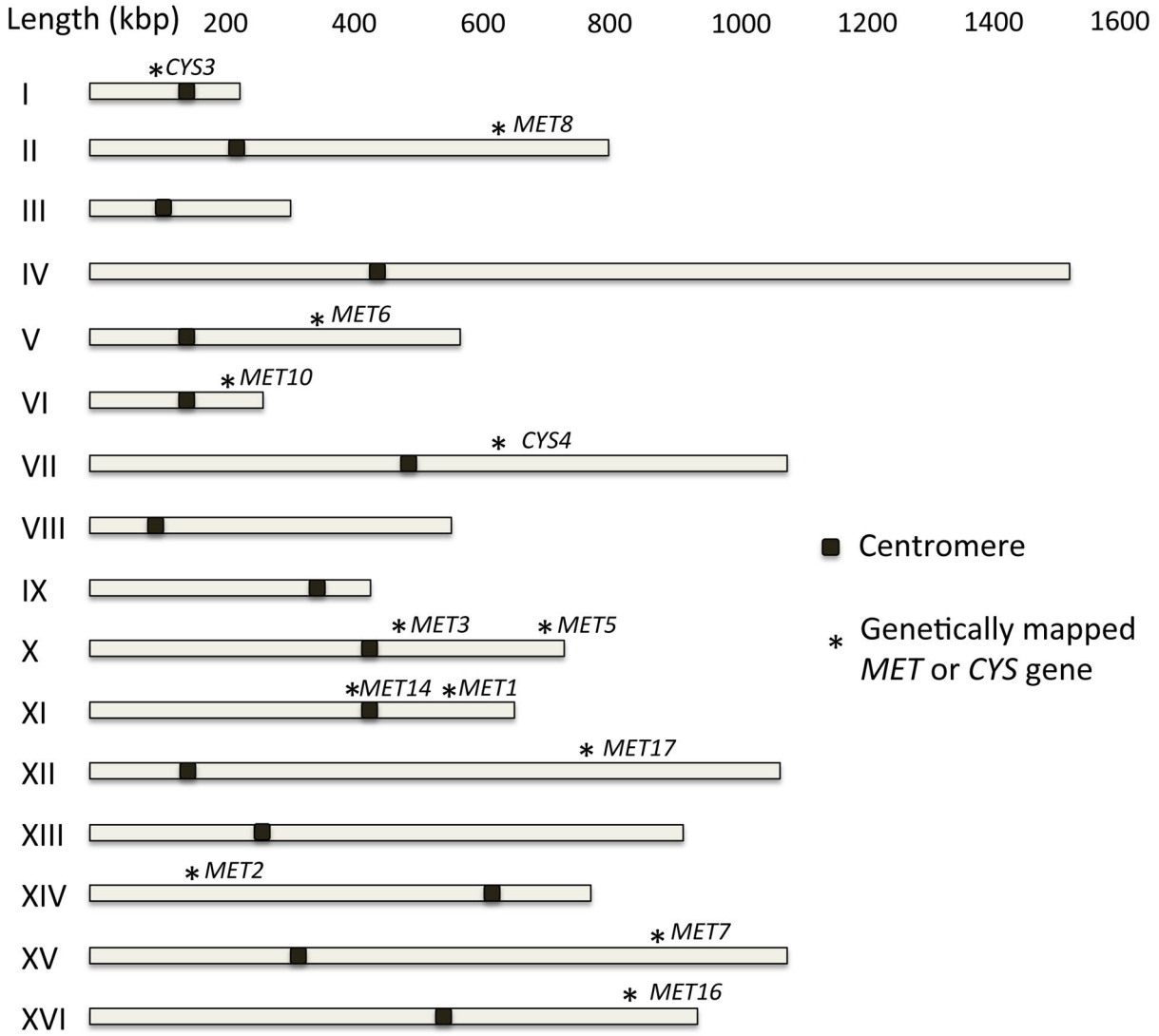
Saccharomyces genome project provided the reference sequence

The completion of the *S. cerevisiae* genome project (Goffeau *et al.*, 1996) represented a milestone in yeast genetics. *S. cerevisiae* had been an important genetic model for over 50 years, but associating genes with phenotypes was a slow process. In classical genetics, researchers generate collections of mutants and then map the genes responsible for mutant phenotypes by monitoring the segregation of traits during meiosis. Traits that are inherited together more than 50% of the time are assigned to the same linkage group, because they are located on the same chromosome. (Recall Mendel's law of independent assortment.) The more frequently that two traits are inherited together, the closer they are on a chromosome and the least likely to be separated by recombination during meiosis.

Prior to the genome project, yeast geneticists had identified hundreds of linkage groups, which were gradually assembled into genetic maps of 16 chromosomes. The genetic maps contained approximately 1000 known genes, including several genes involved in Met biosynthesis. Over several decades, yeast geneticists had isolated over a hundred mutants that were unable to synthesize Met and these mutants had been placed into 21 complementation groups, the functional equivalents of genes (Masselot & DeRobichon-Szulmajster, 1975). However, the chromosomal locations of most *MET* genes was unknown. The figure on the right shows the positions of *MET* and *CYS* genes that were mapped to yeast chromosomes by classical genetic methods (Cherry *et al.*, 1997). By the time that the genome project began, researchers were also using recombinant DNA technology to identify genes that were deficient in mutant strains, so partial sequence information was available for many chromosomal regions. This sequence information proved to be invaluable in the interpretation of the genome project data.

The *S. cerevisiae* genome was the first eukaryotic genome to be decoded. The success of the *S. cerevisiae* genome project can be attributed to the impressive amount of collaboration within the yeast research community. Over 600 researchers in 92 laboratories contributed sequence data that was compiled to generate a highly accurate genome sequence of *S. cerevisiae* strain 288C (Goffeau *et al.*, 1996). A single yeast strain was chosen for DNA sequencing, because *S. cerevisiae* naturally accumulates mutations and laboratory strains can begin to diverge from one another as they are propagated in the lab (Mortimer, 2000). The deletion strains that we are using in this class (Winzeler *et al.*, 1999) are derived from strain 288C.

The ~12 million base pair (Mbp) DNA sequence provides the definitive physical map of the 16 yeast chromosomes. Computational analysis of the sequence predicted ~6000 open reading frames (ORFs), each representing a potential gene. The number of ORFs was considerably greater than the number of genes that had been previously mapped with genetic methods. Many ORFs were identified by their similarity to genes that had been studied in other organisms, while close to half of the ORFs were completely novel. (Over time, additional ORFs have been identified. Today, the number of dubious or uncharacterized *S. cerevisiae* ORFs is close to 1500.) The *S. cerevisiae* genome sequence generally confirmed the gene order predicted by the earlier genetic maps, but provided more accurate spacing for the distances separating individual yeast genes.



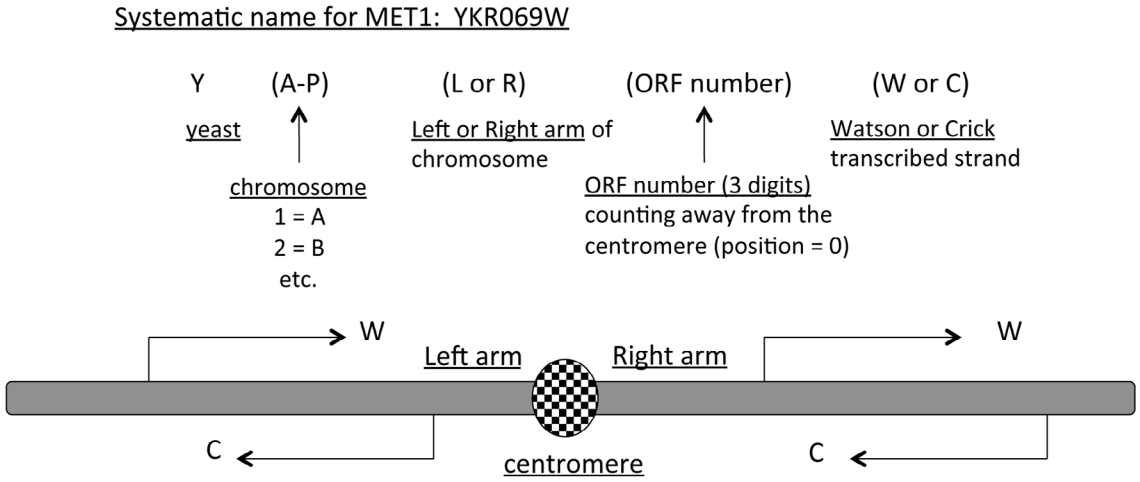
Chromosome map of the *S. cerevisiae* genome.
S. cerevisiae has 16 chromosomes that were originally identified by genetic linkage and subsequently confirmed by DNA sequencing. Chromosome numbers were assigned in the order that they were identified by classical linkage analysis.

Test yourself:
(*First, read the coordinate information on the following page.*)

The *S. cerevisiae* genome contains two genes, *SAM1* and *SAM2*, encoding enzymes that catalyze the conversion of Met to the high energy methyl donor, S-adenosylmethionine. The two genes arose from a gene duplication and remain almost identical to one another. The systematic name for *SAM1* is YLR180W, and the systematic name for *SAM2* is YDR502C. Use the coordinate information below to determine the chromosomal locations of *SAM1* and *SAM2*. Place the two genes on the map above. Draw arrows that indicate the direction of transcription for both genes.

Chapter 5

The genome project data provided the organizing structure for the *Saccharomyces* Genome Database (SGD). The SGD systematically assigned accession numbers to ORFs, based on their location and orientation on yeast chromosomes. The systematic name for each ORF has 7 characters. Each begins with a "Y" for yeast, followed by letters depicting the chromosome number and chromosome arm, followed by a 3-digit ORF number counting away from the centromere. The last letter in the locus name indicates if transcription occurs on the Watson or Crick strand of the DNA.



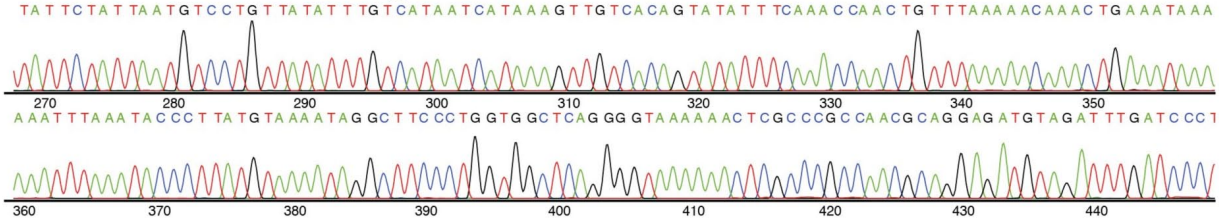
The figure on the opposite page outlines the process used by the genome project to decode and annotate the *S. cerevisiae* sequence. The complete sequences of the 16 yeast chromosomes laid end-to-end are considered the reference genome for *S. cerevisiae*. The genome sequence was submitted to NCBI's GenBank, where curators assigned an NC____ accession number to each of the 16 chromosome sequences, indicating that the sequences are non-redundant chromosome sequences. Potential protein-coding sequences were identified with an ORF-finding algorithm that looks for sequences that begin with an ATG initiation codon and terminate with a stop codon in the same reading frame. ORF finding programs rely on the fact that stop codons are underrepresented in protein coding sequences. Because 3 of the total 64 codons are stop codons, one would predict a stop codon to randomly occur about once in every 21 amino acids in a protein sequence. Most proteins, however, contain 100 amino acids or more. ORF-finders are also able to identify and exclude introns from the ORF. Each potential ORF identified in the project was assigned an NM_____ accession number, consistent with a transcript sequence, or potential mRNA.

Computational methods were used to predict the amino acid sequences of the proteins encoded by the transcripts, and the translated sequences were assigned NP_____ accession numbers. (In fact, the vast majority of protein sequences in NCBI's Protein database have been derived by automated translation of DNA sequences, because chemical sequencing of proteins is much more laborious task than DNA sequencing.) The functions of most proteins predicted by the genome project have still not been experimentally validated. **Your experiments this semester will contribute some of the missing experimental validation, when you transform *met* deletion mutants with plasmids carrying *MET* genes.**

Saccharomyces Genome Project Data Flow

1. DNA Sequencing

Labs were assigned individual chromosomes to sequence



Thousands of sequencing reactions were performed

2. Alignment

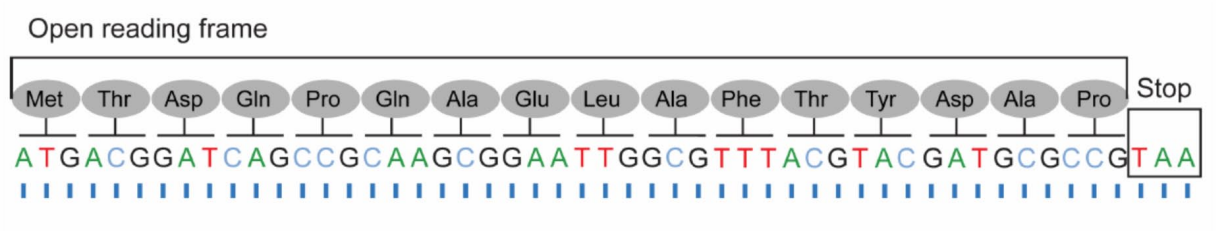
Sequences were aligned to generate complete chromosome sequences

Computer algorithms identify areas of sequence overlap

Process is repeated to align long stretches of sequence

3. Processing and annotation

Chromosome sequences were submitted to NCBI Nucleotide database
Sequences were assigned NC_____ accession numbers



Open reading frames (ORFs) were identified by computational methods

- ORFs begin with an initiation codon for Met
- Methionine is separated from a stop codon in same reading frame by a long stretch of predicted amino acids

ORFs were given NM_____ accession numbers

Automated translation programs predicted amino acid sequences
Sequences were compared to homologs in other species to predict function
Protein sequences were given NP_____ accession numbers

Chapter 5

Exercise 1: Finding gene records in NCBI databases

Homepage: Point your browser to the NCBI homepage: ***ncbi.nlm.nih.gov***
NCBI maintains a large collection of databases. Clicking on the dropdown box brings up the list of databases for more targeted searching. For a comprehensive search, use the "All databases" setting. **Write the name of your *MET* gene in the search box and click "Search."**

Entrez summary page:
The Entrez page summarizes the number of hits in each of the many NCBI databases. The number is probably quite large! Take a look at the results. Note below the number of records in the PubMed, Nucleotide, Protein, and Structure databases.

- **Modify the search term by adding "AND Saccharomyces cerevisiae" to the search box.** Compare the number of records in each category to the numbers retrieved in the unrestricted search (above). The numbers have probably dropped significantly! Why do you think that this happened? This simple comparison may give you some idea of the sheer volume of records in the NCBI databases. You may not receive any hits in the Structure category, since the vast majority of proteins have not been crystallized or studied with NMR.

NCBI Nucleotide:

- Click on the first number under the Nucleotide Sequences grouping, which brings you to your search results in NCBI's Nucleotide database. The Nucleotide database aggregates records from multiple databases, including GenBank and the reference sequence database, RefSeq. Do all the records in your search results refer to *S. cerevisiae* sequences? *Probably not!* Any mention of "Saccharomyces cerevisiae" in a record is enough to bring it up in an unrestricted search, even if the sequence comes from a different organism.
- Narrow down the search to records that actually contain *S. cerevisiae* sequences by clicking the *Saccharomyces cerevisiae* link in the Top Organism list on the right. Note that clicking on the tree adds an additional search term. You will now see RefSeq records in the new list.
- Use the hyperlinks at the right to filter your results to the RefSeq database. You should see a single NC_ and a single NM_ record in the list. (Remember the results are non-redundant!
- Let's look at the NC___ record first.

Record the accession number _____

Which chromosome is represented in the record? _____

How many nucleotides are in the chromosome (bp)? _____

- Click the **GenBank link** associated with the NC_____ *(fill in the numbers)* record. Near the top, you will see links to articles in the primary literature, which will include both the Goffeau et al. (1996) report on the genome project, as well as a more detailed article on the chromosome by the investigators who determined its sequence.

- Scroll down a bit in this very long record and to the FEATURES field and look at a few genes. You will note that the first feature is a telomere, because the sequence begins at the end of the chromosome's left arm. As you scroll down, you are moving from one end of the chromosome to the other, and you will see annotation information for the ORFs identified by the SGP. Each ORF has a description of its gene, mRNA, and coding sequence (CDS).

- Find your *MET* gene in the NC record by typing its name into the "Find" search box of your browser.

  Record the 7-character locus tag (begins with Y) _____
  This is the systematic name assigned to the ORF by the Saccharomyces Genome Project. This serves as the gene's unique identifier in the Saccharomyces Genome Database.

  Is there an intron in your gene? _____
  (Introns will be manifest by interruptions in nucleotide numbers and the word "join" in the mRNA field for your gene.)

- Find the transcript record by clicking on the link to the NM_____ *(fill in the accession number)* record.
  How many nucleotides are in the annotated gene sequence (bp)? _____

  Cursor down to the actual nucleotide sequence at the end of the record. Note that the annotated *S. cerevisiae* ORFs begin with an ATG and end with a stop codon.
  Is the NM_ record the actual sequence of the mRNA for your gene? Why or why not?

- Find the field for the protein coding sequence (CDS) in the NM transcript record.

  The CDS field contains a translation of the the NM nucleotide sequence. Find the NP_____ *(fill in the numbers)* record. Click on the NP record.

- The NP record will give you additional information about the protein, including links to information about its structure, conserved domains and homologs in other organisms. Refer to the right panel on the page. If a 3-dimensional structure is available for your protein, you will see a 4-character PDB accession number under "Protein 3D Structure." Record the accession number, if it is available. (If you would like to see the structure, you can search the Protein Data Bank at www.pdb.org with either this accession number of your gene name.)

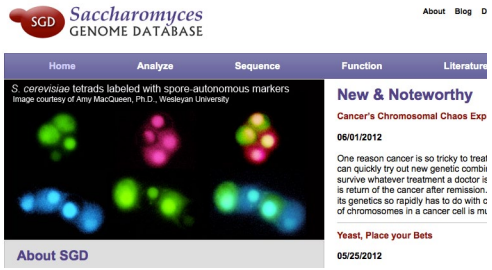
Chapter 5

Exercise 2: Using the Saccharomyces Genome Database

The Saccharomyces Genome Database (SGD) is an important resource for the yeast research community. The SGD is a heavily curated database that draws from many other databases to provide a comprehensive view of a gene and its protein product, regulation of the gene's expression and the role of the protein in cell function. It also serves as a virtual meeting place that welcomes contributions from the community. This exercise will introduce you to some useful resources at the SGD.

Direct your browser to:
http://www.yeastgenome.org

For mobile devices, a Yeast Genome app is available through iTunes



- Type the name of your *MET* gene in the search box. This brings up the summary page for your gene. Note that the systematic name for your gene is identical to the locus name assigned by the genome project.

  Record the standard name for your gene _____
  Are there multiple genes for your name? *What are they?*

  The standard names for many *S. cerevisiae* genes were originally assigned during classic genetic screens. A screen for mutants unable to synthesize Met, for example, might generate a collection of mutant strains that would arbitrarily assigned numbers - *met1*, *met2*..... *metN*. Occasionally, the same gene might appear in screens for different phenotypes, in which case there may be both two different names for a gene. For example, *MET5* is the same as *ECM17*, because it was isolated in a screen for defects in the cell wall. (ECM stands for extracellular matrix.)

What role does the product of your MET gene play in metabolism?

The *MET* genes that we are studying all encode enzymes involved in Met synthesis. These enzymes are parts of highly regulated pathways in cells. Find the **"Pathways"** field in the sidebar on the SGD summary page for your gene. You will see links on the right to one or more pathways that have been identified in yeast by MetaCyc, a nonredundant database of metabolic pathways that are supported by experimental evidence.

- List the MetaCyc pathway(s) that your enzyme is involved in. If the enzyme is involved in more than one pathway, are the pathways related to one another? How?

- Click on one of the pathway links. Locate the position of your gene product in the pathway. What is the official name of the enzyme encoded by your *MET* gene?

  Under the name of your enzyme, you will see a number with 3 decimal points. This is the official classification given to the enzyme by the Enzyme Commission, which categorizes enzymes in very fine detail. The first number indicates the broad class of enzyme, e.g. hydrolase, transferase, oxidoreductase. The subsequent numbers drill down to the kinds of bonds altered in the reaction and finally to specific substrates. Enzymes from different organisms with the same EC number are *expected* to catalyze the same reaction. Record the E.C. number. *This E.C. number will be important later in the semester, when you evaluate whether homologs from other species are likely to catalyze the same reaction as the S. cerevisiae enzyme.*

- Click on the E.C. number to see the reaction catalyzed by the enzyme.

  What are the substrates and products for your enzyme? Draw the structure of the substrate and product that are intermediates in methionine synthesis.

- The position of your enzyme in the pathway will be important in the next experiment, where selective plating is used to identify your group's *met* mutants.

  What gene encodes the enzyme that catalyzes the PRECEDING step in the pathway? This step generates the substrate for the reaction that you are studying.

  What gene encodes the enzyme that catalyzes the NEXT step in the pathway? The product of your reaction will be metabolized in this step.

Chapter 5

Gene expression

The expression of a gene in response to environmental stimuli often provides clues to its physiological importance. The absolute level of gene expression may change in an experiment, and it is also important to see which genes show similar changes in expression, because networks of genes frequently work together to affect the phenotype. Microarray technology offers a tool for researchers to study the simultaneous expression of hundreds of genes. (If you are unfamiliar with microarrays, you may find the following tutorial to be helpful: http://learn.genetics.utah.edu/content/labs/microarray/)

- Click on the Expression tab at the top of your gene's summary page. This brings you to a histogram showing the number of experiments where the level of gene expression in log2 units (0 indicates no change, 1 indicates a 2-fold change, 2 a 4-fold change, etc.) changed in microarray experiments. For most genes, the level of gene expression probably did not change, because the most common value (statistical mode) is zero. Expression levels probably increased in some experiments and decreased in others.

- For more useful and detailed information, click on the SPELL (Serial Pattern of Expression Levels Locator) link at the top of the page. This will bring you to a summary of hundreds of datasets from high-throughput experiments that compare gene expression under a variety of environmental conditions. Cursor down to the list of genes that are most frequently co-expressed with your gene and note their ACS (Adjusted Correlation Score) scores. The ACS score is calculated from the entire range of SPELL datasets. Genes with the highest ACS scores show expression levels that are most closely correlated to your gene-of-interest.

  List below the six genes that were most frequently co-expressed with your gene. Look up the reactions catalyzed by their gene products. Do you notice any patterns in the genes that are coordinately expressed with your *MET* gene? Consult your Pathways data above. Do any other genes in the Met synthetic pathway appear in the list?

- SPELL gives a graphical output of the expression patterns for individual experiments. Green labels indicate increased expression, and red labels indicate reduced expression. Scroll right to see the results of individual experiments (in no particular order) together with their literature citation. A quick visual inspection attests to the co-expression of genes with high ACS scores.

Phenotype information

In keeping with its mission to serve as a resource to the yeast research community, which includes many geneticists, the SGD also collects information on mutant strains and their phenotypes. The deletion strains that we will be using in our experiments are unable to grow in the absence of methionine, but they may demonstrate additional phenotypes.

- Return to the summary page for your gene. Find the field in the sidebar for mutant phenotypes. Focus on the phenotypes that have been identified using **classical genetics**.

- What phenotypes are detected in null mutants (null mutants either lack the gene or produce an inactive protein)? (***Note:*** *you can use the tab at the top of the page to obtain more detailed information about the phenotypes, as well as literature references.*)

- Mutations in some *MET* genes have phenotypes associated with tellurium accumulation and resistance (Ottoson *et al.*, 2010). Find tellurium on the periodic chart. How do you think tellurium is working?

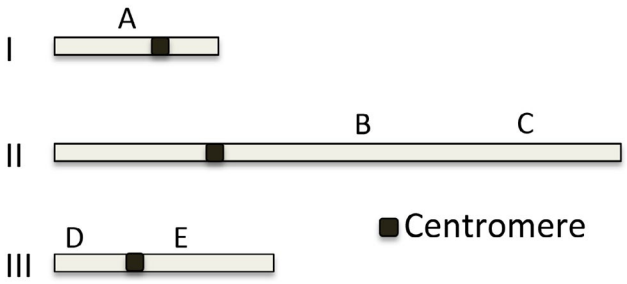
Test yourself

1. *S. cerevisiae* has 16 chromosomes that contain the sequences of ~6000 genes. If an investigator wanted to obtain the complete sequences for all *S. cerevisiae* genes from NCBI, what is the minimum number of records that he/should could obtain that would contain all the information?

   What common feature would be found in all of the accession numbers?

2. How do the *S. cerevisiae* transcript sequences in NM_____ records differ from the actual sequences of mRNAs?

3. Each *S. cerevisiae* gene has a locus name, which is based on its chromosomal location. The figure below shows the positions of yeast ORFs (A-E), which are located on chromosomes I-III. Beside each pair of locus names, write the letter that corresponds to its chromosomal location.



YBR105C _____

YAL026C _____

YCR036W _____

YCL050W _____

YBR213W _____

4. The light chain of cytoplasmic dynein, a microtubule motor protein, is encoded by the *S. cerevisiae DYN2* gene. Find the chromosome record with the *DYN2* gene at NCBI.

   What is the locus tag (systematic name) for the gene?

   Unlike most *S. cerevisiae* genes, the *DYN2* gene contains introns. (You may want to open the *DYN2* record in the SGD and consult the Chromosomal location field for a graphical view.) What is the total length of the transcript in base pairs?

   How many exons are present in the gene?

   How long are the exons (bp)?

   How many amino acids are found in the Dyn2 protein?

References

Cherry JM, Ball C, Weng S *et al.* (1997) Genetic and physical maps of *Saccharomyces cerevisiae*. *Nature* **387**: 67-73S.

Goffeau A, Barrell BG, Bussey H *et al.* (1996) Life with 6000 genes. *Science* **274:** 563-567.

Masselot M & DeRobichon-Szulmajster H (1975) Methionine biosynthesis in *Saccharomyces cerevisiae*. I. Genetical analysis of auxotrophic mutants. *Mol Gen Genet* **139:** 121-132.

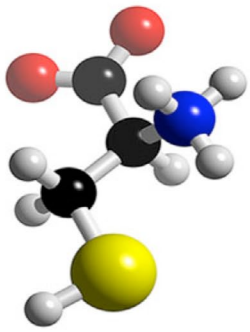
Mortimer RK (2000) Evolution and variation of the yeast (*Saccharomyces*) genome. *Genome Res* **10:** 403-409.

Ottoson LG, Logg K, Ibstedt S, Sunnerhagen P *et al.* (2010) Sulfate assimilation mediates tellurite reduction and toxicity in *Saccharomyces cerevisiae*. *Eukaryot Cell* **9:** 1635-1647.

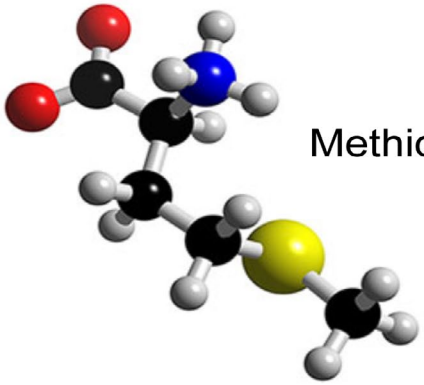
Winzeler EA, Shoemaker DD, Astromoff A *et al.* (1999) Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. *Science* **285:** 901-906.

Chapter 6

Analysis of mutant strains



Cysteine



Methionine

The growth properties of mutant strains can often provide information about the gene products involved in biochemical pathways within cells. The sulfur containing amino acids, Met and Cys, are synthesized in a multi-step pathway in yeast. In this lab, you will use selective and differential media to determine which *MET* genes have been inactivated in *S. cerevisiae met* deletion strains.

Objectives

At the end of this lab, students should be able to:

- use correct genetic nomenclature for genes, proteins and mutants in written reports.
- explain how genetic screens are used to isolate mutants with particular phenotypes.
- distinguish *met* strains by their ability to grow on selective media containing various sulfur sources.
- predict how mutations in the genes involved in Met and Cys synthesis will affect the concentrations of metabolites in the pathway.

Chapter 6

Cells require sulfur-containing amino acids

Mutant organisms provide powerful tools to study biochemical pathways in living cells. This semester, we are working with yeast strains that are unable to synthesize methionine (Met) or cysteine (Cys) because one of the genes involved in the biosynthetic pathway has been inactivated. Met and Cys are essential amino acids for all organisms. The sulfur atoms in their side chains impart distinctive chemistries to Met and Cys, which has important implications for protein function. Unlike us, wild type yeast are able to synthesize both Met and Cys, using only inorganic sulfate as a sulfur source. Each of the *met* mutant strains that we are using this semester is missing a single *MET* gene, and this deletion prevents the strain from growing on media containing only sulfate as the sulfur source. The deleted *MET* genes in the strains have been replaced with a bacterial kanamycin resistance (KAN^R) gene by homologous recombination (Winzeler *et al.*, 1999). Depending on the exact *met* mutation, the strains may be able to synthesize Met and Cys from other sulfur sources that they transport into the cell and convert to Met or Cys. In this lab, you will use selective media containing various sulfur sources and differential media to distinguish between three *met* mutants. In the next lab, you will use the polymerase chain reaction (PCR) to more conclusively identify the mutant *met* strains.

Genetic nomenclature

When referring to genes and strains, it is important to use correct genetic nomenclature. Pay close attention to italics and capital letters as your prepare your reports. Gene names are placed in italics, while proteins and phenotypes are referred to with normal font. Gene names that begin with capital letters refer to dominant alleles, while gene names beginning with lower case letters refer to recessive alleles. (One oddity about budding yeast: *S. cerevisiae* gene names are unique in that dominant alleles are described with three capital letters. In many other eukaryotic species, dominant alleles would be referred to as *Met6* with only the first letter capitalized.) *S. cerevisiae* gene names consist of three letters, followed by a number. There may be many different gene names that begin with the same three letters, e.g. there are over 20 different *MET* genes, but the number at the end of the gene name is specific for a particular gene. If some molecular detail is available for a particular mutant allele, the number may be followed by a hyphen and additional information about the allele.

As an example, let's look at the nomenclature that would be used for the *MET6* gene from *S. cerevisiae*. The *met* prefix is used to describe loss-of-function alleles found in mutant strains, most of which were isolated in genetic screens based on their inability live in the absence of Met. The numbers associated with genes are usually arbitrary. The *MET6* gene acquired its name before its gene product had been identified as homocysteine methyltransferase, the last step in methionine synthesis. The list below describes the naming conventions for genes, proteins, and strains related to *MET6*. These same rules apply for other genes in *S. cerevisiae* as well.

MET6	Dominant allele of the *MET6* gene or the chromosomal locus
met6	Recessive allele of the *MET6* gene (allele found in a *met6* mutant)
met6-12	Recessive allele - number after the parentheses refers to specific mutation
met6-Δ1	Recessive allele - *met6* allele has a specific deletion (*Δ* indicates a deletion)
met6::LEU2	Recessive allele -insertion of a dominant *LEU2* gene into the *MET6* locus on the chromosome has inactivated the host *MET6* gene
Met6p	Protein encoded by the *MET6* gene, *i.e.* homocysteine methyltransferase

We will be working with haploid strains of yeast in this course. To write the genotype of a particular strain, begin with the mating type and follow it with the mutant alleles in the strain. For example, we are using *met* strains constructed by inserting a bacterial kanamycin resistance (KAN^R) gene into yeast strain BY4742, which has the α mating type and carries mutations in genes involved in the synthesis of histidine, leucine, lysine and uracil. BY4742 is derived from strain S288C, which was used for the genome project (Brachmann *et al.*, 1998). Thus, the genotype of our *met6* mutant would include the BY4742 mutations and be written:

MATα his3-Δ1 leu2Δ0 lys2Δ0 ura3Δ0 met6::KAN^R

Auxotrophs and selective media

The *met* mutants are Met auxotrophs, meaning that they are unable to grow in media that does not contain Met. Auxotrophs are microorganisms that are unable to synthesize an essential nutrient because of a gene mutation. Many laboratory strains carry multiple mutations that interfere with the synthesis of essential nutrients. For example, because the BY4742 strain carries mutations in the *HIS3*, *LEU2*, *LYS2* and *URA3* genes, the strain will only grow in media containing histidine, leucine, lysine and uracil. Auxotrophic strains have many uses in genetics. Researchers often use auxotrophic strains as hosts for plasmid transformation (Chapter 12). The plasmids used for transformation carry functional alleles of a gene that is defective in the host strain, making it possible to select transformants by their ability to grow on media lacking the essential nutrient.

Synthetic media are an essential tool for culturing and studying auxotrophs, because all of the components are defined. Yeast researchers have developed a variety of different formulations for synthetic media. All synthetic media contain a carbon source (usually D-glucose), a nitrogen source, and essential vitamins and minerals. The vitamins and minerals are usually purchased in a formulation known as yeast nitrogen base (YNB). The supplements added to synthetic media can be tailored to support or select against the growth of particular genotypes. In this course, we will use Yeast Complete (YC) medium that supports the growth of most *S. cerevisiae* strains. The growth rate of wild type strains in YC is somewhat slower than that in rich media like YPD, but the strains are viable for long periods of time. The table on the following page shows the composition of YC, which includes a rich supply of amino acids and nucleotide bases. In addition to the complete YC medium, we will also use selective media in which some of components have been left out. For example, in this lab, we will use YC-Met "drop-out" media, which contains all of the YC components in the following table, except methionine.

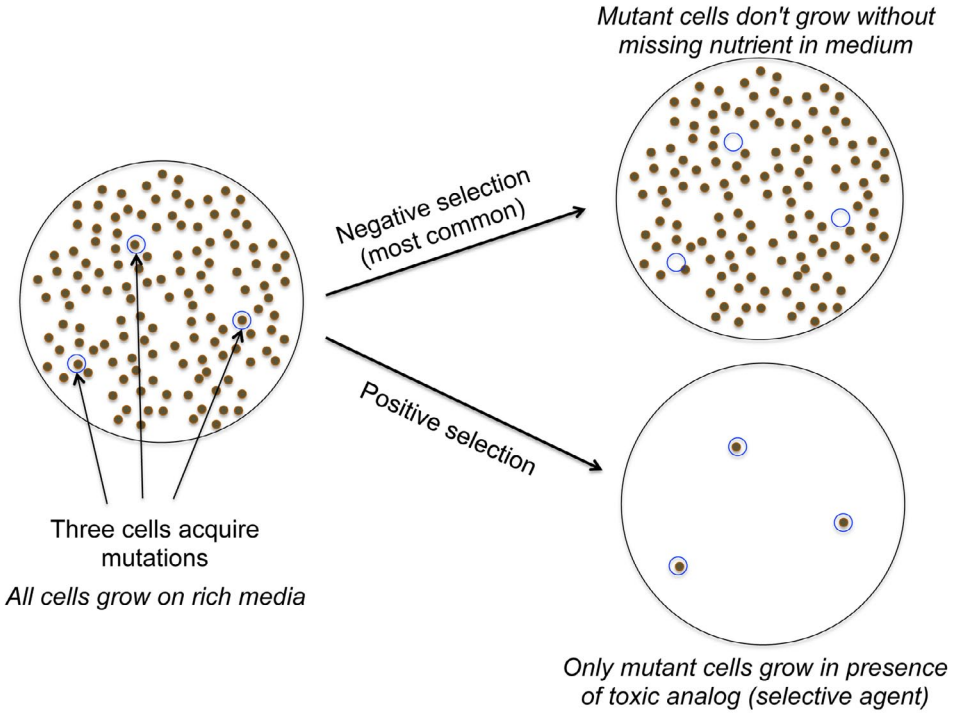
Composition of Yeast Complete (YC) Medium					
Component	**grams/liter**	**Component**	**mg/liter**	**Component**	**mg/liter**
YNB*	1.7	arginine	100	tyrosine	50
$(NH_4)_2SO_4$	5	aspartic acid	50	lysine	100
D-glucose	20	isoleucine	50	methionine	50
		phenylalanine	50	tryptophan	100
		proline	50	leucine	100
		serine	50	histidine	50
		threonine	100	uracil	10
		valine	50	adenine	10

YNB is a complex mixture of vitamins, minerals and salts. Final concentrations in YC:
Vitamins (µg/liter): biotin (2), calcium pantothenate (400), folic acid (2), inositol (2000), niacin (400), p-aminobenzoic acid (200), pyridoxine hydrochloride (400), riboflavin (200), thiamine hydrochloride (400).
Minerals (µg/liter): boric acid (500), copper sulfate (40), potassium iodide (100), ferric chloride (200), manganese sulfate (400), sodium molybdate (200), zinc sulfate (400).
Salts (mg/liter): potassium phosphate monobasic (1000), magnesium sulfate (500), sodium chloride (100), calcium chloride (100).
(Source: http://labs.fhcrc.org/gottschling/Yeast%20Protocols/yc.html)

Genetic analyses of methionine biosynthesis

Looking at the pathway for Met biosynthesis later in this chapter, you may wonder how the gene numbers became associated with specific genes, since the numbers do not correspond to the positions of the reactions encoded by the *MET* gene products in the pathway. The numbering system reflects the discovery process for the *MET* genes. The first studies of Met biosynthesis in yeast were done by geneticists, who used genetic screens to isolate *met* mutants. Genetic screens are important tools for identifying new genes because they are unbiased by prior knowledge of the pathway. In addition, mutation is a random process that should affect all genes involved in producing the phenotype under study. The geneticist begins by treating a parent strain with a chemical or radiation to induce mutations in DNA. The spontaneous mutation rate in yeast is ~10^{-8}/base/generation, which is much too low for a practical genetic screen. Investigators therefore use mutagen doses that kill up to ~50% of the cells. Mutagenesis is random and many mutations have no effect on the phenotype that is being screened. Consequently, large numbers of cells are required to uncover all the genes involved in the phenotype. For example, the yeast genome contains ~6000 genes, so a useful genetic screen might involve 20,000 or more cells.

Selective media provide important tools for identifying mutant phenotypes in genetic screens. Depending on the phenotype being studied, investigators may select for mutants using either a positive or negative selection scheme, as shown on the opposite page. The easiest kinds of screens employ positive selection, because only mutant cells grow on selective media. If investigators are analyzing pathways that are important for cell growth, such as Met synthesis, they would probably use a negative selection scheme. In a negative scheme, cells are first plated



Selection strategies used to isolate yeast mutants.
After the initial mutagenesis, yeast are grown on a plate containing rich (or complete synthetic) media. In this figure, the mutagenesis has generated three different mutants in the gene of interest. The mutant colonies are surrounded by an empty circle. Replicas of the master plate are copied to selective media. In a negative selection scheme, the selective plate lacks a component that is normally present in rich media. In a positive selection scheme, the media contains a selective agent, which is toxic to normal cells, but tolerated by mutant cells. The selective agent is sometimes a toxic analog of a normal cellular metabolite.

http://www.bc.edu/content/bc/offices/stsery/academic/students/uviewplus.html

on media, such as YPD or YC, that allow all cells to grow. Replicas of these master plates are then made on defined media lacking Met. (Replica plating is outlined in Chapter 12.) Since only wild-type cells grow on the selective media lacking Met, researchers look for colonies on the rich media whose counterparts are missing on the selective media.

The number and spectrum of mutants obtained in a genetic screen are unpredictable, because of the random nature of mutation. As you might expect, a screen might produce multiple mutants in one gene and no mutations in other genes involved in the phenotype. After completing a screen, investigators must next determine if the mutations are in the same or different genes. For this, geneticists rely on genetic mapping (Chapter 5) and/or complementation. Complementation is a functional test of gene activity. In a complementation experiment, introduction of a functional gene from another source rescues a mutant phenotype caused by the defective gene. Classic genetic complementation in yeast takes advantage of the two yeast mating types and the ability of yeast to survive as both haploid and diploid strains. In a complementation experiment with *met* mutants, researchers mate a haploid *met* mutant in either the α or **a** mating type (*MAT*α or *MAT***a**) with a haploid *met* mutant of the opposite mating type. If the diploid is able to grow in the absence of Met, complementation has occurred, and the *met* mutations in the two haploid strains must be in different genes. If the diploid is not able to survive on the selective plate, the two haploid strains carry mutations in the same gene (although they are

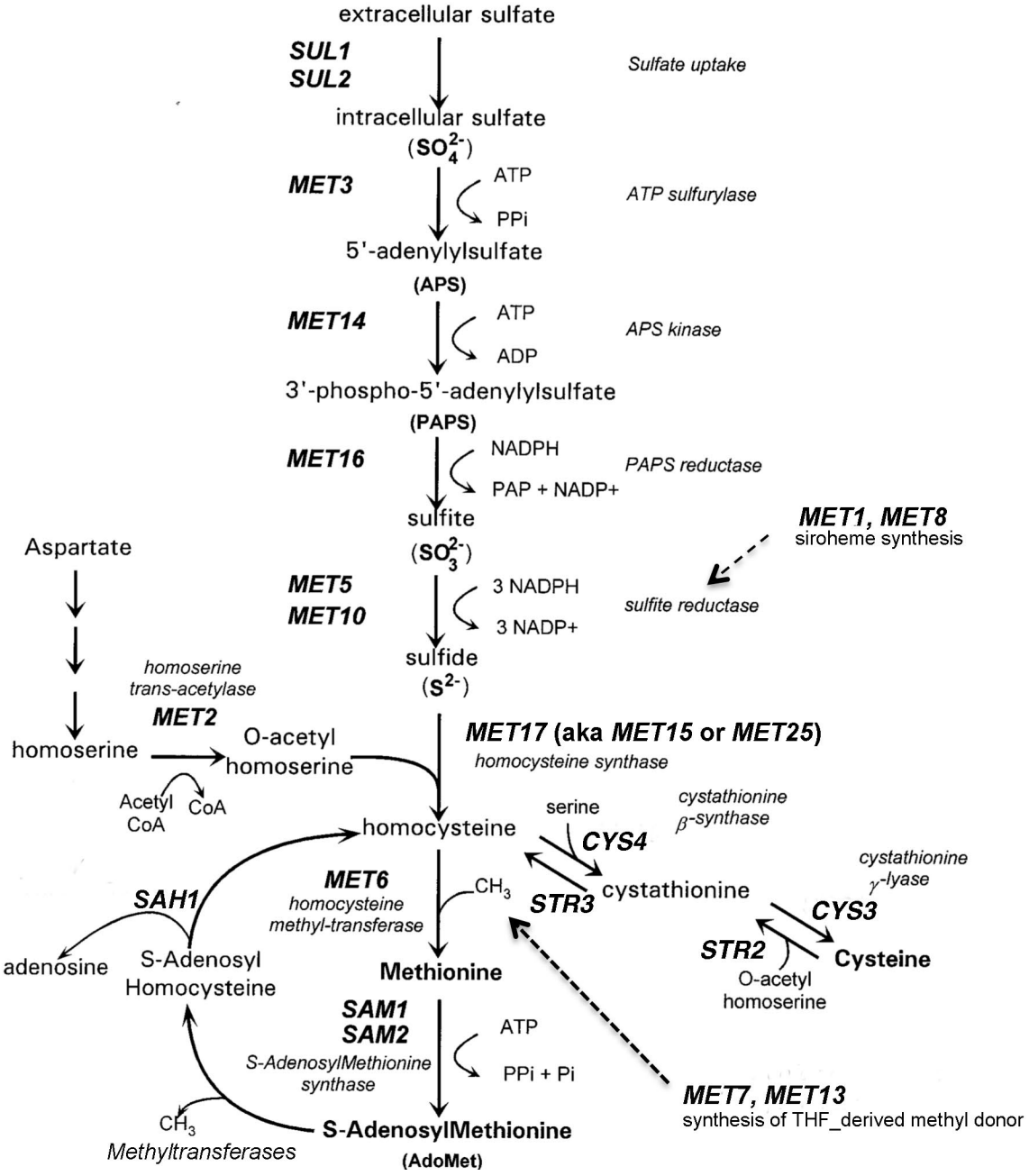
almost certain to be different mutant alleles). A genetic screen can yield multiple mutant alleles of the same gene, which together form a complementation group.

By 1975, yeast labs had isolated collections of *met* mutants and mapped nine of the *met* mutations to chromosomes. In a landmark study, Masselot and DeRobichon-Szulmajster (1975) collected 100 *met* strains from labs around the world and did systematic complementation experiments with all the mutants. Twenty-one complementation groups, representing potential genes, were identified, and the genes were assigned names *MET1* through *MET25*. Many of the *MET* genes encode enzymes in the Met biosynthetic pathway, which is outlined on the opposite page. Some gene products are involved in the synthesis of cofactors and methyl donors used in the pathway, while other *MET* gene products (not shown) are involved in regulation of the pathway (reviewed in Thomas & Surdin-Kerjan, 1992). For the most part, the names assigned in the 1975 study are still used today. A few genes identified in the 1975 study were subsequently shown not to be involved in Met biosynthesis, and others (e.g. *MET15*, *MET17* and *MET25*) were later shown to represent different alleles of the same gene (D'Andrea et al., 1987).

At the time of the 1975 study, the biochemical reactions in the pathway were largely known, and scientists faced the challenge of associating genes with enzymatic activities. You can see from the pathway that mutations in 11 different *MET* genes would produce a phenotype in which strains would grow in the presence of methionine, but not in its absence. The scientists narrowed down possible gene-enzyme relationships by analyzing the ability of *met* strains to use alternative sulfur sources in the place of methionine (Masselot & DeRobichon-Szulmajster, 1975). Yeast are very versatile in their use of both inorganic and organic sulfur sources. Sulfate is efficiently transported into cells by the Sul1p and Sul2p transporters in the membrane. Sulfite and sulfide are also transported into the cells with a reduced efficiency. Yeast are also able to transport and use Met, Cys, homocysteine and S-adenosylmethionine (AdoMet or SAM) as sulfur sources (reviewed in Thomas and Surdin-Kerjan, 1992). In this lab, you will use selective media in which sulfite or cysteine replaces methionine to distinguish between 3 *met* mutants. You will also use a differential medium, BiGGY agar, that distinguishes yeast strains by their production of hydrogen sulfide.

NOTE: The *met* mutants used in this course were **NOT** generated by traditional mutagenesis. Instead, the mutants were constructed by a newer molecular approach that requires detailed knowledge of the yeast genome sequence. After the yeast genome project was complete, researchers were interested in obtaining a genome-wide collection of deletion strains, each of which differed from the parental BY4742 strain at a single site. Their approach, which is discussed in more detail in Chapter 7, takes advantage of the high frequency with which *S. cerevisiae* undergoes homologous recombination. Each ORF in the *S. cerevisiae* genome was systematically replaced with a bacterial KAN^R gene (Winzeler et al., 1999). A major advantage of this strategy, sometimes referred to as "reverse genetics," over the traditional genetic approach is that positive selection can be used to isolate mutants. Only strains with disrupted *MET* genes are able to grow on media containing analogs of kanamycin. Strains with KAN^R-disrupted genes have other advantages over mutant strains generated with chemical mutagens or radiation treatment.

They do harbor secondary mutations caused by mutagen treatments and spontaneous reversion to a wild type phenotype is not possible.



Methionine biosynthesis in yeast.
The proteins catalyzing individual steps in Met and Cys biosynthesis are listed next to each step in the pathway. The names of the genes encoding the activities are shown in italicized capital letters, following *S. cerevisiae* conventions. The *MET1* and *MET8* genes encode proteins that are involved in synthesizing siroheme, an essential cofactor for sulfite reductase. The *MET7* and *MET13* gene products catalyze the last two steps in the synthesis of the methyl donor used by Met6p, homocysteine methyltransferase, to synthesize methionine. (Adapted from Thomas *et al.*,1992)

Chapter 6

Biochemistry of the sulfur amino acids

S. cerevisiae requires three sulfur-containing amino acids to live. In addition to Met and Cys, which are incorporated into cellular proteins, cells also require S-adenosylmethionine (AdoMet), which supplies activated methyl groups for many methylation reactions. The consensus view of the synthesis of these three amino acids on the previous page is now well-supported by biochemical and genetic evidence from many laboratories (reviewed in Thomas & Surdin-Kerjan, 1992). The gene-enzyme relationships could not be definitively established until the development of molecular cloning and DNA sequencing techniques, which enabled investigators to use plasmid complementation to test gene function directly. In these experiments, investigators constructed plasmids with wild type *MET*, *CYS* or *SAM* genes, which were transformed into mutant strains. Transformed strains were only able to survive when the plasmid contained the wild type allele of the inactivated gene in the mutant. (You will use plasmid complementation in this class to confirm the identification of your strains and plasmids.)

Most of the genes that we will be working with this semester encode enzymes that catalyze an interconversion of one sulfur-containing molecule to a second sulfur-containing molecule. Other *MET* genes encode enzymes that do not directly participate in the synthesis of sulfur amino acids, but catalyze the synthesis of a cofactor or a methyl donor required for synthesis of sulfur amino acids. In the brief description below, we will follow the progress of a sulfur atom from inorganic sulfate through its conversion to Met, Cys or AdoMet.

Sulfate assimilation involves sulfur activation and reduction to sulfide

The early steps of the pathway, which encompasses the reactions involved in the conversion of sulfate to sulfide, comprise the sulfate assimilation pathway. Sulfate ions are the source of most sulfur in biological molecules, but considerable metabolic energy is required to activate sulfate from its +6 oxidation state and to convert it into sulfide, which has a -2 oxidation state. The enzymes responsible for sulfate assimilation are widely distributed in microorganisms and plants. In *S. cerevisiae*, sulfate is first activated by ATP sulfurylase, the *MET3* gene product, to form 5'-adenylylsulfate (APS). APS is then phosphorylated by Met14p, or APS kinase, forming 3'-phospho-5'-adenylylsulfate (PAPS). PAPS is an interesting molecule, since it contains an activated sulfur atom that can be used for a variety of sulfur transfer reactions. In mammals, PAPS in used for a variety of sulfation reactions in the Golgi, where the acceptors include lipids, proteins and a variety of small molecules. (Interestingly, APS kinase is the only yeast enzyme involved in sulfate assimilation with homologs in mammals.)

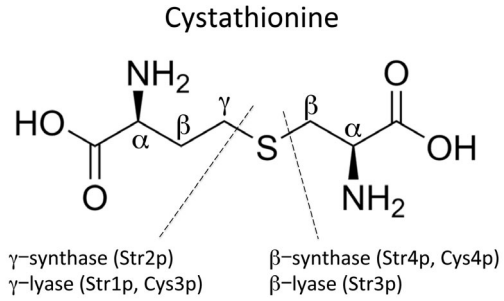
The final two steps in sulfate assimilation are NADPH-dependent reduction reactions. PAPS reductase, or Met16p, catalyzes the first reaction, which adds two electrons to the sulfur atom. The final 6-electron reduction is catalyzed by sulfite reductase. Sulfite reductase is a complex metalloenzyme containing two Met5p and two Met10p subunits as well as multiple prosthetic groups, including siroheme, that participate in electron transfer. (A prosthetic group is a metal ion or organic molecule that is covalently bound to an enzyme and essential for its activity.) In yeast, siroheme is synthesized in a series of reactions catalyzed by Met1p and Met8p.

Siroheme synthesis is not formally considered to be part of the sulfate assimilation pathway, but its function is critical for the assembly of functional sulfite reductase.

Homocysteine synthesis and transsulfuration

In the next step of Met and Cys biosynthesis, sulfide becomes incorporated into the amino acid homocysteine (Hcy). Hcy sits at the branch point between several pathways in yeast. The amino acid backbone of Hcy ultimately derives from aspartic acid, which has been converted in a series of steps to homoserine. (Note: "homo" amino acids have an extra carbon atom in their side chains compared to the namesakes without the prefix.) Met2p activates the homoserine in an acetylation reaction that uses acetyl-CoA. Homocysteine synthase, or O-acetyl homoserine sulfhydryase, then catalyzes the reaction of O-acetylhomoserine with sulfide to form Hcy.

In yeast, Hcy serves as the precursor for either Cys or Met. The pathway connecting Hcy and Cys is referred to as the transsulfuration pathway. Transsulfuration provides *S. cerevisiae* with unusual flexibility with respect to sulfur sources. Four different gene products are involved in the conversion of Hcy to Cys and vice versa, using cystathionine (below) as a common intermediate. Str2p catalyzes cystathionine synthesis from Cys and O-acetylhomoserine, the product of the reaction catalyzed by Met2p. In the opposite pathway, Cys4p (aka Str4p) catalyzes cystationine synthesis from Hcy and Ser. The four genes in the sulfur transfer pathway show different patterns of evolutionary conservation. For example, *E. coli* is unable to synthesize Cys from Met, while mammals are unable to synthesize Met from Cys.



Cystathionine is the intermediate for transsulfuration reactions. Enzymes in the S. cerevisiae transsulfuration pathway are encoded by the *STR1-STR4* genes. Str2p and Str1p (Cys3p) catalyze the synthesis and hydrolysis, respectively, of the cystathionine S-C_g bond. Str3p and Str4p (Cys4p) catalyze the synthesis and hydrolysis, respectively, of the cystathionine S-C_b bond.

Methionine and AdoMet are formed during the methyl cycle

Hcy is also the starting point of a cycle that produces Met and AdoMet. The cycle begins as Met6p catalyzes the conversion of Hcy to Met, using a unusual methyl donor, polyglutamyl 5-methyl-tetrahydrofolate (THF). The *MET13* and *MET7* genes encode the enzymes that catalyze the last two steps in the synthesis of polyglutamyl 5-methyl-THF, which accounts for their inability of *met7* and *met13* cells to synthesize methionine.

As you might expect, most methionine is used for protein synthesis in cells, but an appreciable amount is converted to the high energy methyl donor, AdoMet, by two nearly identical AdoMet synthases, Sam1p and Sam2p. *S. cerevisiae* is able to synthesize large quantities of AdoMet, which is either used for transmethylation reactions or stored in its vacuole. (In

fact, yeast is the source for most commercially-produced AdoMet.) Many different yeast enzyzmes transfer methyl groups from AdoMet to hundreds of different acceptors, which include nucleotide bases and sugars in DNA and RNA, various amino acid side chains in proteins, lipids, small molecules, and more. Each transmethylation reaction generates a molecule of S-adenosylhomocysteine (AdoHcy), which is hydrolyzed to adenosine and Hcy by Sah1p, completing the methyl cycle.

We will not be studying the enzymes involved in the methyl cycle in this class, but it is important to appreciate their importance to cell survival. The amino acid sequences of Sam1p and Sam2p are 93% identical, which is far higher than other proteins that have arisen by gene duplication in *S. cerevisiae*. This redundancy provides a buffer against loss of either function. Cells with a mutation in either the *SAM1* or *SAM2* gene are able to survive, but cells with mutations in both genes are unable to survive. Similarly, the *SAH1* gene is one of the few essential genes in *S. cerevisiae*, probably because the build-up of AdoHcy would inhibit many methyltransferase reactions.

Mutations disrupt biochemical pathways

The *met* mutants that you are analyzing are unable to catalyze one of the reactions required for sulfur amino acid synthesis. In this lab, you will use selective and differential media to determine which genes have been inactivated in your strains. Think of each mutation as erasing one of the arrows shown in the sulfur amino acid pathway. Our selective media contain a variety of sulfur sources. Depending on the position of the *met* mutation relative to the sulfur source, the strain may or may not be able to synthesize the sulfur amino acids.

You will also be using the differential medium, BiGGY agar to distinguish yeast strains by the quantity of hydrogen sulfide that they produce. This is because BiGGY contains bismuth, which reacts with sulfide to form a brownish to black precipitate. All strains are expected to grow on BiGGY, since it contains yeast extract, which is a source of methionine. BiGGY also contains sulfite, rather than sulfate, as the source of inorganic sulfur. Locate the positions of your mutated genes in the pathway relative to sulfide. Mutations in genes upstsream of sulfide should produce lighter colonies, since less sulfide will be produced. Mutations in genes downstream of sulfide should produce darker colonies, because the strains will be unable to metabolize sulfide.

In making your predictions for this experiment, you may find this analogy useful: A metabolic pathway is not unlike a one-way (*due to energetic considerations, most metabolic pathways are unidirectional*) highway with a series of bridges connecting islands. (*The islands are different energy levels.*) The cars passing along the highway are the molecules that are being converted from one form to another. When a car reaches the next island in the pathway, its color changes because it has been converted into a different molecule. The bridges are the enzymes in the pathway. They facilitate the passage of the cars, because they catalyze the reactions that convert one molecule to the next. If a mutation occurs in a gene that encodes a particular enzyme, that particular bridge falls down. Cars begin to pile up before the broken bridge, and very few cars would be found on islands past the broken bridge. In some cases, there may be an alternative route, or salvage pathway, but this is usually a less efficient route.

Exercise 1 - Predicting growth properties of mutant strains

Our class *met* mutants are derived from strain BY4742, which has the genotype *MATa his3-Δ1 leu2Δ0 lys2Δ0 ura3Δ0*. The defined media are based on YC, which contains histidine, leucine, lysine and uracil. YC Complete medium also contains methionine and sulfate, so we will use "dropout" media to test the ability of strains to grow on various sulfur sources. Modifications to YC are as follows:

YC-Met	methionine has been removed from the YC
YC-Met+Cys	cysteine has been added to YC-Met plates
YC-Met+SO_3	sulfite replaces sulfate in YC-Met plates

Predict the ability of *met* mutants to grow on various sulfur sources and complete the table below. Place a plus (+) when you predict that the strain will grow on the plate and a minus (-) when you do not expect the strain to grow.

BiGGY agar plates are used to detect sulfide production. Use upward- and downward-facing arrows to predict whether strains will give rise to darker or lighter colonies than BY4742.

	YPD	**YC Complete**	**YC - Met**	**YC-Met +Cys**	**YC-Met +SO_3**	**BiGGY**
met1	+					
met2	+					
met3						
met5						
met6						
met7						
met8						
met10						
met13						
met14						
met16						
met17						
cys3						
cys4						
str2						
str3						
sam1						

Chapter 6

Exercise 2 - Identifying strains by nutritional requirements

Your team will be given three strains, each of which carries a different *met* mutation. Prepare spot plates (Chapter 4) to distinguish between the three strains. Each member of the team should prepare serial dilutions of a single strain.

1. Spot your dilution series on each of the plates that your team received. Spot the complete dilution on one plate before proceding to the second plate. Use the same pattern of strains/rows on each of the different selective plates. ***Make sure that the plates are properly labeled!***

2. Incubate the plates at 30 °C until colonies should become apparent. Note that some colonies grow slowly on defined media and may require more than 3 days to appear. When colonies reach the desired size, transfer the plates to the cold room for storage.

3. Scan the plates as you did in Chapter 4 to record your data. ***These data will become the focus of your first lab report for the semester.*** Think of how you would like your figure to look as you place the pates on the scanner.
   - Scan the plates containing variations of YC media together, taking care to orient the plates in the same direction.
   - Scan the plate containing BiGGY agar separately using the color settings.

4. Use the predictions from the previous exercise to identify your team's mutant strains. This information will be compiled into a table in your lab report.

Consult the "Write It Up!" chapter for instructions on preparing lab reports.

References

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P & Boeke JD (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruptions and other applications. *Yeast* **14:** 115-132.

D'Andrea R, Surdin-Kerjan Y, Pure G, & Cherest H (1987) Molecular genetics of *met17* and *met 25* mutants of *Saccharomyces cerevisiae*: intragenic complementation between mutations of a single structural gene. *Mol Gen Genet* **207:** 165-170.

Masselot M & DeRobichon-Szulmajster H (1975) Methionine biosynthesis in *Saccharomyces cerevisiae*. I. Genetical analysis of auxotrophic mutants. *Mol Gen Genet* **139:** 121-132.

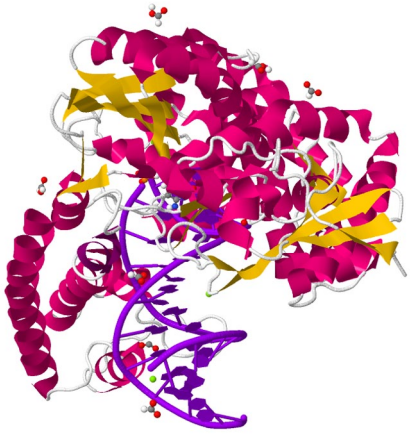
Sherman F (2002) Getting started with yeast. *Method Enzymol* **350:** 3-41.

Thomas, D & Surdin-Kerjan Y (1997) Metabolism of sulfur amino acids in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **61:** 503-532.

Winzeler EA, Shoemaker DD, Astromoff A *et al.* (1999) Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. *Science* **285:** 901-906.

Chapter 7

Yeast Colony PCR



The *S. cerevisiae* strains that we are using this semester were constructed by the *Saccharomyces* Gene Deletion Project. In the project, yeast investigators systematically replaced every ORF in the yeast genome with a bacterial kanamycin resistance (KAN^R) gene. In this lab, you will use the polymerase chain reaction (PCR) to identify the disrupted *MET* genes in your deletion strains. Thermostable DNA polymerases (above) play a key role in PCR.

Objectives

At the end of this lab, students should be able to:

- describe the reactions that are occuring at the different temperatures used in PCR cycles.
- design oligonucleotide primers to amplify sequences with PCR.
- explain how changes to the annealing temperature and extension time affect the production of PCR products.
- design and carry out a PCR strategy that distinguishes three *met* deletion strains.

Chapter 7

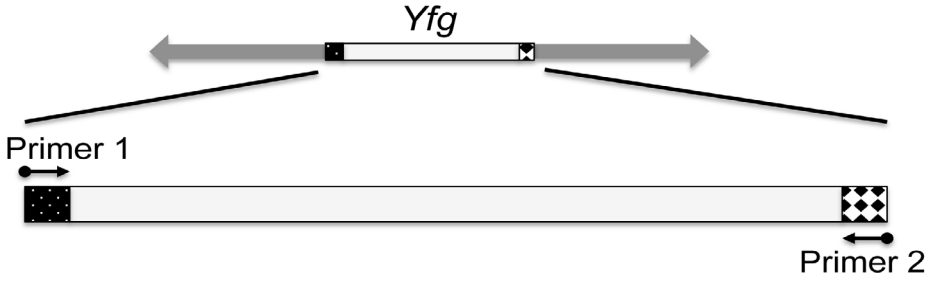
In a previous lab, you used different culture media to distinguish between different *S. cerevisiae met* mutants. Your results may have allowed you to tentatively identify your strains. In this lab, you will use the polymerase chain reaction (PCR) to more conclusively identify the mutant strains. This chapter begins with an overview of the PCR and the *Saccharomyces* Gene Deletion Project. You will use this knowledge to design and carry out a strategy for identifying *met* deletion strains by yeast colony PCR.

Polymerase chain reaction overview

The polymerase chain reaction (PCR) revolutionized molecular biology. With PCR, researchers had a tool for amplifying DNA sequences of interest from extremely small amounts of a DNA template. Indeed, billions of copies can be synthesized from a single DNA molecule in a typical PCR reaction. The development of PCR grew out of research on DNA polymerases and the discovery of thermostable DNA polymerases able to withstand extended heat treatments that denature most proteins (Sakai *et al.*, 1988). Today, PCR is a standard technique that is widely used to analyze DNA molecules and to construct novel recombinant molecules.

Thermostable DNA polymerases are central to PCR. The first description of PCR used a DNA polymerase from *E. coli*, which denatured and had to be replaced after each round of DNA synthesis (Sakai *et al.*, 1985). The procedure was much-improved by replacing the *E. coli* polymerase with a DNA polymerase from *Thermus aquaticus*, a bacterium that thrives in thermal springs at Yellowstone National Park. The *T. aquaticus* DNA polymerase, or *Taq* polymerase, functions best at temperatures of 70-75˚C and can withstand prolonged (but not indefinite) incubation at temperatures above 90˚C without denaturation. Within a few years, the *Taq* polymerase had been cloned and overexpressed in *E. coli*, greatly expanding its availability. Today, the selection of polymerases available for PCR has increased dramatically, as new DNA polymerases have been identified in other thermophilic organisms and genetic modifications have been introduced into *Taq* polymerase to improve its properties.

PCR involves multiple rounds of DNA synthesis from both ends of the DNA segment that is being amplified. Recall what happens during DNA synthesis: a single-stranded oligonucleotide primer binds to a complementary sequence in DNA. This double-stranded region provides an anchor for DNA polymerase, which extends the primer, ***ALWAYS*** traveling in the 5' to 3' direction. Investigators control the start sites for DNA replication by supplying oligonucleotides to serve as primers for the reaction (shown below for *Your favorite gene Yfg*). **To design PCR primers, investigators need accurate sequence information for the primer binding sites in the**

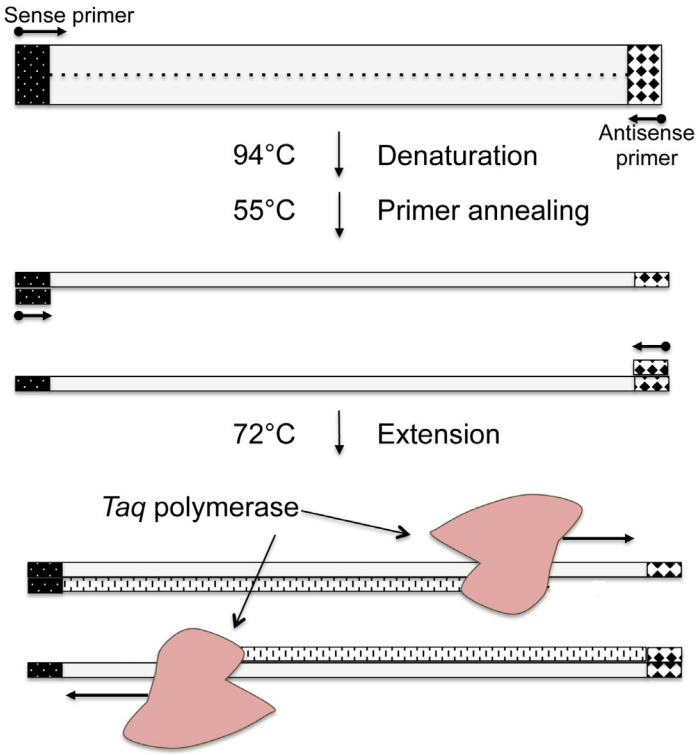


target DNA. (Note: Sequence information is not needed for the entire sequence that will be amplified. PCR is often used to identify sequences that occur between two known primer binding sites.) Two primers are required for PCR. One primer binds to each strand of the DNA helix.

PCR begins with a denaturation period of several minutes, during which the reaction mixture is incubated at a temperature high enough to break the hydrogen bonds that hold the two strands of the DNA helix together. Effective denaturation is critical, because DNA polymerase requires single-stranded DNA as a template. The initial denaturation segment is longer than subsequent denaturation steps, because biological templates for PCR, such as genomic DNA, are often long, complex molecules held together by many hydrogen bonds. In subsequent PCR cycles, the (shorter) products of previous cycles become the predominant templates.

Following the initial denaturation, PCR involves a series of 30-35 cycles with three segments, performed at different temperatures. PCR reactions are incubated in thermocyclers that rapidly adjust the temperature of a metal reaction block. A typical cycle includes:

- a denaturation step - commonly 94˚C
- a primer annealing step - commonly 55˚C
- an extension step - commonly 72˚C

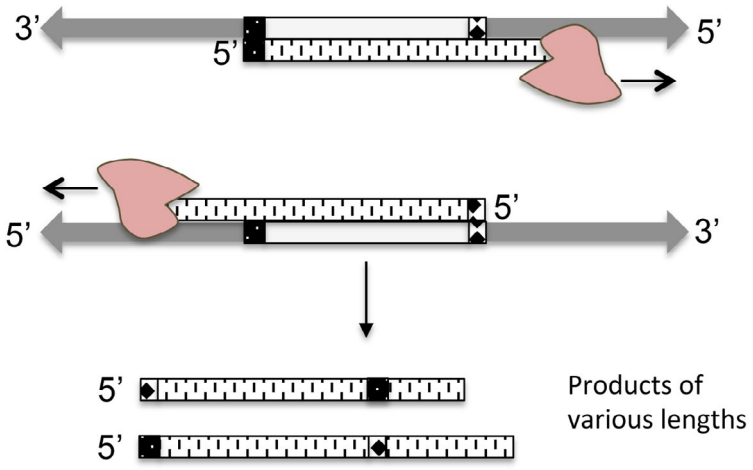


PCR reactions include multiple cycles of denaturation, annealing and extension.
Each cycle of PCR includes three different temperatures. In the denaturation step, the hydrogen bonds holding DNA helices together are broken. In the annealing step, oligonucleotide primers bind to single-stranded template molecules, providing starting points for processive DNA polymerases that extend the primer sequence. DNA polymerases become more active at the extension temperature, which is closer to their optimal temperature. Investigators adapt the temperatures and timing of the steps above to accommodate different primers, templates and DNA polymerases.

PCR products of the intended size accumulate exponentially

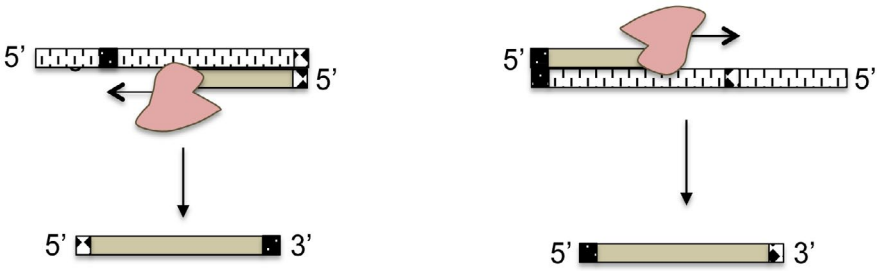
PCR is indeed a chain reaction, since the DNA sequence of interest roughly doubles with each cycle. In ten cycles, a sequence will be amplified ~1000 fold (2^{10}=1024). In twenty cycles, a sequence will be amplified ~million fold. In thirty cycles, a sequence can be theoretically amplified ~billion fold. PCR reactions in the lab typically involve 30-35 cycles of denaturation, annealing and extension. To understand PCR, it's important to focus on the first few cycles. PCR products of the intended size first appear in the second cycle. Exponential amplification of the intended PCR product begins in the third cycle.

During the first cycle, the thermostable DNA polymerases synthesize DNA, extending the 3' ends of the primers. DNA polymerases are processive enzymes that will continue to synthesize DNA until they literally fall off the DNA. Consequently, the complementary DNA molecules synthesized in the first cycle have a wide variety of lengths. Each of the products, however, has defined starting position, since it begins with the primer sequence. These "anchored" sequences will become templates for DNA synthesis in the next cycle, when PCR products of the intended length first appear. The starting template for PCR will continue to be copied in each subsequent cycle of PCR, yielding two new "anchored" products with each cycle. Because the lengths of the "anchored" products are quite variable, however, they will not be detectable in the final products of the PCR reaction.



First cycle of PCR
DNA polymerases synthesize complementary strands of the template DNA, beginning at the primer site. The lengths of the products are quite variable and depend on the processivity of the DNA polymerase.

DNA strands of the intended length first appear during the second cycle. Replication from the "anchored" fragments generates PCR products of the intended length. The number of these defined length fragments will double in each new cycle and quickly become the predominant product in the reaction.



PCR products of the intended length first appear in the second cycle.
The "anchored" fragments generated during the first PCR cycle begin with either the primer 1 or primer 2 sequence. During the second cycle, replication begins at the other primer site, generating a PCR product is capped at both ends with primer sequences.

Most PCR protocols involve 30-35 cycles of amplification. In the last few cycles, the desired PCR products are no longer accumulating exponentially for several reasons. As in any enzymatic reaction, PCR substrates have become depleted and the repeated rounds of incubation at 94˚C have begun to denature *Taq* polymerase.

Primer annealing is critical to specificity in PCR

Good primer design is critical to the success of PCR. PCR works best when the primers are highly specific for the target sequence in the template DNA. Mispriming occurs when primers bind to sequences that are only partially complementary, causing DNA polymerase to copy the wrong DNA sequences. Fortunately, investigators are usually able to adjust experimental parameters to maximize the probability that primers will hybridize with the correct targets.

PCR primers are typically synthetic oligonucleotides between 18 and 25 bases long. When designing a primer, researchers consider its T_m, the temperature at which half of the hybrids formed between the primer and the template will melt. In general, the thermal stability of a hybrid increases with the length of the primer and its GC content. (Recall that a GC-base pair is stabilized by three H-bonds, compared to two for an AT pair.) The following formula provides a rough estimate of the T_m of oligonucleotide hybrids. In this formula, *n* refers to the number of nucleotides, and the concentration of monovalent cations is expressed in molar (M) units.

$$T_m = 81.5^{\circ}C + 16.6 (\log10[K+ + Na+]) + 0.41 (\%[G + C]) - (675/n)$$

When possible, researchers design primers that are similar in length and have a 40-60% GC composition. The sequences are designed so that the T_ms of the primer-DNA hybrids are within a few degrees of the annealing temperature. Adjusting the T_ms of the primers to be close to the annealing temperature favors specific hybrids over less specific hybrids that may contain a few mismatched bases. A hybrid formed between a primer and a non-target sequence with even one mismatched base has a T_m that is lower than that of the fully hydrogen-bonded hybrid. If DNA polymerase extends the mismatched primer, incorrect PCR products will be generated. When mispriming appears to be a problem in a PCR reaction, investigators have several options to

increase the yield of the desired product. They can increase the length and/or GC content of the primers, alter the salt concentrations (results may be hard to predict) or increase the annealing temperatures.

When designing a PCR reaction, investigators also consider the nature of the template DNA. A variety of DNA templates can be used for PCR. Depending on the purpose of the experiment, investigators could choose to use genomic DNA, a plasmid or a cDNA (complementary DNA generated by a reverse transcriptase from mRNA). PCR can also be done with much cruder preparations of DNA, such as a bacterial or yeast colony. The more complex the template (its length in bp), the greater the probability that it will contain another sequence that is very similar to a primer sequence. For example, the haploid yeast genome is 12 Mbp long and contains only one copy of each *MET* gene. The probability that a non-target sequence in the yeast genome is similar enough to a 25-nucleotide *MET* primer to cause mispriming is reasonably good. Furthermore, these sequences with small mismatches may outnumber the target sequence. With complex templates such as genomic DNA, therefore, investigators can sometimes reduce the impact of mismatched hybrids by decreasing the amount of template DNA in the reaction. (Using too much template is the most common error in yeast colony PCR.)

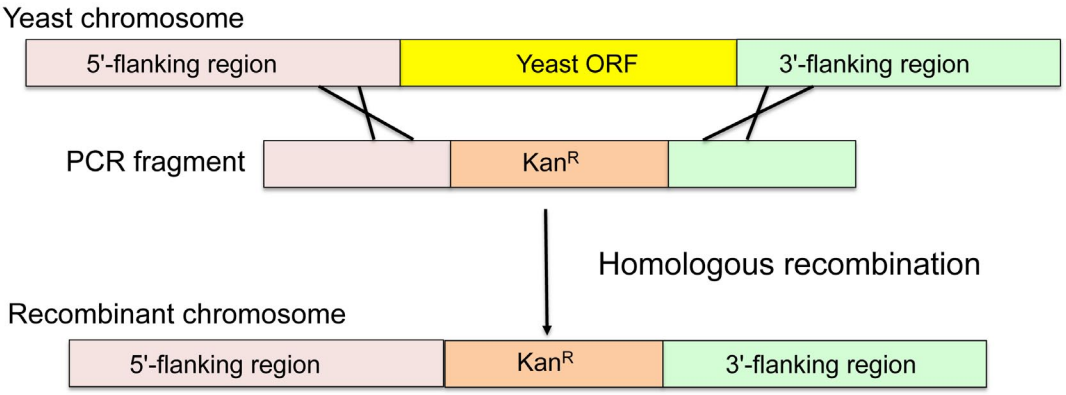
The components of a PCR reaction are simple, consisting of the DNA template, primers, dNTPs, a buffer containing $MgCl_2$ (polymerases use dNTPs complexed with Mg^{2+}), and the thermostable polymerase. ***For our experiments, we will be using a master mix that contains all of the components except the template DNA and the primers.*** The use of a master mix ensures that all reactions have identical reagents and it also reduces the number of transfers requiring micropipettes. The smaller number of transfers is particularly advantageous, because it reduces the opportunities for cross-contamination of reagents. PCR is an exquisitively sensitive procedure. Some researchers even use special barrier tips for their micropipettes, which contain filters that prevent samples from reaching the barrel of the micropipettes.

Saccharomyces Genome Deletion Project

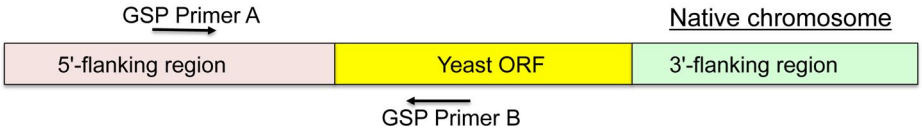
The publication of the yeast genome sequence opened new opportunities for yeast geneticists. Knowing the DNA sequence of the yeast genome, geneticists could now take advantage of the high frequency with which yeast exchange genes by homologous recombination to generate mutants of their own design. Homologous recombination normally occurs during meiosis and during certain kinds of DNA repair. During homologous recombination, two closely related DNA sequences align with one another, breaks appear in the DNA molecules and strand exchange occurs when the breaks are repaired. Homologous recombination allows researchers to replace a chromosomal gene with a DNA construct of their own design. To use this strategy, investigators first construct a replacement cassette in which a marker gene is flanked by sequences that are identical to chromosomal DNA sequences at either side of the the yeast target gene. The sequence is then introduced into cells by chemical transformation (Chapter 12) or by

electroporation. Transformed cells that have incorporated the replacement cassettes into chromosomal DNA can be identified by selecting for the marker gene in the replacement cassette.

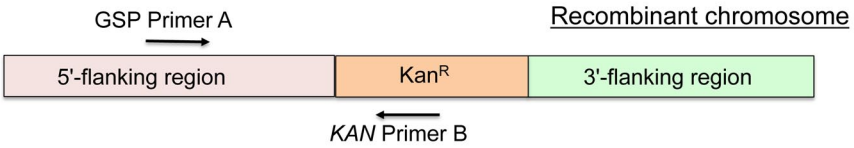
The *Saccharomyces* Genome Deletion Project (SGDP) used this approach to systematically replace each of the predicted ORFs in the *S. cerevisiae* genome with a kanamycin resistance (*KAN*^{R}) gene. For each ORF, researchers used a series of PCR reactions (below) to construct cassettes in which the *KAN*^{R} gene was flanked by short DNA sequences that occur upstream and downstream of the targeted ORF on the *S. cerevisiae* chromosome. The cassettes were then used to transform the BY4742 strain (Brachmann *et al.*, 1998). Strains that had incorporated the *KAN*^{R} gene were selected on plates containing analogs of kanamycin (Winzeler *et al.*, 1999).



The *S. cerevisiae* strains that we are using in our experiments are part of the SGDP collection of deletion mutants. In this lab, you will use PCR to identify which *MET* genes have been disrupted in your strains. The PCR primers were designed and tested by the SGDP to verify strain identities. Available primers include two gene-specific primers (GSP) for each of the *MET* genes. One of the GSPs, GSP Primer A, is located 200-400 bp upstream of the initiation codon. The second GSP, GSP Primer B, is an antisense primer that binds within the ORF. We also have an antisense primer that binds 250 bp within the *KAN*^{R} gene (*KAN* Primer B).



Gene-specific (GSP) primers A and B give a product from the native chromosome, but not the recombinant chromosome



Gene-specific (GSP) primer A and KAN primer B give a product from the recombinant chromosome, but not the native chromosome

Chapter 7

Exercise 1 - Physical properties of PCR primers

The table below lists the primer pairs used by the SGDP to analyze the deletion mutants. In designing the primers, researchers aimed for primer pairs that had similar physical properties, *i.e.* length and T_m, and would generate PCR products that were several hundred base pairs long. Like most genomes, the *S. cerevisiae* genome has both AT-rich and GC-rich regions (Goffeau *et al.*, 1996). Non-coding regions, in particular, tend to be enriched in AT base pairs, which are stabilized by fewer hydrogen bonds than GC pairs.

Gene	**Primer A**	**Primer B**
MET1	TTCTATTTTCGTTATTGGTTTCTCG	AAATGAACCTGATCAGTAGCAAAAC
MET2	AAGTCATGTTAATCGTTTGGATTTG	GTCCAAGTAGTTGGGATCTGAGTAA
MET3	GTAATTTTGTAACTCACCGCATTCT	CATTCTTCTTTAACGCATCTCTAGC
MET5	TTCATCACGTGCGTATTATCTCTTA	GGTATTCAATGGATCTTGATTGTTC
MET6	ATGCGATAGATGCACTAATTTAAGG	AAAACTTGGTCGTAAAAGGAGAAGT
MET7	GTTGGTTAACAGAAAAAGGCAACTA	TCATGCATTTCCAATAATGTCATAG
MET8	ATGCCATTTCAGTTACAACCTAGTC	GAATAATGGATTTGTGTAGGTCAGG
MET10	AAAGAAAACACTATCAACATTCCCA	AGTTTAAAGCACCAACATTCAAAAG
MET14	AAAGAATACAGTTGCTTTCATTTCG	GATTGTACTTTTACCTGACGCACTT
MET16	GCTGACAAAAGAATTGGATAAAAGA	ATATACTGTTTAACCTGCTCGAACG
MET25	CATCCTCATGAAAACTGTGTAACAT	GCAGAATGTGTTACAATATCAGCAC
CYS3	ACCCCATACCACTTCTTTTTGTTAT	ATAGGGTTAGCTGGAGAAGATTGTT
CYS4	ACAACTTCAACTTCACCCAAGTAAG	TCAAGTCTTCTAGCTGTCTTTGGAT
KAN^R		CTGCAGCGAGGAGCCGTAAT

Note: The primer sequences are written in Courier font, a nonproportional font that is often used for nucleic acids, because all letters have the same width. What is immediately apparent about the primer lengths?

In this lab, you will use the SGDP primers to identify your three deletion strains. Fill in the table below with information about the length, GC-content and T_m for the primers designed for your strain. To find this information, use one of several online tools for primer analysis:

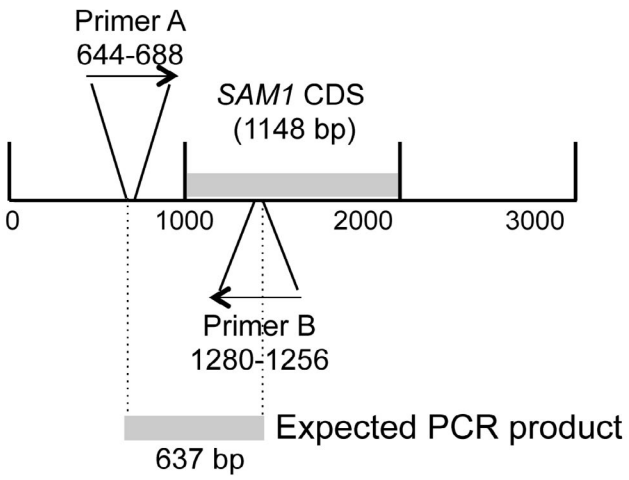
http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/
http://www.basic.northwestern.edu/biotools/oligocalc.html

	Primer A			**Primer B**		
Gene	**Length**	**% GC**	**T_m**	**Length**	**% GC**	**T_m**
KAN^R						

Exercise 2 - Map the PCR product to the genome

In the next lab, you will analyze your PCR reaction products on an agarose gel that separates DNA molecules according to their sizes. To interpret those results, it will be important to have calculated the expected sizes of the PCR products from your reactions. To assist those calculations, prepare a simple map of the genomic region containing your gene with the primer sequences aligned against the genome sequence. All PCR products should contain a portion of the *MET* gene's 5'-flanking region because of primer A. A PCR product may or may not contain portions of the *MET* gene's CDS, depending on whether you are analyzing a strain with the native or disrupted *MET* gene. To constsruct your map, you can take advantage of special genome sequence records prepared by SGD curators. These records contain the CDS for your *MET* gene together with 1 kb of upstream and 1 kb of downstream sequence. SGD curators generated these records because researchers are often interested in studying regulatory elements that control transcription of a gene and the processing of gene transcripts. In *S. cerevisiae*, these regulatory elements are usually located within 1 kb of the CDS.

The figure on the right shows a map for the *SAM1* gene. The binding site for primer A is 357 nucleotides upstream of the *SAM1* initiation codon (nucleotide 1001). Primer B-anchored products add 280 bp of CDS to the PCR product. The expected size of the PCR product is 357 + 280 bp, or 637 bp. If the deletion strain had been used for PCR, the *SAM1* primers A and B would not generate a PCR product. Instead, *SAM1* primer A and *KAN*^{R} primer B would generate a 607 bp product. (The *KAN*^{R} primer B adds 250 bp to PCR products.)



You will need two browser windows for this exercise. Each member of the group should work with a single gene.

Find the genomic sequence for your gene.

- Navigate to your gene's summary page in the SGD (yeastgenome.org)
- Cursor down the sidebar to "Retrieve sequences" (Under Sequence Information)
- Retrieve the "DNA +/- 1kb" sequence for your gene.
- Note below the starting and ending coordinates for the sequence and calculate the length of the sequence.

Starting and ending coordinates:

Length of sequence (bp) Length of the coding sequence

Align the primer sequences with the genomic sequence.

To find the position on the primers in the genomic sequence, we will use NCBI's BLAST tool. BLAST stands for Basic Local Alignment Search Tool and can be used to align either protein or nucleic acid sequences. You will learn more about the BLAST algorithms in Chapter 9.

- Direct your browser to the NCBI site and access the BLAST tool. Select Nucleotide BLAST.
- Click the box "Align two or more sequences."
- Copy the genomic sequence from SGD and paste the sequence into the Subject Sequence box.
- Type the Primer A for your gene in query box.
- Adjust the BLAST algorithm for a short sequence. The primer sequences that we are using are 25 nucleotides long. This is shorter than the default value of 28 for "words" in BLASTN (the algorithm for comparing nucleotide sequences). BLAST will not align two sequences if the match is smaller than 28 nucleotides. Expand the "algorithm parameters" at the bottom of the page. Select a "word size" less than 28.
- Click BLAST. The BLAST results bring up a table that shows each match between your primer and the genome sequence. The top result should be a perfect match between your primer and the genome sequence. (Check your typing if it isn't a perfect match!) Record the starting and ending nucleotides in the genomic DNA sequence where it matches the primer sequence.

- Repeat the BLAST alignment for primer B. Click "Edit and Resubmit" at the top of the BLAST results page. Clear the query box and type in the sequence of primer B. Click BLAST and record the alignment results. In the results, note that the primer nucleotide numbers are ascending, while the genomic DNA nucleotide numbers are in descending order. This is because Primer B sequence is the reverse complement of the gene sequence.

Draw a map of your gene and primer binding sites in the space below. Include distances in bp.

Calculate the sizes of the PCR products that would be generated with:

Primer A and Primer B

Primer A and *KAN*^{R} primer B

Exercise 3 - Design yeast colony PCR

In this lab, each team will be able to perform **six** PCR reactions to identify your three *S. cerevisiae* strains. You will have the option of using any of the *MET* primers listed above, as well as the *KAN*^{R} primer B. It will not be possible to test every strain with both the GSPA-GSPB and GSPA-*KAN*^{R} B combinations. Use your results from the selective plating experiment as you work with your team to devise a strategy that will allow you to positively identify your *met* strains.

List the primer pairs that you will use for the reactions, together with the predicted sizes of the PCR products in your notebook.

Exercise 4 - Yeast colony PCR

To prepare the reactions, you will first mix the primer pairs with a ***VERY SMALL*** number of yeast cells that you transfer from a colony to the tube with the tip of a P-20 or P-200 micropipette. The colony and primers will then be heated at 98˚C for 15 minutes to disrupt the yeast cells. At that point, you will add an equal volume of a PCR master mix, containing nucleotides and the *Taq* polymerase, to each tube. The tubes will then be returned to the thermocycler for a typical PCR reaction.

1. ***Label the PCR tubes.*** The tubes are very small, so develop a code that you can use to identify the tubes. (Don't forget to include the code in your notebook. The code should indicate which primers and strains are mixed in each tube.)

2. ***Prepare the primer mixtures.*** The final volume of the PCR reactions will be 20 µL. The primer mixture accounts for half the final volume, or 10 µL. The primers stock concentrations are 2.0 µM each. Pipette 5.0 µL each of the two primers that you would like to use into each PCR tube. *What will the final concentration of each primer be in the actual PCR reaction?*

NOTE: Because of the extraordinary sensitivity of PCR reactions, it is very important not to cross-contaminate tubes with unintended primers. Change tips between every primer transfer.

3. ***Transfer a small quantity of yeast cells to each PCR tube.*** Lightly touch the tip of a P20 or P200 micropipette to a yeast colony. Twirl the micropipette tip in the tube containing your primer mix to release the cells. The most common error is transferring too many yeast cells, which will interfere with the PCR reaction. The right amount of yeast will fit on the tip of a pin.

4. ***Lyse the yeast cells.*** Place the tubes in the thermocycler for 15 min at 98˚C.

5. ***Set up the PCR reactions.*** Remove the tubes from the thermocycler and add 10 µL of PCR master mix to each tube.

6. ***Amplify the target gene sequences.*** Return the tubes to the thermocycler and start the PCR program.

Our thermocyclers are set for the following program:

- One cycle of denaturation: 95°C for 2 minutes
- 35 cycles of denaturation, annealing and extension:
  95°C for 30 sec.
  55°C for 30 sec.
  72°C for 1 minute
- One cycle of extension: 72°C for 10 minutes

Test yourself

1. A researcher plans to amplify the CDS of the *SAM1* gene. The partial sequence below shows the coding sequences for the N-terminus and C-terminus of Sam1p. Design two 18-nucleotide long primers to amplify the sequence the *SAM1* CDS.

   5'-ATGGCCGGTACATTTTTATTC.....(CDS).....TCCAAGACTTTGAAGTTCTAA - **3'**

2. Most of the primers that we are using for PCR have T_ms that are slightly lower than the annealing temperature used in the PCR reactions. Thus, less than half of the target sites are expected to anneal with the primers in each cycle. Why would investigators choose to design primers that will not be fully annealed with the DNA template?

References

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P & Boeke JD (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruptions and other applications. *Yeast* **14:** 115-132.

Goffeau A, Barrell BG, Bussey H *et al.* (1996) Life with 6000 genes. *Science* **274:** 563-567.

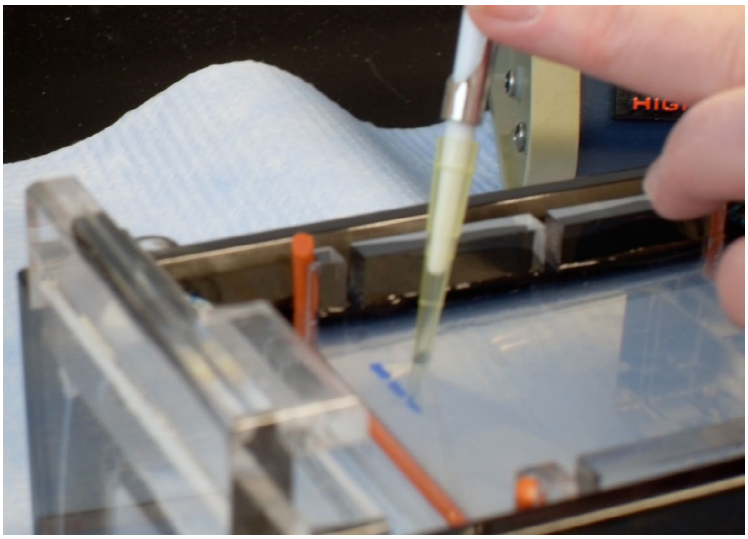
Sakai RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA & Amheim N (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230:** 1350-1354.

Sakai RK, Gelfand DH, Stofel S, Scharf SJ, Higucki R, Horn GT, Mullis KB & Erlich HA (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239:** 487-491.

Winzeler EA, Shoemaker DD, Astromoff A *et al.* (1999) Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. *Science* **285:** 901-906.

Chapter 8

Agarose gel electrophoresis



Agarose gels are used to analyze DNA molecules. These gels are simple to construct, because they rely only on the gelling properties of agarose. Molecules are separated by size and visualized with fluorescent intercalating dyes. In this lab, you will analyze the products of the PCR reactions from the previous lab.

Objectives

At the end of this lab, students should be able to:

- prepare an agarose gels for separating DNA molecules.
- separate DNA molecules by electrophoresis.
- visualize DNA molecules on agarose gels using intercalating dyes.
- estimate the approximate sizes of DNA molecules using size standards.

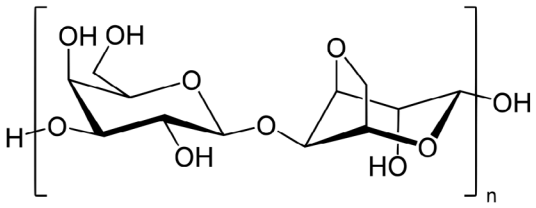
Chapter 8

Background

Agarose gels provide a simple method for analyzing preparations of DNA. DNA molecules share the same charge/mass ratio, which imparts similar electrophoretic properties to linear DNA molecules of widely varying lengths. We will use "molecule" to refer to a linear piece of DNA, but keep in mind that a single DNA "molecule" may contain the sequences of multiple genes or parts of genes.

Agarose gels are porous matrices

Agarose is a polysaccharide purified from red algae, or seaweed. Agarose is more highly purified (and significantly more expensive!) than agar, which is obtained from the same seaweed. Agarose molecules are long, linear polymers of a repeating disaccharide D-galactose and 3,6-anhydro-α-L-galactopyranose (right). A typical agarose molecule contains over one hundred subunits. The agarose used for electrophoresis has been highly purified. The purification process removes contaminants that would interfere with the enzymes used in molecular cloning, such as restriction endonucleases. The process also generates an agarose preparation with desirable electrophoretic properties and minimal background fluorescence, which is important for visualizing DNA molecules.



Agarose molecules are able to form gels with relatively defined pore sizes because of the chemical properties of agarose molecules. Agarose demonstrates hysteresis - its melting temperature is higher than its gelling temperature. Agarose molecules dissolve at about 90°C, forming random coils in solution. Gels form when the temperature falls to approximately 40°C. As the gel forms, the agarose molecules first assemble into helical fibers, which then further aggregate to form networks of supercoiled helices stabilized by hydrogen bonds. The sizes of the pores, which typically range from 50 to 200 nm, depend on the concentration of agarose. As the agarose concentration increases, the average diameter of the pore decreases.

Several factors affect the migration of DNA through agarose gels

Because of the negative charge of the phosphate residues in the DNA backbone, DNA molecules move toward the positive pole (anode) of the electrophoresis apparatus. The actual migration rate of DNA molecules in a particular experiment is affected by multiple factors. Some of these factors are intrinsic to the DNA molecules, while other factors relate to the electrophoretic conditions. Intrinsic factors include both the length and conformation of the DNA molecules that are being analyzed. Within a certain size range dictated by the gel conditions, the migration rate of linear DNA molecules is inversely proportional to the $\log_{10}$ (number of base pairs). The migration of more structured DNA molecules, such as supercoiled plasmids, is much less predictable. The migration rates of these more highly structured DNAs are influenced by the density of coils, the presence of nicks, and other structural features.

The migration rates of DNA molecules in agarose gels are also affected by the composition of the gel. The migration rate of a DNA molecule decreases as the concentration of agarose in the gel increases. Researchers commonly adjust the agarose concentration to optimize the resolution of DNA molecules within a particular size range. The two buffers commonly used in labs, TAE (Tris: acetate: EDTA) and TBE (Tris: borate: EDTA), also affect electrophoresis rates.

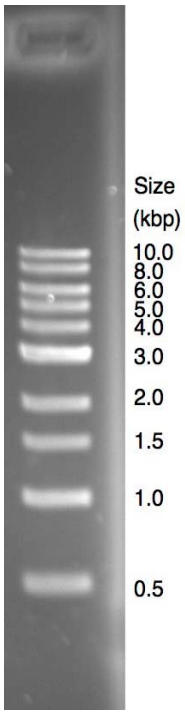
Because of this inherent variability, researchers ALWAYS include a lane of DNA standards with known sizes on the same gel as the samples being analyzed. Importantly, these standards need to have a similar structure (e.g. linear or supercoiled) and to be subjected to the same chemical modifications as the DNA samples being analyzed.

Fluorescent intercalating agents are used to visualize DNA molecules in gels

Nucleic acids are visualized by fluorescent dyes that bind strongly to DNA. The dyes are intercalating agents that insert into the DNA helix and into structured regions of single-stranded nucleic acids. The fluorescence of these dyes increases by an order of magnitude when they bind nucleic acids, so the background fluoresence on agarose gels is usually low. In this class, we will use ethidium bromide (EtBr) to visualize DNA fragments. EtBr absorbs light in the ultraviolet (UV) range and emits orange light. Gels are viewed with special transilluminators containing UV lights. EtBr is a light-sensitive compound, so stocks are stored in the dark.

Standards are used to estimate the sizes of DNA molecules

The figure on the right shows a lane from an EtBr-stained 1% agarose gel containing DNA size standards. The 1 kb ladder shown in this student gel is a proprietary mixture of linear DNA fragments ranging in size from 0.5 to 10 kilobases (kb). The staining intensity of bands on agarose gels reflects the quantity of DNA in the band, because EtBr intercalates fairly evenly along the length of linear DNA molecules. As you can see, this particular mixture contains similar quantities of each DNA fragment, with the exception of the 3 kb fragment, which has ~2.5 more DNA than the other fragments. The greater intensity of the 3 kb fragment serves as a useful orientation marker in situations where smaller fragments might have run off the end of the gel or when some markers are not well resolved from one another. (These are common occurrences!)



Note that the shortest molecules on the gel are more well separated from one another than the longer molecules. The PCR products that you will be analyzing in this lab are mostly in the 400-1000 bp range. To increase the resolution of these molecules, we will use 1.25% agarose gels. (This will reduce the resolution of larger DNA molecules). Note also that smaller bands appear fuzzier on the stained gel, because they have been more affected by random diffusion as they have migrated through the network of agarose polymers.

Chapter 8

Prepare the agarose gel

In this lab, you will use agarose gels to separate DNA molecules produced in PCR reactions. These PCR products should be well-resolved on 1.25% agarose gels prepared in TAE buffer, which provide good separation of molecules that are smaller than 2 kb. Place the casting tray into the gel apparatus. If you are using the BioRad apparatuses, position the black wedges at each end of the casting tray.

1. Determine the amount of agarose that you will need for a 1.25% (1.25 g/100 mL) gel that fits your casting platform. Most of the gel apparatuses in the lab are the BioRad Mini-Sub GT systems with a 7 cm x 7 cm casting tray that accommodates a 30 mL gel. You may need to adjust the size of your gel if you are using a different apparatus. Check your calculations with your teammates before you proceed.

2. Fill a graduated cylinder with the appropriate volume of TAE buffer. Pour the solution into a small flask.

3. Weigh out the appropriate amount of agarose. Sprinkle the agarose onto the surface of the TAE in the flask. Note: the agarose will not dissolve until it is heated.

4. Dissolve the agarose by heating the solution for intervals of 15-20 seconds in a microwave oven. After each interval, remove the flask and gently swirl it around a bit to disperse the contents. Note if the agarose particles are still apparent or if the agarose has dissolved. The best gels are made from agarose that has NOT been overcooked.

SAFETY NOTE: The agarose solution will be very HOT when you remove it from the microwave! Use caution when handling the flask. Be particularly careful not to contact the steam that will be coming through the opening of the flask. Fold several paper towels and wrap them around the neck of the flask when you handle it. If you do happen to spill some hot agarose on your skin, wash it immediately with cold water and alert your TA.

5. Allow the agarose solution to cool until you can comfortably touch the flask with your hands. Agarose solutions over 60˚C will warp the casting tray! Pour the gel. Place the sample comb in place. Do not move the casting platform until the gel sets. You will know that the gel is set when it becomes opaque. Allow the gel to cure for about 20 minutes after it sets.

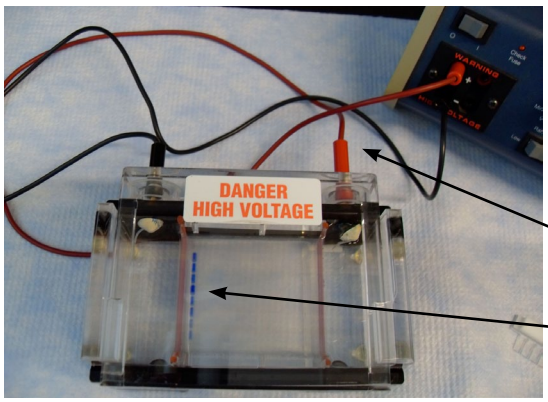
Sample preparation

Prepare your samples for electrophoresis while the gel is curing by adding concentrated loading buffer. The loading buffer contains two dyes, bromophenol blue and xylene cyanol. During electrophoresis, the dyes will migrate with "apparent" molecular weights of ~5 kb and ~0.5 kb, respectively. The loading buffer also contains glycerol, which makes the sample dense enough to sink to the bottom of the sample well.

Add 4 µL of 6X loading buffer directly to each of the 20 µL PCR reactions from the last lab. Briefly, centrifuge each tube to mix the dye and samples, if necessary. You will use half of each sample in your gels. ***Store the remaining sample in the refrigerator.***

Load and run the agarose gel

1. When the gel has set, ***carefully*** remove the comb and the black wedges.
2. Orient the gel in the electrophoresis tank such that the wells (holes made by the comb) are oriented toward the black (negative) electrode. DNA molecules will move from the well toward the red (positive) electrode. Fill the tank with enough TAE buffer to submerge the gel (approx. 275-300 mL).
3. Load one sample to each well, which can accommodate ~20 µL. Load 10-12 µL of a PCR sample to each lane. Try to avoid air bubbles as you load the samples.
4. Load 5 µL molecular weight standard to one lane of the gel. Make sure that you have accurately recorded the location of each sample in the gel.
5. Place the lid on the electrophoresis tank and connect the electrodes to the power supply (black-to-black and red-to-red). ***Make sure that the polarity is correct before continuing!***
6. Turn on the power and apply a constant voltage of 125 V.
7. Pay careful attention to the gel as it runs. Turn off the power when the bromophenol blue is ~ 1 cm from the end of the gel. ***Do not allow the dye to run off the gel, since small DNA molecules will be lost. (Think about the size of your PCR products.)***



Proper set-up of an agarose gel
The gel is submerged in running buffer. Samples will electrophorese toward the postivie (red) pole.

Red electrical leads are attached to the positive pole of the power pack.

Samples are loaded into wells.

Chapter 8

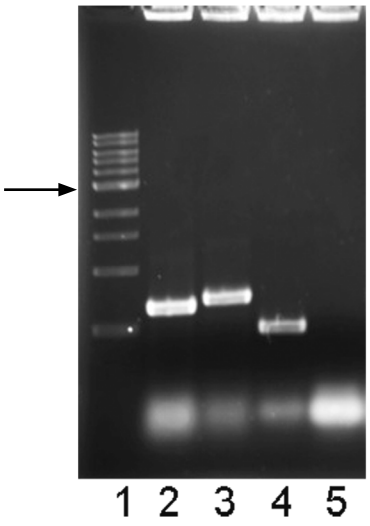
Stain and analyze the agarose gel

SAFETY NOTE: Wear disposable gloves when staining gels. Gloves are important when working with intercalating dyes, which are potential mutagens.

1. Remove the gel from the apparatus and transfer the gel to a small tray. Cover the gel with deionized water. Add 5 µL of EtBr solution (10 mg/mL) to the tray. *What is the approximate concentration of EtBr in the staining solution?*
2. Place the tray on a rocking platform and rock gently for 20-30 minutes.
3. Drain the EtBr solution in to the appropriate waste container in the fume hood.
4. Cover the gel with deionized water and rock gently for 1 minute.
5. With a spatula, carefully place the gel on the transilluminator and close the cover to the Gel-Logic apparatus. (Drain the wash solution into the waste container.)
6. Turn on the transilluminator light and photograph the gel according to the posted instructions. Turn off the transilluminator immediately after you photograph the gel. Save the picture and email a copy to yourself. (If no bands are apparent, the staining can be repeated for a longer period of time.)
7. Open the door of the GelLogic apparatus. Use the spatula to transfer the gel to a waste container set up for EtBr-stained gels.
8. Determine the approximate length of the DNA molecules in your samples by comparing their migration to that of the standards. Are the sizes consistent with your expectations?

Test yourself

A research group has just received four strains from a one of their collaborators. Following their usual practice, a student is using colony PCR to confirm the identities of the four strains A-D before starting work. The lab has diagnostic primers for all four strains. The expected sizes of the PCR products for strains A-D are 680 bp, 1100 bp, 572 bp and 462 bp, respectively. Unfortunately, the student who runs the gel forgets to write down which sample has been run in each lane. Analyze the gel on the right. No PCR product can be detected in lane 5. The position of the 3 kb marker is indicated. Place size labels for a few of the other markers. Which strain's PCR product has been loaded into:

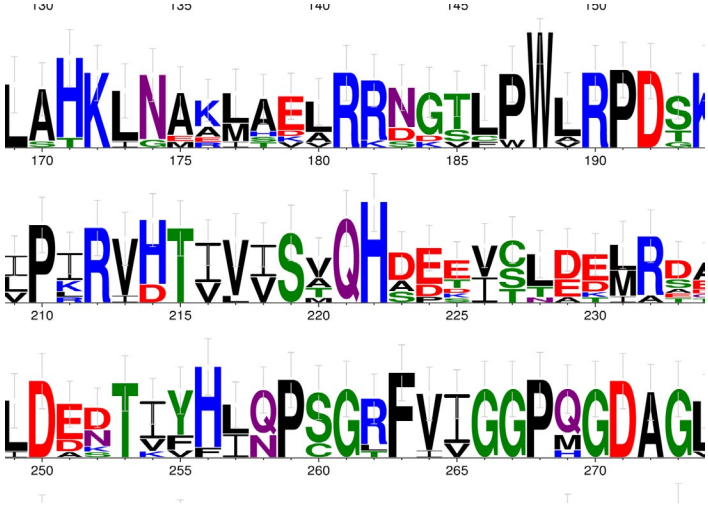


Lane 2 Lane 3 Lane 4 Lane 5

Explain your reasoning. What might the lack of a product in Lane 5 indicate?

Chapter 9

Protein conservation



Our experiments this semester will test whether gene involved in Met and Cys synthesis have been functionally conserved during the evolutionary divergence of *S. cerevisiae* and *S. pombe*. A variety of algorithms offer researchers tools for studying the evolution of protein sequences. In this graphic depiction of aligned Sam2p sequences from nine divergent model organisms, the height of the letter reflects the frequency of a particular amino acid at that position.

Objectives

At the end of this laboratory, students should be able to:

- identify amino acids by their 1-letter code.
- explain the differences between high and low scores on the BLOSUM 62 matrix.
- use the BLASTP algorithm to compare protein sequences.
- identify conserved regions in a multiple sequence alignment.

Chapter 9

As species evolve, their proteins change. The rate at which an individual protein sequence changes varies widely, reflecting the evolutionary pressures that organisms experience and the physiological role of the protein. Our goal this semester is to determine if the proteins involved in Met and Cys biosynthesis have been functionally conserved between *S. pombe* and *S. cerevisiae*, species that are separated by close to a billion years of evolution. In this lab, you will search databases for homologs of *S. cerevisiae* sequences in several species, including *S. pombe*. **Homologs** are similar DNA sequences that are descended from a common gene. When homologs are found in different species, they are referred to as **orthologs**.

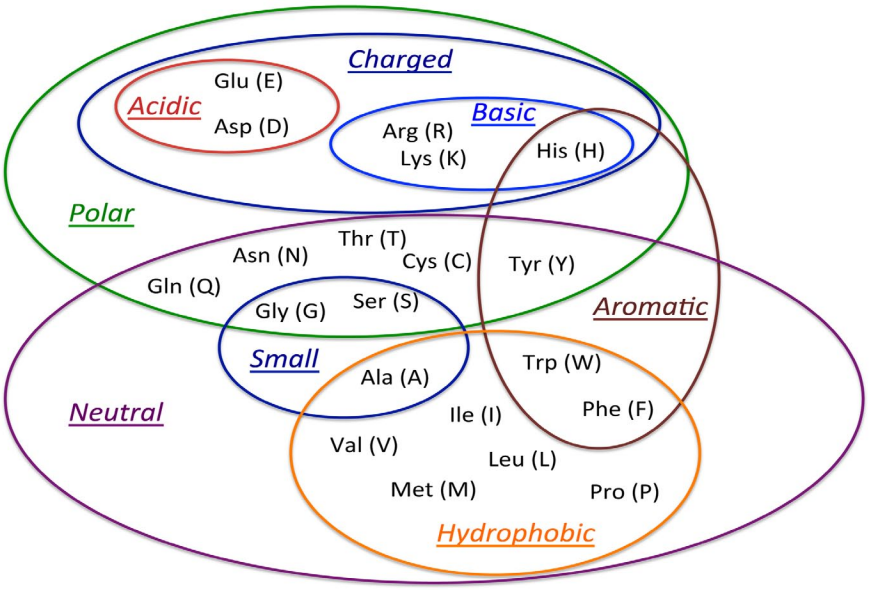
Homologs within the same genome are referred to as paralogs. **Paralogs** arise by gene duplication, but diversify over time and assume distinct functions. Although a whole genome duplication occurred during the evolution of *S. cerevisiae* (Kellis *et al.*, 2004), only a few genes in the methionine superpathway have paralogs. Interestingly, *MET17* is paralogous to three genes involved in sulfur transfer: *STR1* (*CYS3*), *STR2* and *STR4*, reflecting multiple gene duplications. The presence of these four distinct enzymes confers unusual flexibility to *S. cerevisiae* in its use of sulfur sources. The *SAM1* and *SAM2* genes are also paralogs, but their sequences have remained almost identical, providing functional redundancy if one gene is inactivated (Chapter 6).

Protein function is intimately related to its structure. You will recall that the final folded form of a protein is determined by its primary sequence, the sequence of amino acids. Protein functionality changes less rapidly during evolution when the amino acid substitutions are conservative. Conservative substitutions occur when the size and chemistry of a new amino acid side chain is similar to the one it is replacing. In this lab, we will begin with a discussion of amino acid side chains. You will then use the BLASTP algorithm to identify orthologs in several model organisms. You will perform a multiple sequence alignment that will distinguish regions which are more highly conserved than others.

As you work through the exercises, you will note that protein sequences in databases are written in the 1-letter code. Familiarity with the 1-letter code is an essential skill for today's molecular biologists.

Amino acid R groups have distinct chemistries

Each of the 20 amino acids commonly found in proteins has an R (reactive) group with its own distinctive chemistry. R groups differ in their size, polarity, charge and bonding potentials. When thinking about evolutionary changes in proteins, it is helpful to group the amino acids by their chemistry in a Venn diagram, shown on the opposite page. In general, replacing one amino acid with a second amino acid from the same sector can be considered a conservative change. The size of an R group is also important. Substitution of a large R group for a small one can significantly alter the function of a protein.



Exercise 1 - The 1-letter code for amino acids

You may find NCBI's Amino Acid Explorer helpful for this exercise. Access it at:
http://www.ncbi.nlm.nih.gov/Class/Structure/aa/aa_explorer.cgi

1. Under the amino sequence below, write the same sequence using the 1-letter code.
   Met-Glu-Asn-Asp-Glu-Leu-Pro-Ile-Cys-Lys-Glu-Asp-Pro-Glu-Cys-Lys-Glu-Asp

2. What is the net charge of this peptide? (Assign -1 for each acidic amino acid and +1 for each basic amino acid. Add up the total charges.)

3. Using the Venn diagram above, propose a conservative substitution for:

   Trp - His - Arg - Leu -

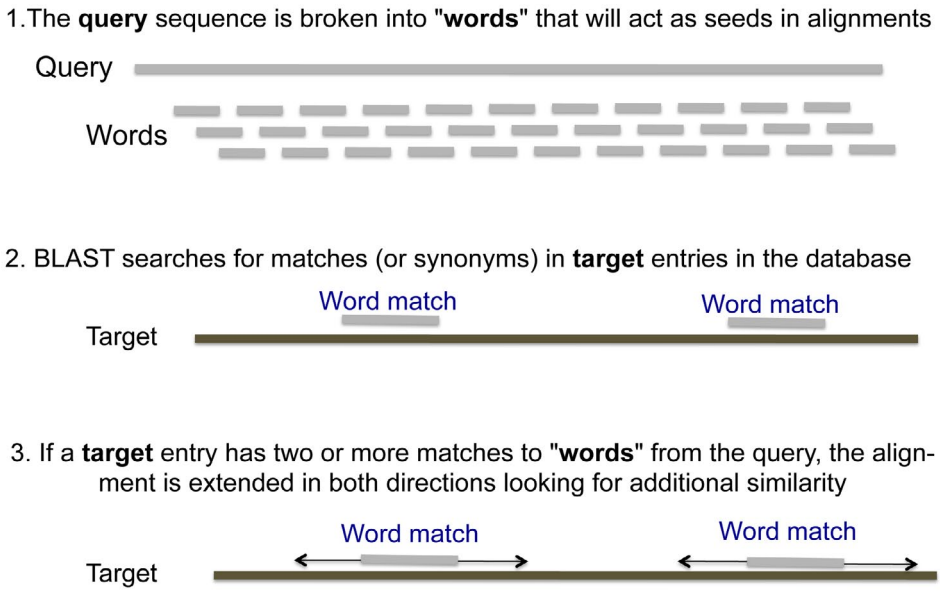
4. Write the name of a music group that you enjoy. Then transpose the name into an amino acid sequence written with the 3-letter code. Pass the amino acid sequence to a friend and have him/her decode it. (Note: the 1-letter code uses all of the alphabet, except B, J, O, U, X and Z).

Chapter 9

BLAST algorithms are used to search databases

There are many different algorithms for searching sequence databases, but BLAST algorithms are some of the most popular, because of their speed. As you will see below, the key to BLAST's speed is its use of local alignments that serve as seeds for more extensive alignments. In fact, BLAST is an acronym for Basic Local Alignment Search Tool (Altschul *et al.*, 1990). A set of BLAST tools for searching nucleotide and proteins sequences is available for use at the NCBI site. You have already used the BLASTN algorithm to search for nucleotide matches between PCR primers and genomic DNA (Chapter 7). In this lab, you will use the BLASTP algorithm to search for homologs of *S. cerevisiae* Met proteins in other organisms.

BLAST searches begin with a query sequence that will be matched against sequence databases specified by the user. As the algorithms work through the data, they compute the probability that each potential match may have arisen by chance alone, which would not be consistent with an evolutionary relationship. BLAST algorithms begin by breaking down the query sequence into a series of short overlapping "words" and assigning numerical values to the words. Words above a threshold value for statistical significance are then used to search databases. The default word size for BLASTN is 28 nucleotides. Because there are only four possible nucleotides in DNA, a sequence of this length would be expected to occur randomly once in every 4^{28}, or 10^{17}, nucleotides, which is far longer than any genome. The default word size for BLASTP is three amino acids. Because proteins contain 20 different amino acids, a tripeptide sequence would be expected to arise randomly once in every 8000 tripeptides, which is longer than any protein. The figure below outlines the basic strategy used by the BLAST algorithms.



Overview of the strategy used in BLAST algorithms
BLASTN and BLASTP use a rolling window to break down a query sequence into words and word synonyms that form a search set. At least two words or synonyms in the search set must match a target sequence in the database, for that sequence to be reported in the results.

In this lab, we will use the BLASTP algorithm, which is more useful than BLASTN for studying protein evolution. Unlike BLASTN, BLASTP overlooks synonymous gene mutations that do not change an amino acid. Synonymous substitutions do not affect the function of a protein and would therefore not be selected against during evolution. BLASTP uses a weighted scoring matrix, BLOSUM 62 (Henikoff & Henikoff, 1999), that factors in the frequencies with which particular amino acid substitutions have taken place during protein evolution.

We will return to this discussion of BLASTP after an introduction and chance to work with the BLOSUM62 matrix.

BLOSUM62 scoring matrix for amino acid substitutions

The results obtained in a BLASTP search depend on the scoring matrix used to assign numerical values to different words. A variety of BLOSUM (BLOcks SUBstitution Matrix) matrices are available, whose utility depends on whether the user is comparing more highly divergent or less divergent sequences. The BLOSUM62 matrix is used as the default scoring matrix for BLASTP. The BLOSUM62 matrix was developed by analyzing the frequencies of amino acid substitutions in clusters of related proteins. Within each cluster, or block, the amino acid sequences were at least 62% identical when two proteins were aligned. Investigators computationally determined the frequency of all 2^{10} possible amino acid substitutions that had occurred in these conserved blocks of proteins. They used this data to construct the BLOSUM62 scoring matrix for amino acid substitutions. The BLOSUM62 score for a particular substitution is a log-odds score that provides a measure of the biological probability of a substitution relative to the chance probability of the substitution. For a substitution of amino acid *i* for amino acid *j*, the score is expressed:

$$S_{ij} = (1/\lambda)\ log\left(\frac{p_{ij}}{q_i q_j}\right)$$

where p_{ij} is the frequency of the substitution in homologous proteins, and q_i and q_j are the frequencies of amino acids i and j in the database. The term (1/λ) is a scaling factor used to generate integral values in the matrix.

The BLOSUM62 matrix on the following page is consistent with strong evolutionary pressure to conserve protein function. As expected, the most common substitution for any amino acid is itself. Overall, positive scores (shaded) are less common than negative scores, suggesting that most substitutions negatively affect protein function. The most highly conserved amino acids are cysteine, tryptophan and histidine, which have the highest scores. Interestingly, these latter amino acids have unique chemistries and often play important structural or catalytic roles in proteins.

BLOSUM 62 scoring matrix

(positive values are shaded)

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	**4**																			
R	-1	**5**																		
N	-2	0	**6**																	
D	-2	-2	**1**	**6**																
C	0	-3	-3	-3	**9**															
Q	-1	**1**	0	0	-3	**5**														
E	-1	0	0	**2**	-4	**2**	**5**													
G	0	-2	0	-1	-3	-2	-2	**6**												
H	-2	0	**1**	-1	-3	0	0	-2	**8**											
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	**4**										
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	**2**	**4**									
K	-1	**2**	0	-1	-3	**1**	**1**	-2	-1	-3	-2	**5**								
M	-1	-1	-2	-3	-1	0	-2	-3	-2	**1**	**2**	-1	**5**							
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	**6**						
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	**7**					
S	**1**	-1	**1**	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	**4**				
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	**1**	**5**			
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	**1**	-4	-3	-2	**11**		
Y	-2	-2	-2	-3	-2	-1	-2	-3	**2**	-1	-1	-2	-1	**3**	-3	-2	-2	**2**	**7**	
V	0	-3	-3	-3	-1	-2	-2	-3	-3	**3**	**1**	-2	**1**	-1	-2	-2	0	-3	-1	**4**

The values for amino acid substitutions were obtained from Henikoff S & Henikoff JG (1992) Amino acid substitutions matrices from protein blocks. *Proc. Natl. Acad. Sci.* **89**: 10915-10919.

Exercise 2 - The BLOSUM62 matrix

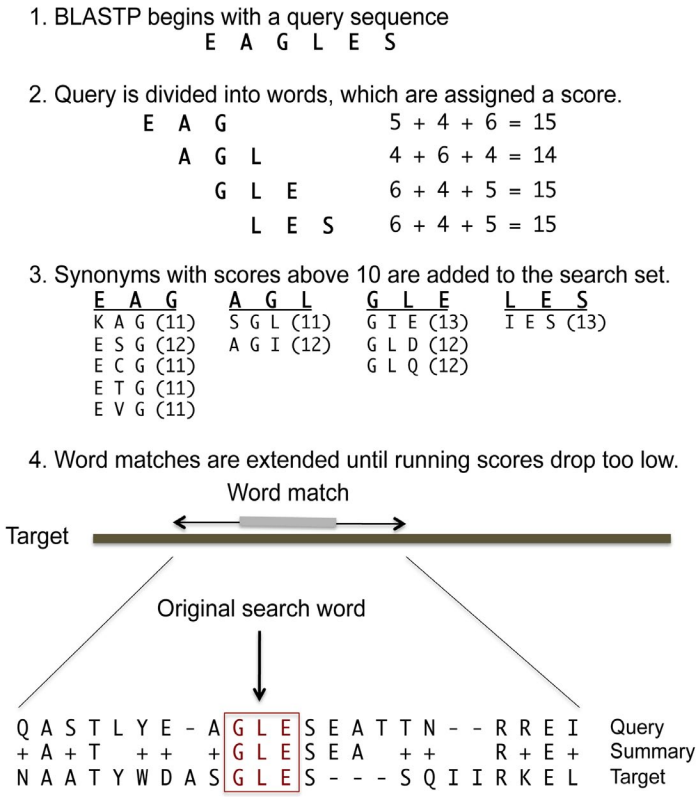
1. Find the BLOSUM scores for the conservative substitutions that you sggested in Exercise 1. Does the BLOSUM data support your hypotheses?

2. Find the two **substitutions** with the highest BLOSUM scores. In what ways are the biochemical properties of the substituted amino acid similar or dissimilar to the amino acid that it replaces?

3. Find the three amino acids for which there is no evidence of amino acid substitutions that have occurred more frequently than predicted by chance alone. What special features do these amino acids have?

The BLASTP algorithm

In BLASTP, the query sequence is broken into all possible 3-letter words using a moving window. A numerical score is calculated for each word by adding up the values for the amino acids from the BLOSUM62 matrix. Words with a score of 12 of more are collected into the initial BLASTP search set. BLASTP next broadens the search set by adding synonyms that differ from the words at one position. Only synonyms with scores above a threshold value are added to the search set. NCBI BLASTP uses a default threshold of 10 for synonyms, but this can be adjusted by the user. Using this search set, BLAST rapidly scans a database and identifies protein sequences that contain at two or more word/synonyms from the search set. These sequences are set aside for the next phase of the BLASTP process, where these short matches serve as seeds for more extended alignments in both directions from the original match. BLAST keeps a running raw score as it extends the matches. Each new amino acid either increases or decreases the raw score. Penalties are assigned for mismatches and for gaps between the two alignments. In the NCBI default settings, the presence of a gap brings an initial penalty of 11, which increases by 1 for each missing amino acid. Once the score falls below a set level, the alignment ceases. Raw scores are then converted into bit scores by correcting for the scoring matrix used in the search and the size of the database search space.



Overview of the BLASTP process.
The query sequence EAGLES into broken into three-letter words or synonyms that are used as a search set against records in a protein or translated nucleotide database. See the text for additional details.

The output data from BLASTP includes a table with the bit scores for each alignment as well as its E-value, or "expect score". The E-value indicates the number of alignments with that particular bit score that would be expected to occur solely by chance in the search space. Alignments with the highest bit scores (and lowest E-values) are listed at the top of the table. For perfect or nearly perfect matches, the E-value is reported as zero - there is essentially no possibility that the match occurs randomly. The E-value takes into account both the length of the match and the size of the database that was surveyed. The longer the alignment, and/or the larger the database search space, the less likely that a particular alignment occurs strictly by chance.

BLASTP provides additional data for each alignment. In some cases, the alignment may not extend along the entire length of the protein. The "Query cover" refers to the fraction of the query sequence where the alignment score is above the threshold value. BLASTP also reports the percentage of aligned amino acids that are identical in two sequences as "% Identity".

Exercise 3 - Using BLASTP

In this exercise, we will use BLASTP to find the *S. pombe* homolog for Met3p. Direct your browser to the NCBI site and click BLAST under Popular Resources on the right.

1. Choose Protein BLAST
2. Enter the NP_ number for S. *cerevisiae* Met3p in the query box (NP_012543)
3. Choose the records to be searched. For the database, select **reference proteins**. For the organism, type **Schizosaccharomyces pombe**. (This is taxid 4896 from the dropdown box.)
4. Expand the algorithm parameters. We will use the default word size of 3 with a threshold value of 10 and a gap penalty of 11. The search could be made more stringent by increasing the threshold value or gap penalty. The search could be made less stringent by decreasing the word size.
5. Click BLAST and wait for the results to appear.
6. Analyze the results page:
   - The **graphic summary** at the top gives you an instant overview about the extent and strength of the match with *S. pombe* sequences. The sequence is an excellent match, with an alignment score above 200 throughout all but the C-termini of the proteins.
   - The **summary table** provides the numerical data. Using an E-value of 1E-10 as a threshold for significance, there is a single high quality match with *S. pombe* sequences that extends over 87% of the sequence.
   - Cursor down to see the actual **alignment** between the sequences. Dashes have been introduced to either the S. *cerevisiae* or *S. pombe* sequence where gaps interrupt the alignments. The center row summarizes the homology between the protein sequences. If an amino acid is conserved between the two species, its 1-letter code is shown. Plus signs indicate conservative substitutions, *i.e.* substitutions with BLOSUM values of 1 or more.
   - Click on the link to the NP_ record for the *S. pombe* ortholog. Find its EC number (2.7.7.4) in the Features field, which identifies the protein as a sulfur adenylyltransferase. This is the same EC number as that of *S. cerevisiae* Met3p, suggesting that the two homologs catalyze the same reaction.

Exercise 4 - Multiple sequence alignments

BLASTP gives a pairwise alignment of sequences that is very useful for identifying homologs. Multiple sequence alignments compare a larger number of sequences simultaneously. By comparing a larger number of sequences over a wider evolutionary range, multiple sequence alignments allow researchers to identify regions of a protein that are most highly conserved, and therefore, more likely to be important for the function of a protein. In this exercise, we will study conservation of protein sequences in a number of model organisms that are widely used in genetic studies. The genomes for model organisms have been sequenced, and techniques for genetic analysis are well-developed. In addition, database and clone resources are available to support research with model organisms. The organisms below have been selected because they represent important branches of evolution and because they are potential candidates for future research in this course.

Bacteria - these represent two major divisions of the bacteria
Escherichia coli strain K-12 (gram negative; K-12 is the standard laboratory strain)
Bacillus subtilis strain 168 (gram positive reference strain)
Eukaryotes - model organisms
Saccharomyces cerevisiae - needs to be included in trees and alignments!
Schizosaccharomyces pombe
Arabidopsis thaliana - thale crress; model organism for flowering plants
Caenorhabditis elegans - nematode model organism used in developmental studies
Mus musculus - laboratory mouse

Collect the sequences and BLAST data

The first step in a multiple sequence alignment is to collect the sequence data and analyze the BLASTP data that compare the sequences with the *S. cerevisiae* sequence. We will be using the reference sequences for the organisms, which begin with a NP___ number. Since you already know how to find NP____ records and use BLASTP, we will take some shortcuts to finding the remaining numbers and BLASTP statistics. The accession numbers for the bacterial species will be available on the class website and in the lab. We will use the Homologene database at NCBI (Sayers *et al.*, 2012) to find the accession numbers for the eukaryotes. Homologene is a database that uses automated BLASTP searches to identify homologs in 21 sequenced eukaryotic genomes.

Access Homologene at: http://www.ncbi.nlm.nih.gov/homologene

Click on Release Statistics to see the species that have been included in the BLASTP searchers. Enter the name of your gene into the search box. This brings up the various Homologene groups that have a gene with that name. If search brings you to a page with more than one Homologene group list, click on the Homologene group that contains the *S. cerevisiae* gene.

Record the accession number for the Homologene group:

At the top of the record, Homologene describes the taxonomic distribution of homologs in eukaryotes ("Gene conserved in _____") A narrowly conserved protein might only be found in the Ascomycota, while a widely-districuted protein would be found in the Eukaryota.

What phylogenetic divisions have homologs of your gene?

The left column of each Homologene record has links to comprehensive gene summaries prepared by NCBI curators. The right column has links to the NP___ records. The Homologene output always provides a graphical depiction of protein domains in each of the homologs.

How many domains are found in the S. cerevisiae protein?

Are the domains equally well-conserved between species?

Record the NP___ numbers for *S. cerevisiae* and *S. pombe*, as well as the model organisms listed on the previous page. Depending on the phylogenetic distribution of your gene and the possibility of paralogs in a species, you will have different numbers of entries in the table. Add the NP_ numbers for *E. coli* and *B. subtilis* homologs from the posted data sheet. If you have less than five entries, *e.g.* the protein is narrowly restricted to Ascomycota, add two additional species of your choice from the Homologene group that contains your gene.

NOTE: Does the S. pombe homolog of your MET gene have a different name! You will need this information later in this chapter.

Species	**NP Accession #**	**Total score**	**Coverage**	**E-value**
S. cerevisiae				
S. pombe				

Next, perform a pairwise BLASTP alignment for each sequence against the *S. cerevisiae* sequence. Collecting BLASTP data is easy with Homologene: Use the grey box on the lower hand side of the page to set up each BLASTP comparison. Record the total score, % coverage and E-value for each match.

In the next step, you will prepare a multiple sequence alignment using the sequence information in the NP___ records. Using the BLASTP data, it may be possible to exclude some sequences from further study. The best matches will have high total scores and % coverage (fraction of the two proteins that are aligned) and low E-values. ***For the rest of this assignment, exclude sequences where the total score is less than 100 and E-values are higher than 1E-10.***

Prepare the multiple sequence alignment.

We will use the Phylogeny suite of programs to construct a multiple sequence alignment and phylogenetic tree. Phylogeny describes itself as providing "Robust Phylogenetic Analysis for the Non-Specialist." You will be working with material at two different sites, so you need two operational browser pages. One browser tab should remain at NCBI, where you will retrieve records. Direct the other browser page to ***http://www.phylogeny.fr***

1. Click the **Your Workspace** tab and set up an account to store your work.
2. Under the heading **Phylogeny Analysis** tab, select **One Click**. After you enter the data, your sequences will be automatically brought through multiple alignment and phylogenetic tree building algorithms. The advanced option on this page would allow you to adjust the parameters associated with each program. *We will let Phylogeny make these decisions for us!*
3. Enter the data in FASTA format. Do this by **pasting** individual sequences from NCBI directly into the Phylogeny data entry box. **IMPORTANT: Do NOT use a text editor or Word**, as these programs introduce hidden punctuation that interfere with Phylogeny programs.
4. Obtain a FASTA file for each sequence. Find the record in the NCBI Protein Database, using the NP_ number as a search term. (Alternatively, you can click to the NP_ record from the Homologene summary page.) When the record comes up, click the FASTA link at the upper left side of the record. Note that the title line of a FASTA file begins with a > character and ends with a hard return. The title lines of NCBI FASTA files contain a GenBank reference number (gi_____), the RefSeq accession number (NP_xxxxxx), protein name and the species name in brackets. ***Edit out all the characters except the species name.*** (You will see why later!) The first file in our comparisons will be the *S. cerevisiae* protein sequence. Copy the title line, beginning with > and the entire amino acid sequence. Paste the FASTA record into the Phylogeny data entry box
5. Repeat step 4 with each of the sequences that you would like to compare.
6. When you are finished, give your project a title, enter your email address and click the Submit button. Your results will be posted on a web page.

Export and print the multiple sequence alignment

1. Click on the **Alignment** tab to view the multiple sequence alignment.
2. Ask for the output in ClustalW format. The output appears on a new web page. Note that the bottom line of each cluster indicates if an amino acid is invariant at the position by an asterisk. The positions of conserved amino acids are indicated by colons in the bottom line.
3. Right-click on the page and download the Clustal alignment with a new filename that makes sense to you. The page will download as a text file that you will open in Word or a text editor.
4. Open the file in a word processor. Adjust the font size and page breaks so that sequences are properly aligned and all members of a cluster fit on the same page. Choose a non-proportional font such as Courier so that the amino acids line up properly.
5. Print the file and **check that the format is correct**! Turn it in with the Phylogeny assignment.

Chapter 9

Construct a phylogenetic tree.

1. Click the **Tree Rendering** tab to access your phylogenetic tree.
2. Use the editing tools to alter the appearance of your tree. Pay particular attention to the legends in the "leaves" of the tree, which should have the species names.
3. Download the file in a format of your choice. Print the file and turn it in with the phylogeny assignment.

Investigate S. pombe homologs using Pombase

Like *S. cerevisiae*, *S. pombe* is a model organism with its own large community of researchers. Pombase serves as the central database for information on *S. pombe*, functioning much like the SGD. Access Pombase at ***http://www.pombase.org***

- Enter the name for the *S. pombe* homolog that you obtained in Homologene.
- Record the systematic name for your gene, which refer to the position of the gene on the chromosome. Systematic names begin with SPAC, SPBC or SPCC, corresponding to chromosomes I, II, and III, respectively.
- Pombase stores all the information on a single page that is divided into fields. Individual fields can be expanded or minimized with an arrow by the field name. Quick links on the lower right also help with navigation.
- Navigate to the **Transcipt** field. The graphic will indicate whether the homolog contains an intron. The field also contains information about the exon/intron boundaries and information on 5'- and 3'-UTRs in mRNAs, when that information is available. Does the homolog to your *MET* gene contain introns?
- Spend a little time seeing what kind of information (*e.g.* gene/protein size, function, protein interactions) is available for your *S. pombe* ortholog. You may find Pombase to be a helpful resource when writing lab reports.

References

Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool *J. Mol. Biol.* **214**: 403-410.

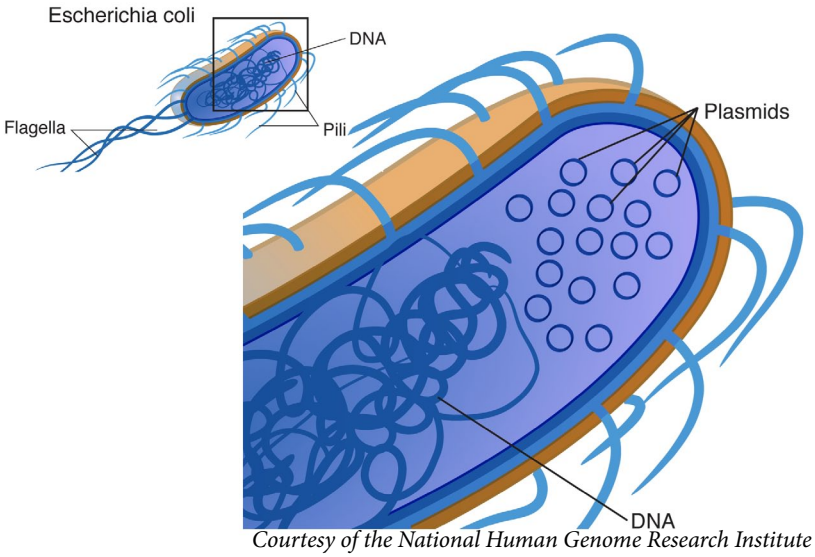
Henikoff S, and Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* 8**9**: 10915-10919.

Kellis M, Birren BW & Lander ES (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428:** 617-624.

Sayers EW, Barrett T, Benson DA *et al.* (2012) Database resources of the National Center for Biotechnology Information. *Nucl Acids Res* **40:** D1-D25. (Note: this is an online publication that is updated annually.)

Chapter 10

Plasmids



Courtesy of the National Human Genome Research Institute

Plasmids are small, circular pieces of DNA that replicate independently of the host chromosome. Plasmids have revolutionized molecular biology by allowing investigators to obtain many copies of custom DNA molecules. In this lab, you will isolate plasmids from non-pathogenic strains of *Escherichia coli*, which you will use in subsequent experiments to transform *S. cerevisiae met* strains.

Objectives

At the end of this laboratory, students should be able to:

- describe the structure of plasmids and their mechanism of replication.
- identify functional elements that have been engineered into laboratory plasmids.
- explain how the physical properties of plasmids are used in their purification.
- isolate plasmids from transformed strains of *Escherichia coli*.

Chapter 10

Plasmids are the workhorses of molecular biology. Plasmids are small, circular DNA molecules that replicate independently of the chromosomes in the microorganisms that harbor them. Plasmids are often referred to as vectors, because they can be used to transfer foreign DNA into a cell. The plasmids used in molecular biology have been constructed by molecular biologists, who used recombinant DNA technology to incorporate many different functional elements into naturally-occurring plasmids. Plasmids have been engineered to carry up to 10 kb of foreign DNA and they are easily isolated from microorganisms for manipulation in the lab. For the next few labs, your team will be working with yeast overexpression plasmids. Your team will work with three plasmids: a plasmid carrying an *S. cerevisiae MET* gene, a plasmid carrying its *S. pombe* homolog, and a plasmid carrying the bacterial *lacZ* gene, which will act as a negative control. In this lab, you will isolate plasmids from a bacterial culture. In the next few weeks, you will characterize the plasmids and then use them to transform mutant yeast strains, testing whether *MET* gene function has been conserved between *S. cerevisiae* and *S. pombe*.

Plasmids are composed of functional elements

Plasmid replication depends on host cell polymerases

Plasmids are found naturally in many microorganisms. Plasmids can be transferred between species by transformation or conjugation, but they generally have a restricted host range. When you think of plasmids, you probably also think of bacteria, but plasmids are not restricted to bacteria. In fact, most *S. cerevisiae* strains carry a large plasmid known as the 2 micron or 2 μm plasmid. Multiple copies of the 2 μm plasmid are usually present in the nucleus of a yeast cell, and the plasmid number is stable through many rounds of cell division.

Although plasmids replicate independently of the chromosomal DNA, they rely on host enzymes to catalyze their replication. Host DNA polymerases bind to an origin of replication (*ori*) sequence in the plasmid. Plasmids that replicate in bacteria have *ori* sequences that bind bacterial DNA polymerase, while plasmids that replicate in yeast have distinct *ori* sequences that bind yeast DNA polymerase. The plasmids that we are using are sometimes referred to as "shuttle vectors," because they are able to replicate in more than one kind of cell. Our plasmids contain the *ori* of plasmid, pBR322, which is replicated in *E. coli* to a copy number of 30-40. The plasmids also contain the *ori* of the *S. cerevisiae* 2 μm plasmid described above. In this class, we will propagate the shuttle vectors in bacteria, because bacteria grow more rapidly than yeast and because the yield of plasmid from bacteria is higher than the yield from yeast. We will harvest the plasmids from bacteria and then use them to transform yeast cells.

Laboratory plasmids carry selectable markers

Plasmids place a toll on the host cell's metabolism, and they would normally be lost from their host cells if they did not confer some selective advantage to the host. The plasmids used in molecular biology therefore carry genes for selectable markers, which allow transformed cells to grow under conditions where untransformed cells are unable to grow. Our plasmids contain the β-lactamase (amp^R) gene, which allows *E. coli* to grow in the presence of ampicillin, an antibiotic

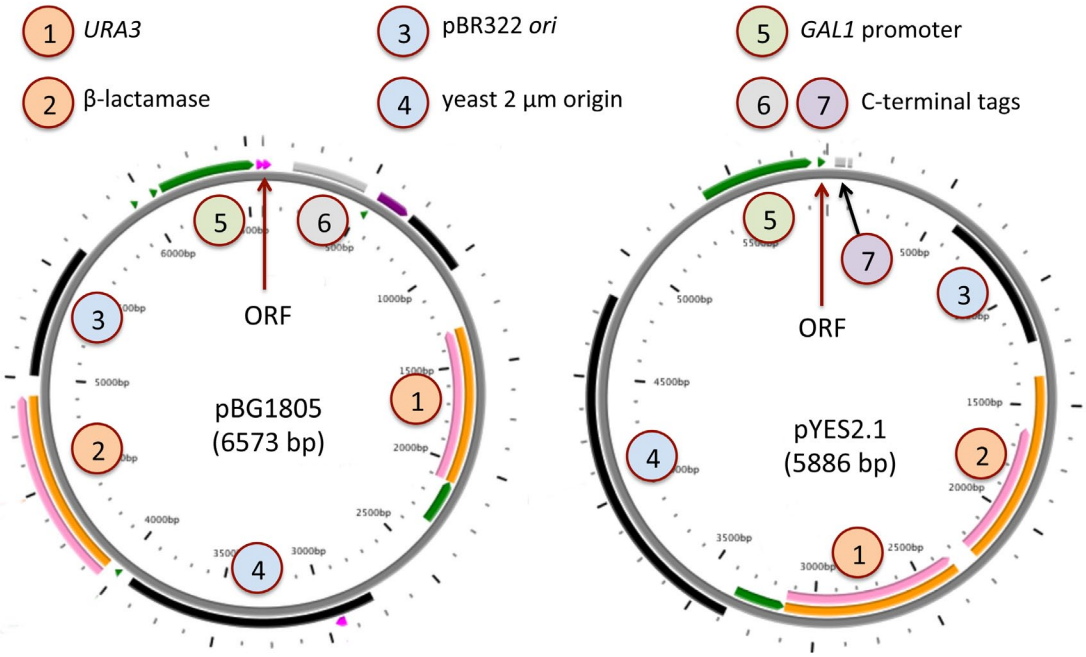
that interferes with cell wall synthesis in bacteria. The plasmids also contain the *S. cerevisiae URA3* gene, which allows *ura3* mutants like BY4742, the parent strain of our mutants, to grow in the absence of uracil after they are transformed with plasmids (Chapter 12).

Promoters control transcription of plasmid sequences

The plasmids that we will use this semester contain *MET* genes that have been cloned into plasmids directly downstream of the promoter sequence for the yeast *GAL1* gene (Johnston, 1987). Transcription from the *GAL1* promoter is normally controlled by regulatory proteins that sense glucose and galactose levels in yeast (Chapter 13). In the plasmids, the *GAL1* promoter has been placed at the 5'-ends of protein coding sequences for *S. cerevisiae* Met proteins, their *S. pombe* homologs or bacterial LacZ. The presence of the *GAL1* promoter will allow you to manipulate expression of the Met proteins or LacZ in transformed yeast cells.

Comparison of yeast overexpression vectors

The figure below compares the plasmids that you will be using to overexpress *S. cerevisiae* and *S. pombe* proteins. The plasmids have many similarities, but some significant differences. ORFs for the *S. cerevisiae* proteins were cloned into the plasmid pBG1805 cloning vector in a genome-wide experiment (Gelperin *et al.*, 2005). ORFs for the *S. pombe* proteins were individually cloned into the pYES2.1 plasmid by the BI204 staff. In all cases, the ORFs were cloned downstream of the *GAL1* promoter (element 5 in the diagram). Both plasmids contain the yeast *URA3* gene (element 1) and the bacterial β-lactamase gene (element 2). Both plasmids also contain the pBR322 *ori* sequence for bacterial replication and the 2 μm *ori* sequence for yeast replication. The placement of the functional elements in the two plasmids is quite distinct, consistent with their independent development by different laboratories.



As you can see from the figure on the previous page, the plasmids also differ in the C-terminal tags that are added to overexpressed proteins. The proteins expressed from both plasmids are fusion proteins that are longer than the natural coding sequences. During the cloning processes used to construct the overexpression plasmids, researchers deleted the natural stop codons of the ORFs so that transcription would continue into plasmid-encoded sequences. The plasmid sequences encode tags that can be used for procedures such as western blots (Chapter 15) or protein purification. The pBG1805 plasmid adds a 19 kDa extension (element 6) to expressed proteins, while the pYES2.1 plasmid adds a smaller 5 kDa extension (element 7). The tags will be discussed in greater detail in Chapter 13.

Plasmid nomenclature

Proper nomenclature is important for distinguishing plasmids. The "p" in pBG1805 denotes that it is a plasmid, while the remainder of the plasmid name is a code used by the researchers who constructed the plasmid. Often, the letters in a plasmid's name contain the initials of the researcher who performed the final step in its construction. In this case, "BG" refers to Beth Greyhack, who was one of the senior authors on the paper (Gelperin *et al.*, 2005). The pYES2.1 plasmid was developed by researchers working for a commercial source. In this course, we will follow the convention of denoting the plasmid backbone in normal font, followed by a hyphen and then the name of the cloned ORF in italics. In our experiments:

pBG1805-*MET1* would designate the *S. cerevisiae MET1* gene cloned into pBG1805.

pYES2.1-*Met1* would designate the *S. pombe Met1* gene cloned into pYES2.1. (Recall from Chapter 6 that *S. cerevisiae* genes are unusual in using 3 capital letters for their names.)

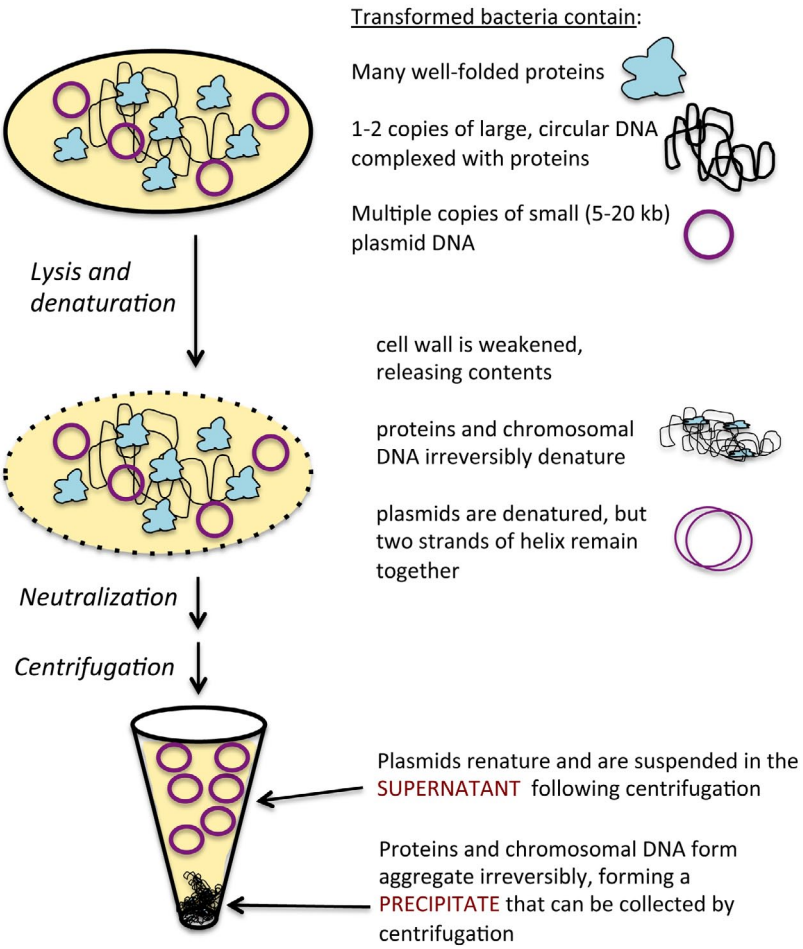
pYES2.1 -*lacZ*$^{+}$ would designate the bacterial LacZ gene cloned into pYES2.1.
(*Note: bacterial gene names are not capitalized. Plus sign indicates wild type gene.*)

NOTE: The S. pombe homolog of an S. cerevisiae gene may not share the same gene number. Many of the *S. pombe* genes received their names due to their homology to previously identified *S. cerevisiae* genes, but some *S. pombe* genes had been named before their DNA sequences were known. For example, the ortholog of *S. cerevisiae MET2* in *S. pombe* is *Met6*.

Plasmids are easily isolated from bacterial cells

Plasmid isolation takes advantage of the unique structural properties of plasmids. Plasmids are small, supercoiled circular pieces of DNA. Unlike the much larger bacterial chromosome (which is also circular), plasmids are quite resistant to permanent denaturation. Today, most laboratories use commercial kits for plasmid isolations, because the kits are convenient and relatively inexpensive. The kits give good yields of high-quality DNA, while avoiding the need for organic denaturatants. A variety of less expensive, but somewhat more time-consuming, procedures have been described for investigators who want to make their own reagents. These procedures generally give good yields of DNA that is slightly less pure than

DNA purified with the kits. Whatever the isolation procedure, the general principles of plasmid isolation are the same. The figure and paragraphs below summarize the steps and general principles used for plasmid isolation.



1. ***Lysis and denaturation*** - Strong denaturating conditions are used to weaken the tough bacterial cell wall. The most common procedures use a combination of strong base and a detergent. The detergents help to solubilize lipids in the cell wall, allowing the denaturants to enter the cell. Proteins, because of their fragile structures, are irreversibly denatured. The treatment also breaks the hydrogen bonds holding together the two strands of DNA helices.
2. ***Neutralization*** - Neutralization allows complementary DNA strands to reanneal and causes proteins to precipitate. Plasmids renature because they have supercoiled structures that have held the two strands of the helix together during denaturation. Chromosomal DNA is not able to renature, however, because its longer strands have become mixed with denatured proteins. ***Samples must be mixed gently at this step to prevent fragmentation of the long, chromosomal DNA into pieces that might be able to reanneal and co-purify with the plasmids.***
3. ***Centrifugation*** - Plasmid DNA is separated from large aggregates of precipitated proteins and chromosomal DNA by centrifugation.
4. ***Additional purification*** - Plasmids are further purified by organic extraction or adsorption to a resin.

Chapter 10

The plasmids that we are working with in this lab have been maintained in a laboratory strain of *Escherichia coli*. *E. coli* is a proteobacterium that normally inhabits the intestinal tract of warm-blooded mammals. The virulent strains of *E. coli* that appear in the news have acquired, often by lateral gene transfer, pathogenicity islands containing genes for virulence factors, toxins and adhesion proteins important for tissue invasion. Pathogenicity islands are not present in lab strains, which are derivatives *E. coli* strain K12. K12 is a debilitated strain that is unable to colonize the human intestine. It has been propagated and used safely in the laboratory for over 70 years. The *E. coli* strains used in molecular biology have also been selected to tolerate large numbers of plasmids.

We will use the Zyppy™ (Zymo Research) kit to purify plasmids from the transformed *E. coli* strains. The final purification step in the procedure involves a spin column of silica resin. Nucleic acids absorb strongly to silica in the presence of high concentrations of salt. Following a wash step that removes any residual proteins, the plasmids are eluted with a low salt solution. Plasmid solutions are very stable and can be stored for long periods of time in the refrigerator or freezer.

Exercise 1 - Plasmid isolation with the Zyppy™ kit

Concentrate the plasmid-bearing bacterial cells

1. Collect the three bacterial cultures that your group has been assigned. The bacteria have been transformed with plasmids carrying either an *S. cerevisiae MET* gene, its *S. pombe* homolog or bacterial *LacZ*. Each culture contains 1.5 mL Luria Bertani (LB) medium with 100 µg/mL ampicillin. Cultures were grown overnight at 37 degrees C, and the cell density is expected to be 3-4 X 10^9 cells/mL.

   What is the purpose of the ampicillin? How does it work?

2. Concentrate the cells in the 1.5 mL culture by centrifuging the cultures at maximum speed (~14,000 rpm) for 1 min. The cells will form a white pellet at the bottom of the tube.
3. Use a P200 to aspirate off as much of the culture medium as possible.

Alkaline lysis of bacterial cells harboring the plasmids

4. Re-suspend the pellet in 600 µl of TE buffer (Tris-HCl, EDTA pH=8.0) using the vortex mixer.
5. Add 100 µL of 7X Blue Zyppy Lysis buffer to the tube. Mix the buffer and cells by gently inverting the tube 4-6 times. Be gentle! Too much mechanical stress could fragment the bacterial chromosomal DNA and contaminate your plasmid preparation. The solution should turn from a cloudy blue to a clear blue.

   NOTE: ***This step is time-sensitive!! Proceeed to the next step within 2 minutes.***

Separate plasmid DNA from denatured proteins and chromosomal DNA

6. Add 350 µL of cold Yellow Zyppy Neutralization buffer (w/RNAase A) to the tube, and mix the contents thoroughly by inverting several times. The solution will turn yellow when neutralization is complete, and a yellowish precipitate will form. Invert the sample an additional 3-4 times to ensure complete neutralization. ***Be sure there is no more blue color.***
7. Centrifuge the mixture at maximum speed for 3 minutes to remove denatured proteins and chromosomal DNA. Notice that the tube contains a yellow precipitate that has collected on one side of the tube. The pale yellow supernatant contains the plasmid DNA.

Purify plasmid DNA by adsorption to a silica resin.

8. Using a pipette, carefully transfer the pale yellow supernatant (~900 µL) onto a Zyppy spin column. Be careful not to transfer any of the yellow precipitate!
9. Place the column with the collection tube attached into a centrifuge and spin at maximum speed for about 15 seconds. It is best to use the "pulse" button on the centrifuge and count to 15 or 20 seconds for this centrifugation step.
10. Remove the column and discard the flow through in the collection tube.
11. Place the column back into the collection tube and add 200 µL of Zyppy Endo-Wash solution. (Endo-Wash contains guanidine hydrochloride and isopropanol, which will remove denatured proteins from the resin.)
12. Centrifuge for 15-20 seconds, and discard the flow through.
13. Place the column back into the collection tube then add 400 µL of Zyppy Column Wash buffer. (This steps removes contaminating salts.) Centrifuge for 30-40 seconds. Empty the collecting tube.
14. Repeat the centrifugation step to remove any residual ethanol.

Elute the plasmid DNA

15. Transfer the Zyppy column to a clean (and appropriately labeled) 1.5 mL centrifuge tube, leaving the lid of the tube open.
16. Carefully, add 100 µL of TE buffer directly on top of the white column bed. Place the pipette tip as close as you can to the white column bed without poking it. Slowly dispense the water on top of the resin bed.
17. Allow the buffer to percolate into the column by letting the column sit in the microcentrifuge tube for 1 minute.
17. Centrifuge the column at maximum speed for 30 seconds. Again, it's fine to leave the cap open during this spin.
18. Remove the column, cap the tube and place it on ice. This tube should now contain plasmid DNA. ***Label the tube. SAVE the DNA for future experiments.***

Chapter 10

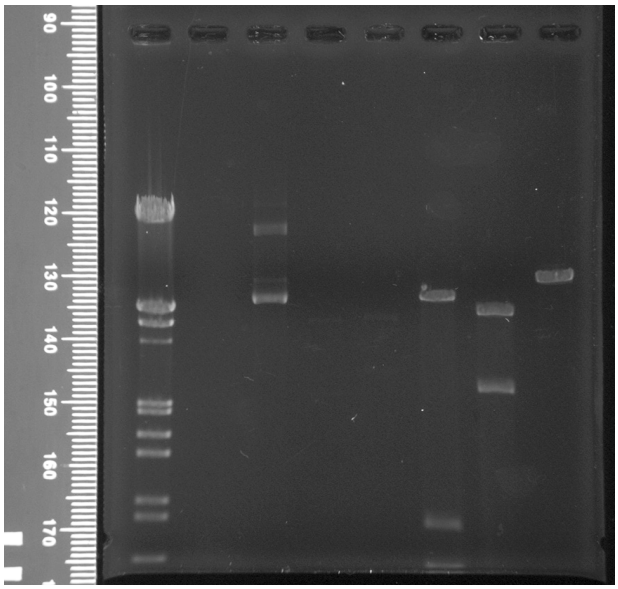
References

Gelperin DM, White MA, Wilkinson ML *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

Johnston M (1987) A model fungal regulatory mechanism: the *GAL1* genes of *Saccharomyces cerevisiae*. *Microbiol Rev* **51:** 458-476.

Chapter 11

Restriction mapping



Restriction endonucleases (REs) are part of bacterial defense systems. REs recognize and cleave specific sites in DNA molecules. REs are an indispensable tool in molecular biology for both analyzing and constructing DNA molecules. In this lab, you will prepare restriction digests to distinguish which of your plasmids carry the *S. cerevisiae* or *S. pombe* ORFs or the bacterial *lacZ* gene.

Objectives

At the end of this lab, students will be able to:

- describe the biological origins and functions of REs.
- use online tools to identify recognition sites for REs in a DNA molecule.
- devise a strategy to distinguish DNA molecules by selecting REs to use in DNA digests.
- interpret the patterns of restriction fragments separated on agarose gels.

Chapter 11

In the last experiment, your group isolated three different plasmids from transformed bacteria. One of the three plasmids, carries the *S. cerevisiae MET* gene that has been inactivated in your yeast strain. This *MET* gene was cloned into the pBG1805 plasmid (Gelperin *et al.*, 2005). A second plasmid carries the *S. pombe* homolog for the *MET* gene, cloned into the the pYES2.1 plasmid. The third plasmid is a negative control that contains the bacterial *lacZ*$^{+}$ gene cloned into pYES2.1. In this lab, your team will design and carry out a strategy to distinguish between the plasmids using restriction endonucleases. In the next lab, you will separate the products of the restriction digests, or restriction fragments, by agarose gel electrophoresis, generating a restriction map.

Restriction endonucleases

Bacterial restriction/modification systems protect against invaders

The discovery of restriction enzymes, or restriction endonucleases (REs), was pivotal to the development of molecular cloning. REs occur naturally in bacteria, where they specifically recognize short stretches of nucleotides in DNA and catalyze double-strand breaks at or near the recognition site (also known as a restriction site). To date, thousands of REs with distinct specificities have been described. You might wonder why bacteria harbor these potentially destructive enzymes. REs are part of a bacterial defense system against foreign DNA, such as an infectious bacteriophage. The bacterial DNA is protected from cleavage because it has been modified by a methyltransferase that modifies RE recognition sites in the bacterial DNA. The combined activities of the endonuclease and methyltransferase are referred to as a restriction/modification system. Today, most commercially available REs are not purified from their natural sources.. Instead, REs are usually isolated from bacteria that overexpress large quantities of REs from plasmids. These recombinant REs have often been engineered by molecular biologists to include amino acid changes that increase the catalytic activity or stability of the RE.

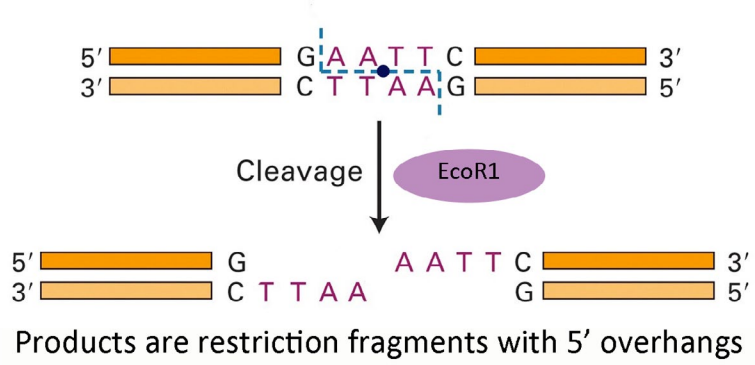
To understand how REs work, we will use EcoRI, one of the best-studied REs, as an example. Although the names of individual REs may sound a bit like baby talk, the nomenclature is actually very systematic and is based on its biological source. EcoRI is found naturally in the RY13 strain of *Escherichia coli*. Its name begins with the genus and species (Eco for *E. coli*), followed by a strain identifier (R for RY13), and ends with a Roman numeral that distinguishes the different REs found in the strain. Strain RY13 of *E. coli* contains multiple REs, but only EcoRI and EcoRV, are widely used in molecular biology.

Restriction enzymes cleave specific sites in DNA

Restriction enzymes like EcoRI are frequently called 6-cutters, because they recognize a 6-nucleotide sequence. Assuming a random distribution of A, C, G and Ts in DNA, probability predicts that a recognition site for a 6-cutter should occur about once for every 4000 bp (4^6) in DNA. Of course, the distribution of nucleotides in DNA is not random, so the actual sizes of DNA fragments produced by EcoRI range from hundreds to many thousands of base pairs, but

the mean size is close to 4000 bp. DNA fragments of this length are useful in the lab, since they long enough to contain the coding sequence for proteins and are well-resolved on agarose gels.

EcoRI recognizes the sequence G A A T T C in double stranded DNA. This recognition sequence is a palindrome with a two-fold axis of symmetry, because reading from 5' to 3' on either strand of the helix gives the same sequence. The palindromic nature of the restriction site is more obvious in the figure below. The dot in the center of the restriction site denotes the axis of symmetry. EcoRI catalyzes the hydrolysis of the phosphodiester bonds between G and A on both DNA strands. The restriction fragments generated in the reaction have short single-stranded tails at the 5'-ends. These ends are often referred to as "sticky ends," because of their ability to form hydrogen bonds with complementary DNA sequences.



EcoRI catalyzes the cleavage of a palindromic recognition site.

The recognition site for EcoRI has a two-fold axis of symmetry. Cleavage generates two fragments with 5′-sticky ends.

REs are sometimes referred to as molecular scissors because of their ability to generate restriction fragments that terminate with defined sequences. These "sticky ends" are important for recombinant DNA technology, because they enables researchers to construct designer DNA molecules. Any two DNA molecules with compatible sticky ends can be joined together by DNA ligases that serve as the "paste" by resealing broken phosphodiester bonds. We will not be generating recombinant molecules in this class, but it is important to understand their importance to modern biology. Consider the pBG1805 and pYES2.1 plasmids. From the plasmid maps in Chapter 10, you can see that these complex plasmids were constructed by stitching together DNA sequences from evolutionary distinct sources.

DNA molecules have unique restriction maps

The sequence of a DNA molecule determines the distribution of recognition sites for REs. Hundreds of REs with unique specifities have been described, so researchers can use the distribution of these recognition sites in a DNA molecule to construct a "map" of the sequence. In these experiments, DNA samples are digested with various REs to produce a restriction digest, a collection of smaller restriction fragments that have been cleaved at either end by the RE. The molecules in the digest are then separated by agarose gel electrophoresis (Chapter 8). From the sizes of the restriction fragments that are resolved on the gel, investigators are able to identify the original DNA molecule used in the restriction digest.

Chapter 11

Careful planning is required for meaningful restriction maps. The first step in a mapping experiment is to identify the sizes of restriction fragments that will be generated from a target DNA molecule with different REs. A variety of software programs generate these restriction maps and provide tabular data with details about the lengths and positions of the restriction fragments in the DNA sequence. The list of enzymes that cut a particular sequence is always impressive, but only a few enzymes usually turn out to be practical for the purpose of the experiment. When choosing REs for a restriction map, there are many things to consider:

- *How many restriction fragments will be generated?*
- *What are the predicted sizes of the restriction fragments?*
- *Will all the restriction fragments be clearly resolved on 1% agarose gels?*
- *Will the RE generate a distinctive set of fragments from each DNA sample?*
- *How expensive is the RE?*

In this lab, you will use the NEB Cutter to identify REs sites in plasmid sequences. (NEB Cutter is provided by New England Biolabs, a commercial supplier of REs.) You will recall that plasmids are supercoiled circles. Digestion with a RE opens up a plasmid and relaxes its structure. (Without RE digestion, the apparent sizes of plasmids on agarose gels are unreliable.) The plasmids that we are using for our experiments are complex plasmids based on pYES2.1 (5886 bp) or pBG1805 (6573 bp). Search the results for REs that will generate clearly distinguishable restriction fragments from your plasmids. ***It is highly recommended that you select the same RE for all three digests!*** Since two plasmids are based on pYES2.1, it would not be surprising to observe some common restriction fragments in those digests, which could be a helpful diagnostic.

Handling restriction endonucleases in the laboratory

The REs that we are using in the lab are highly purified (and expensive!) proteins that have been purified from recombinant bacteria. Like all enzymes, each RE functions optimally under a defined set of reaction conditions, including temperature, pH, and the concentrations of metal ions and salts. The manufacturer of our REs has developed buffers that support high levels of activity for more than 200 REs. Each buffer contains 0.1 mg/mL bovine serum albumin (BSA), an abundant protein from cow serum, which helps to stabilize denaturation-prone REs and to prevent nonspecific absorption of REs to test tubes and tips.

Like all enzymes, REs are subject to spontaneous denaturation, so REs need to be handled with care. (By comparison, DNA is an exceptionally stable molecule.) The rate of protein denaturation increases with temperature and at air/water interfaces. Some simple precautions will minimize denaturation. Follow these simple rules when you prepare the restriction digests:

- Use the recommended buffer for a particular RE.
- Keep the reactions on ice until the incubation begins.
- Be careful not to introduce bubbles. Do not use the vortex mixer.
- Add the RE last, after the other components of the reaction mixture have been assembled.

Exercise 1 - Plan the restriction digest

Assign each person in your group a different plasmid to analyze.

You will first need to assemble the complete sequences of your overexpression plasmids by combining the plasmid and *MET* gene sequences. Recall that the *S. cerevisiae* genes have been cloned into the pBG1805 vector and that the *S. pombe* genes and *LacZ* have been cloned into the pYES2.1 vector. You will then generate a restriction map that can be used to predict restriction fragments generated in an RE digestion. To generate the map, we will use one of several online tools that is available at the website of New England Biolabs, a commercial supplier of REs.

Use the table below to calculate the length of your plasmid as you complete the first few steps of this exercise.

Plasmid	**Vector size (bp)**	**Inserted CDS (bp)**	**Final length (bp)**
pBG1805-*MET*___	6573		
pYES2.1-	5886		
pYES2.1-*LacZ*	5886	3078	8964

1. ***Locate the coding sequence of your gene.***
   - Open the record for your *MET* gene or its homolog at SGD or Pombase, respectively.
   - Find the number of amino acids in your protein. Multiply this number by 3 to find the number of nucleotides in the CDS. Add this number to the table above.
   - Find the CDS sequence for your gene. You will paste this sequence to the end of the plasmid sequence in step 2.
   - The pYES2.1- *lacZ* sequence is posted on Canvas, so steps 1 and 2 are already done for this plasmid. Note that plasmid *lacZ* gene has been modified from naturally-occurring *lacZ* genes.

2. ***Assemble the complete nucleotide sequence of your plasmid.***
   - Open the Word document containing the pBG1805 or pYES2.1 vector sequence posted on Canvas. (Alternatively, you can find the sequence record for pBG1805 in the NCBI Nucleotide database. The accession number is JN560956.) The plasmid sequences are numbered so that the *GAL1* promoter is at the 3'-end of the DNA sequence.
   - Copy the CDS from SGD or Pombase and paste it at the end of the plasmid sequence.
   - Delete the last three nucleotides of the CDS, which comprise the gene's stop codon. The overexpression plasmids are designed to encode fusion proteins with C-terminal extensions. (Note: the stop codon in the *LacZ* sequence has been removed by the manufacturer.)

3. ***Prepare a restriction map of the complete plasmid sequence.***
   - Paste the sequence from step 2 into the search box in the NEBCutter tool:

     tools.neb.com/NEBcutter2/

   - **Check the box to indicate that the plasmid is CIRCULAR, rather than linear.**
   - You might also want to give your plasmid a name. The NEB site will store your queries for 24 hours, which can be very convenient. Click submit.
   - The search tool will return results for a bewildering number of REs. The vast majority of the RE sites are not useful, because the fragments are too large or too small, the enzyme is not available in the lab, or the endonuclease is sensitive to DNA methylation (which can be unpredictable).

4. ***Perform custom digests with enzymes that look promising.***
   - Click the custom digest link. This brings up a chart of REs that cut the plasmid, their recognition sites, the number of recognition sites, and the amount of enzyme activity in each of four buffers.
   - Analyze the fragments that would be produced by the four REs below by checking the box and clicking the green Submit button.

     AccI HincII ScaI XbaI

     (Note: There may be some changes to this list of available REs before the lab.)

5. ***Prepare a table summarizing the restriction maps for your three plasmids.***
   - Complete the table below, indicating the sizes of the restriction fragments generated with each RE.
   - Include the total length of the plasmid in the table. The sum of the restriction fragment lengths should sum up to this number.

Plasmid name			
Length (bp)			
Restriction Enzyme	*(Fragment lengths)*	*(Fragment lengths)*	*(Fragment lengths)*

6. ***Choose a RE that distinguishes your three plasmids.***

   The team should use the data table above to select the RE that best allows you to distinguish the three plasmids. We will be analyzing the restriction fragments on 1% agarose gels, which do a good job of resolving fragments ranging in size from ~500 bp to ~5000bp. Refer to the figure in Chapter 8, which shows the distribution of molecular size markers on 1% agarose gels. Choose an RE that will produce restriction digests with a nice range of fragment sizes.

Exercise 2 - Set up the restriction digests

The concentrations of RE and plasmid DNA need to be matched in a restriction digest. Manufacturers assay the activity of each batch of RE and express the activity in units of enzyme activity per µL. A unit of activity (U) is assessed in a standardized assay for measuring RE concentrations. Restriction digests are usually set up to contain at least 2-5 U per µg plasmid DNA. The Zyppy™ kits typically yields plasmid concentrations ranging from 10 to 30 ng/µL. (You will be able to estimate your plasmid DNA concentrations when you run the agarose gels in the next lab.) In this lab, we are using 7 µL of plasmid miniprep DNA in each reaction and 1 U of RE. This should be more than enough RE to ensure complete digestion of the plasmid DNA.

In your lab notebook, note which RE(s) you have decided to use.

- Prepare a separate tube for each of your RE digests.
- Combine the following components in each tube in order listed:

  7.0 µL plasmid
  1.0 µL 10X CutSmart™ buffer or 10X Buffer 3.1 (for HincII digests only)
  2.0 µL (1.0 U) restriction enzyme

  The total reaction volume should be 10 µL.

- Ensure that the components of each reaction are well-mixed by centrifuging them for a few seconds in the microcentrifuge.
- Incubate the samples at 37 °C for at least 2 hr.
- Store the reactions in the freezer.

Exercise 3 - Analyze the restriction digests on agarose gels

This exercise will be performed in the next laboratory session.

1. Plan your gel. Each group of students will prepare one agarose gel with 8 sample lanes.

- Each student will run one lane with undigested plasmid and a second lane with plasmid that has been digested with RE. It is important to run a lane with undigested plasmid to determine if the REs have effectively cleaved the plasmids. *Keep in mind that undigested plasmids will electrophorese with anomalous sizes because of their supercoiled structures.*
- One lane should be reserved for molecular weight standards.
- Record in your notebook how you plan to load your gel.

2. Prepare a 1% agarose gel in TAE buffer as described in Chapter 8.

3. Prepare the samples for electrophoresis. Use the entire restriction digest on the gel. For undigested plasmid samples, set up tubes containing 7 µL plasmid and 3 µL deionized water.

IMPORTANT: Return any unused plasmid to the freezer.
You will use the plasmids for yeast transformation in a later lab.

Combine:

10 µL restriction digest or undigested plasmid
5 µL deionized water
3 µL 6X sample buffer

4. Load and run the agarose gel as described in Chapter 8.

5. Estimate the DNA concentrations in your plasmid preparations. After the gel has been run and stained, attempt to estimate the DNA concentrations in your three plasmid preparations, using the size standards as a reference. Each of the markers, with the exception of the 3 kb marker, contains ~40 ng. The 3 kb band contains 125 ng. Use these values to visually estimate the amount of DNA in your lane. Correct for the volume of sample in the lane to calculate the concentration of DNA in each plasmid preparation. (Although imprecise, this value will be useful for calculating your transformation efficiency in the next lab.)

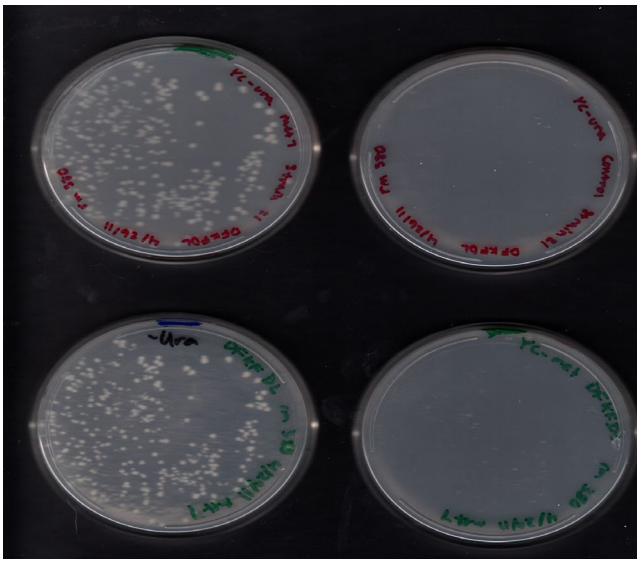
Test yourself

The pBG1805 plasmid was derived from an earlier yeast shuttle vector, pRS426. The pRS426 plasmid is 3419 bp long and has restriction sites recognized by AvaII at positions 1514, 1736 and 2683 in its sequence. Draw a map of pRS426 below, showing the positions of the AvaII restriction sites. (Be sure to show the position of nucleotide 1 in the sequence.)

Next to the drawing, list the sizes of the restriction fragments that would be generated by digesting pRS426 with AvaII.

Chapter 12

Yeast Transformation



Techniques for transforming microbial organisms with foreign DNA are essential in modern molecular biology. In this lab, you will transform a *S. cerevisiae met* strain with three different plasmids and use *ura3* complementation to detect transformed cells. You will then use replica plating to determine if *S. pombe Met* genes are functionally equivalent to *S. cerevisiae MET* genes.

Objectives

At the end of this lab, students should be able to:

- explain the processes of transformation and complementation at the molecular level.
- design a selection strategy to isolate transformed strains
- transform *S. cerevisiae* with plasmids and isolate transformants on selective media
- use replica plating to analyze the ability of plasmid-encoded *MET* genes to complement *met* deficiencies

Chapter 12

In this lab, you may receive a preliminary answer to the semester's research question about the functional conservation of Met proteins in the Ascomycota . During the first part of the semester, your team used selective plating and colony PCR to identify yeast deletion mutants. You then isolated and characterized plasmids that can be used to overexpress Met proteins. These two sets of experiments come together in this lab, when you transform the *S. cerevisiae* deletion strain with the expression plasmids. Through a series of complementation experiments, you will determine if the genes carried on the plasmids are able to compensate for the missing *MET* genes in the mutants. In complementation, the introduced gene restores the normal phenotype to a mutant with a defective gene.

Transformation alters the phenotype of a cell

Transformation refers to the uptake of DNA by a cell, causing a change in its phenotype. Naturally-occurring transformation was first described in 1928 by Frederick Griffith, who described a heat-stable transforming principle from virulent *Streptococcus pneumoniae* that could transform non-virulent *S. pneumoniae* to an encapsulated, virulent form. The transforming principle was subsequently identified as DNA by Avery and colleagues in 1944. Since then, transformation has become an indispensable tool in the molecular biology laboratory. The physical basis for yeast transformation is still incompletely understood, but researchers have empirically developed conditions that give fairly consistent transformation in the lab. Reliable transformation techniques have been developed for bacteria and many eukaryotes, ranging from yeast to mammalian cells.

Transformation conditions have been developed empirically

The challenge in laboratory transformation is to devise conditions under which DNA will pass across the cell wall and plasma membrane of living cells, which are normally impermeable to DNA. Very few cells are naturally competent, or able to take up DNA on their own. Consequently, researchers use a variety of chemical treatments to render cells competent. In general, these chemical treatments have some kind of destabilizing effect on the plasma membrane. The introduction of DNA into these competent cells can be further encouraged by a physical stress, such as a pulse of electric current or temperature elevation. Transformation is not a very efficient process, but because large numbers of microorganisms can be cultured in the laboratory, useful numbers of transformants can be obtained with most microorganisms.

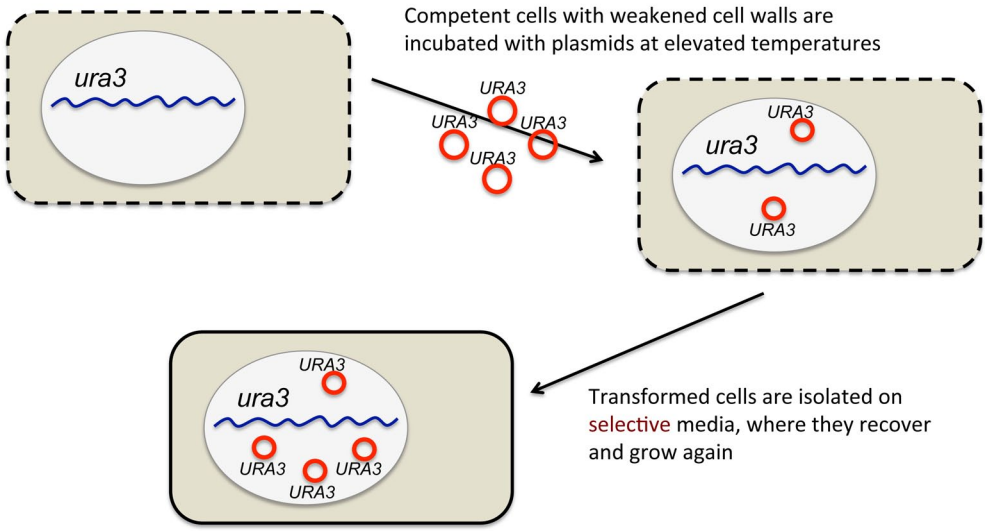
Techniques for yeast transformation are now standard in the laboratory. Depending on the details of the experimental procedure, reactions can yield as many as 10^6 transformants per µg DNA. The structure of the DNA used for transformation greatly affects the transformation efficiency. Transformation efficiencies are considerably higher with supercoiled plasmid DNA than with linear pieces of DNA, possibly because plasmids enter the cell more readily and/or plasmids are less susceptible to endonuclease digestion.

The most commonly used yeast transformation methods use a combination of lithium acetate, single-stranded carrier DNA and polyethylene glycol (PEG). Although no one knows

exactly how these components promote transformation, a number of hypotheses have been advanced. Lithium ions neutralize the negative charges on DNA molecules to be transformed and the phospholipid bilayer of the yeast cell, and they may also generate small holes in the plasma membrane that allow the passage of nucleic acids. Single-stranded DNA acts as a carrier for the plasmid DNA to be transferred into the cell and it may help to protect the latter from endonucleases. The source of the carrier DNA is unimportant. Since the carrier DNA concentration is considerably higher than that of the DNA to be introduced into the cell, the carrier DNA is usually isolated from an inexpensive source, such as salmon sperm. ***It is imperative that the carrier DNA for transformations be single-stranded.*** In our experiments, we will boil the carrier DNA for 5 minutes and then rapidly chill it to prevent reanneling of the DNA helix. PEG may help bring the DNA into closer apposition with the membrane. PEG is often used to promote membrane fusion and is thought to alter water structure around plasma membranes.

Complementation is often used to isolate transformants

The DNA used for transformation must carry a selectable marker whose presence can be detected by screening. Following a transformation, cells are plated on selective media that will allow transformed, but not untransformed, cells to grow. All the pBG1805- (Gelperin *et al.*, 2005) and pYES2.1-derived plasmids that we are using carry a normal copy of the yeast *URA3* gene, as well as the *URA3* promoter, so the gene is regulated much like a normal chromosomal gene. Our yeast deletion strains were derived from strain BY4742, which has the *ura3Δ0* allele (Winzeler *et al.*, 1999) Complementation will occur because the plasmid carries a functional copy of the gene that is defective in the mutant host strain. The Ura3p protein produced from the plasmid-encoded *URA3* gene compensates for the *ura3* deletion in the yeast chromosome, allowing transformed cells to grow in the absence of uracil, as shown below. Because of its reliability, many yeast transformation schemes rely on *URA3* complementation to isolate transformants.



Transformation and plasmid complementation
Competent *ura3* yeast cells are transformed by incubating cells with a plasmid containing the yeast *URA3* gene at an elevated temperature. Transformed cells are selected on media that does not contail uracil.

Chapter 12

Experimental considerations

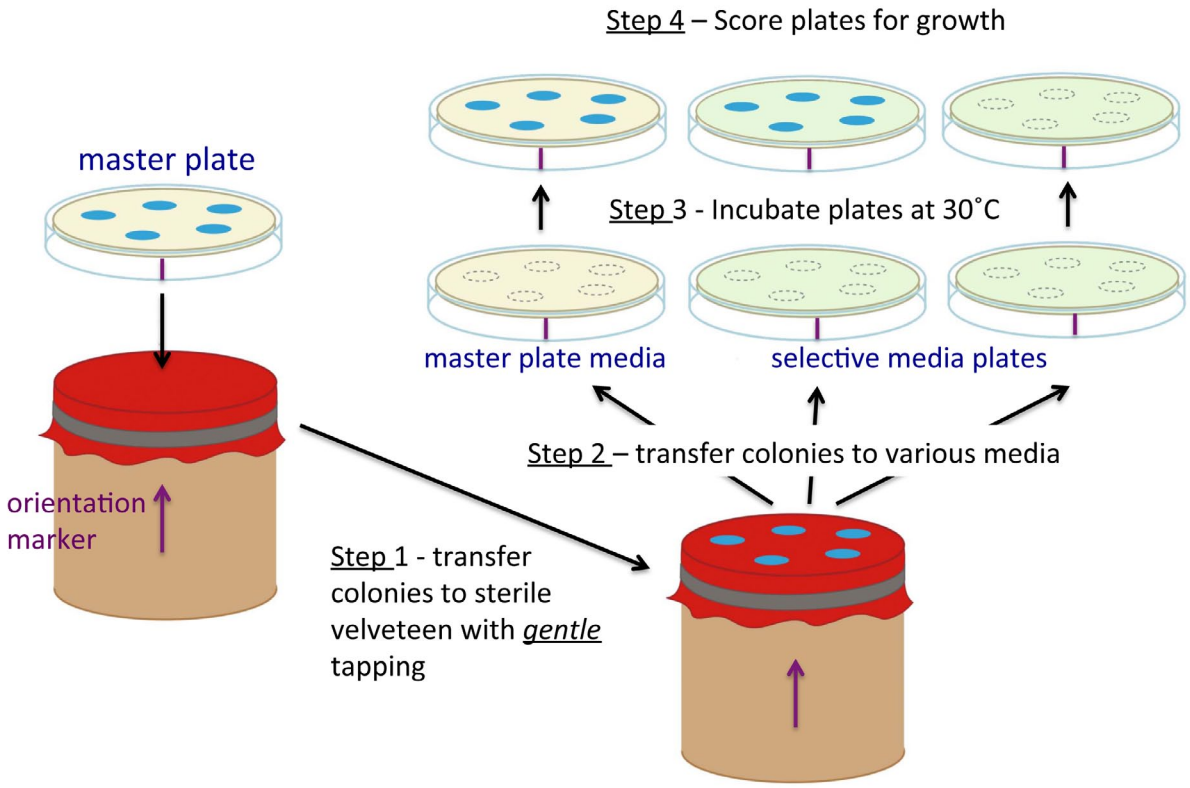
You may be wondering why we are not using *MET* gene complementation to isolate transformants, since this is the goal of our semester's project. There are several reasons why we are using YC-Ura plates, rather than YC-Met plates, to isolate tranformants. First, we need to insure that the overexpression plasmids have successfully transformed the deletion strains and it is possible that the Met fusion proteins encoded by the plasmids will be unable to complement the *met* deficiencies in transformants. *URA3* gene complementation offers a well-tested and reliable means to assess successful transformation that is independent of methionine metabolism.

A second issue relates to uncertainties associated with regulation of the plasmid ORFs by the *GAL1* promoter (Johnston, 1987). The *GAL1* promoter is an inducible promoter that is normally repressed when cells are grown in glucose and induced when galactose replaces glucose as the carbon source. In its normal chromosomal location, the *GAL1* promoter responds to a variety of positive and negative transcription regulators (Chapter 13). Although a large number of studies have established that the *GAL1* promoter functions well in ectopic locations, such as plasmids, the promoter is not as tightly regulated in plasmids as in the yeast chromosome. Some of this difference may relate to copy number. Both pBG1805 and pYES2.1 are multi-copy plasmids.

Following the isolation of transformants of YC-Ura plates, you will analyze *MET* gene complementation on YC-Met plates containing either D-glucose and D-galacatose. Keep in mind that galactose and glucose may not function as simple "ON" and "OFF" switches because the regulatory balance is altered in transformed cells. It is possible, for example, that "leaky" gene transcription could occur in the presence of the normal repressor, D-glucose. In this case, *MET* genes would complement *met* mutants grown in D-glucose. It is also possible that transformed cells could produce excessive quantities of Met and proteins that are detrimental, or even fatal, to transformed cells.

Replica plates accelerate the screening process

As noted above, transformation is an inefficient process, so researchers want to make the most of every cell that has been transformed. In our experiments, we will be isolating transformed cells for their ability to grow in the absence of uracil, but we are really interested in their ability to grow in the absence of Met. Replica plating offers a means to quickly screen a plate of cells for their ability to grow in a wide range of media, while retaining information about individual colonies. As shown on the opposite page, the original plate of transformants becomes the "master plate." An imprint of the master plate is made by *GENTLY* tapping the inverted plate on a piece of sterile cotton velveteen immobilized on a block. This imprint can then be transferred to plates with different kinds of selective media, establishing the genotype of the transformants. In our experiments, we will make transfer replicas of the transformation reactions (isolated on YC-Ura plates) to YC-Ura plates that are also lacking Met, with either glucose or galactose as a carbon source.



Replica plating provides a rapid screening method for analyzing phenotypes.
Colonies on a master plate are transferred to a sterile piece of velveteen. Copies of the mater plate are transferred to additional selective or indicator media to monitor phenotypes under additional conditions.
Adapted from http://en.wikipedia.org/wiki/Replica_plating#mediaviewer/File:Replica-dia-w.svg

Exercise 1 - Yeast transformation

The following protocol is a slight modification of the "Quick and Dirty" transformation protocol described by Amberg *et al.* (2005). With careful attention to detail and cooperative strains, this procedure can yield thousands of transformants per µg plasmid DNA. Modifications to this method can increase its efficiency by several orders of magnitude (Gietz and Schiestl, 2007), which would be required if linear pieces of DNA were being used to tranform yeast.

Prepare a transformation master mix

1. Prepare a transformation master mix. The following ingredients provide enough reagents for five transformation reactions. Combine in a microcentrifuge tube:

   100 µL sterile 2 M lithium acetate (freshly prepared)
   400 µL sterile 50% PEG-3350
   4 µL 2-mercaptoethanol (STINKY!! add this in the fume hood!)

Set up individual transformation reactions - <u>for each transformation</u>:

2. Add 15 µL of the denatured salmon sperm DNA (2 mg/mL) to a new microcentrifuge tube ***labeled*** with the name (or code) of the plasmid.

Note: It is important for the salmon sperm DNA to be single-stranded for this procedure to work well. Boil the DNA for 5 minutes to denature the DNA. Quick chill the DNA by placing it immediately on ice. Keep the DNA on ice until you are ready to use it.

3. Add 5 µL of miniprep plasmid DNA to the appropriately labeled microcentrifuge tube.
4. Add 100 µL of transformation mix from step 1 to each microcentrifuge tube. Vortex for 10-15 seconds to mix the contents.
5. Using a sterile toothpick, scrape a large yeast colony (or the equivalent of a "match head" of yeast) from a YPD plate. Transfer the yeast to the microcentrifuge tube containing the transformation/DNA solution (step 4) by twirling the toothpick several times. Be sure that the cells are uniformly suspended before proceeding.

Repeat steps 2-5 for each of your transformation reactions. ***Be sure to include a control that contains no plasmid DNA.***

6. Incubate the transformation mixtures at <u>37˚C</u> with shaking for <u>30-45</u> minutes.

Plate the transformed cells on selective media lacking uracil.

7. Remove 10 µL of the resuspended cells to 90 µL of sterile water in a microcentrifuge tube. This sample will be serially diluted for a spot plate (step 9) that you will use to calculate the transformation efficiency.
8. Plate the remainder of the mixture on a selective media lacking uracil. Transfer the transformation reaction to the plate, and then shake out ~4 sterile glass beads that will spread the cells. Cover the plates and spend 0.5-1 minutes agitating the plates so that the beads spread the transformation mixture evenly over the surface of the plate. Discard the glass beads into the appropriate waste containers, so they can be used again. Incubate the plates at 30˚C until colonies can be detected. The earliest that colonies will be visible is usually 2 days. If the colonies are small, allow them to grow an additional day(s) at 30˚C. Count the number of cells on the plate.

Determine the number of viable cells in the transformation mixture.

9. Prepare a series of 4 additional dilutions of the cells set aside in step 7. Use these dilutions for a spot plate on YPD media. Each row on the plate should contain cells from a different transformation reaction. Incubate the cells at 30˚C or room temperature until individual colonies can be detected. ***Do not allow the plate to overgrow, because you need to distinguish individual colonies.***

Calculate the transformation efficiency. The efficiency of transformation is influenced by both the quality of the DNA used and the precise details of the transformation procedure.

10. ***Calculate the fraction of cells that were transformed as shown below.*** Total volume of transformation mixture was ~110 µL, including yeast cells. Ten µL was used for spot plating and the remaining 100 µL was used for the transformation.

$$\text{Fraction cells transformed} = \frac{\text{Transformed cells (cell on YC-Ura plate)}}{\text{Number of cells in the transformation mixture}}$$

Calculate the number of cells in the transformation mixture (N) from the spot plate:

$$N = \text{\# cells in most dilute spot (5 µL)} \times \text{Dilution factor} \times (20)$$

correct for spot plate dilutions

100 µL of transformation mixture plated on YC-Ura

11. Transformation efficiencies are usually expressed by the number of cells transformed per µg DNA. In the last lab (Chapter 11), you analyzed your plasmid preparations on agarose gels and obtained a rough estimate of the DNA concentrations of your plasmid preparations. Note that you analyzed 7 µL of plasmid prep on those gels. In this transformation lab, you used 5 µL of your plasmid preps.

Calculate the transformation efficiency:

A. Multiply that number of transformed cells on the YC-plate by 1.1
(Only 100 out of 110 µL in the transformation reaction were plated.)

B. Convert the ng of plasmid in your transformation reaction to µg.

C. Divide the calculated value in A by that in B.

Exercise 2 - Replica plating and complementation

This exercise will be performed at the next lab session after transformants will have had a chance to grow.

Your initial selection of transformants was done on plates that lacked uracil, but contained methionine. You next will test the ability of your transformed strains to grow on media lacking methionine using replica plating. We will use methionine-free media containing either glucose or galactose for replicas, and you will also prepare a fresh master plate. Predict which transformants will grow on each of the plates.

It is important to have a light touch during replica plating!! The goal is to transfer a small portion of cells from each colony on the master plate (the plates carrying your transformants) to a number of plates containing different media.

1. Place an orientation mark with a magic marker on the perimeter of your master plate as well as the plates that will be used for replicas.

2. Remove the lid from your master plate and invert the plate on the block, aligning the orientation marker on the plate with the marker on the block. *GENTLY* and *EVENLY* tap on the bottom of the plate to transfer cells to the velveteen. Remove the master plate and replace the lid.

3. Repeat step 3 with plates containing the following media:
    - Medium without uracil or methionine, containing glucose
    - Medium without uracil or methionine, containing galactose
    - Medium without uracil, containing glucose and methionine

Test yourself

An investigator transforms strain BY4742 (*MATα his3-Δ1 leu2Δ0 lys2Δ0 ura3Δ0 met6::KAN*^{R}) with the plasmids listed in the table below. The transformants are selected on appropriate media and then replica plated to the media listed in the table below. Complete the table below by indicating when cells will grow (+ or Y) or not grow (- or N).

Plasmid	**YC**	**YC-His**	**YC-Leu**	**YC-Lys**	**YC-Ura**	**YC-Met**
pYES2.1-*MET6 LYS2*						
pYES2.1-*met6Δ2 HIS3*						
pBG1805-*MET3 LEU1*						
pBG1805-*HIS3 LEU2*						
pYES2.1-*HIS2 LEU2 LYS2*						
pBG1805-*LEU2 LYS3*						

References

Amberg DC, Burke DJ & Strathern JN (2005) *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Gelperin DM, White MA, Wilkinson ML *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

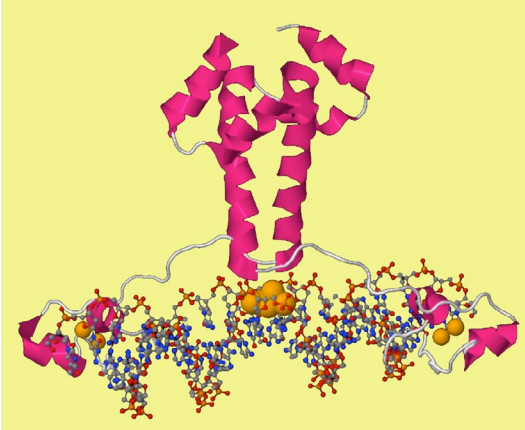
Gietz RD & Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* **2**: 31-34.

Johnston, M (1987) A model fungal regulatory mechanism: the *GAL1* genes of *Saccharomyces cerevisiae*. *Microbiol Rev* **51:** 458-476.

Winzeler, EA, Shoemaker, DD, Astromoff, A *et al.* (1999) Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. *Science* **285:** 901-906.

Chapter 13

Protein overexpression



In this lab, you will use various carbon sources to manipulate the expression of Met and LacZ fusion proteins in cells that have been transformed by overexpression plasmids. The Gal4p transcription factor (above) binds to the *GAL1* promoter in the plasmids and controls protein expression. In this lab, you will prepare protein extracts from cells growing under both repressed and induced conditions for later analysis.

Objectives

At the end of this lab, students will be able to

- compare the effects of different carbon sources on transcription of genes controlled by the yeast *GAL1* promoter.
- explain the roles of heat and detergents in the preparation of yeast cell extracts.
- culture yeast with different carbon sources to induce or repress expression from the *GAL1* promoter.
- prepare extracts from transformed yeast strains suitable for use in SDS-PAGE gels and western blots.

Chapter 13

Over the next few weeks, you will be analyzing the expression of *S. pombe* and *S. cerevisiae* Met or bacterial LacZ fusion proteins in your transformed strains. You have already tested the ability of the overexpression plasmids to complement the *met* mutation in your yeast strains. Complementation depends on the presence of functional Met proteins. If you observed a failure to complement the *met* deficiency, this could indicate that proteins were not expressed from the plasmids. **Alternatively, the overexpressed proteins may be present, but not function normally.** Remember that the proteins expressed from the BG1805 and pYES2.1 plasmids are fusion proteins with additional sequences at their C-termini (Gelperin *et al.*, 2005). The biochemical activities of these fusion proteins have not been previously evaluated. (We are the first to test whether the Met fusion proteins function similarly to the normal *S. cerevisiae* proteins!)

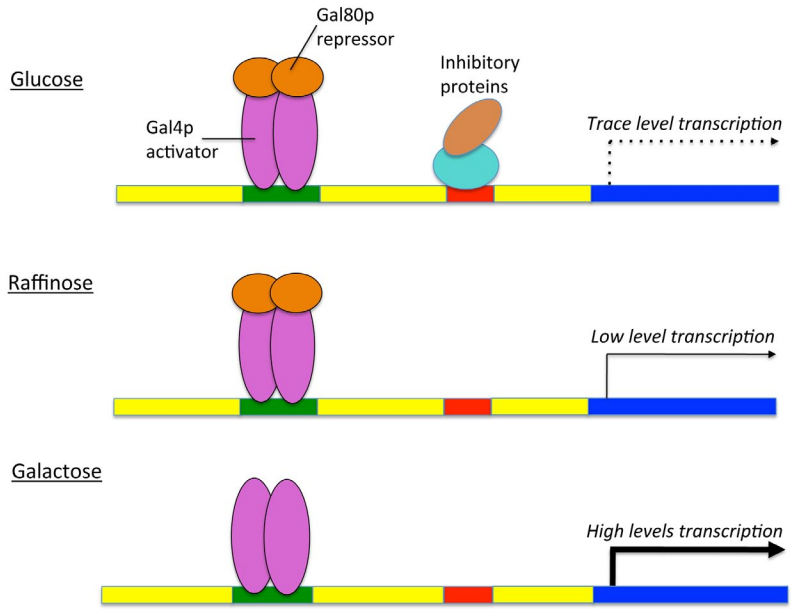
In this experiment, you will prepare extracts from yeast for later experiments (Chapters 14 and 15) in which you will determine if the Met and LacZ fusion proteins are being successfully expressed in the transformed yeast strains and how the expression of the fusion proteins varies with carbon sources.

Regulation of the GAL1 promoter

In yeast, glycolysis plays a major role in energy production, and glucose is far and away its preferred carbon source. Genes involved in the metabolism of other carbon sources are usually repressed when glucose is available. When glucose is not available, however, yeast activate genes that metabolize other available energy sources, such as galactose. Galactose increases the transcription of several genes for enzymes that transport galactose into cells and ultimately convert it into glucose-6-phosphate (G6P), an intermediate in glycolysis. The first gene in the pathway induced by galactose, *GAL1*, encodes galactokinase, which phosphorylates galactose to galactose-1-phosphate. (Check out the *GAL1* pathways link in SGD.) The *GAL1* promoter has been incorporated upstream of the ORF site in both the pBG1805 and pYES2.1 plasmids and therefore controls transcription of plasmid-encoded *MET* and *lacZ* genes in transformed cells.

The figure on the opposite page provides a simple overview of gene expression from the *GAL1* promoter in the presence of glucose, raffinose and galactose. The promoter contains both negative and positive regulatory sites encoded within its DNA sequence. In the presence of glucose, repressor proteins bind to the negative regulatory sites and repress transcription. The Gal4p transcriptional activator binds to positive regulatory sites. Gal4p is a zinc-finger transcription factor that binds to DNA as a dimer. (The figure at the beginning of this chapter shows the crystal structure of the DNA binding and dimerization domains of Gal4p complexed with DNA.) In the presence of glucose, Gal4p is inactive, because it is bound to the repressor protein, Gal80p.

Glucose repression can be relieved by growing cells in a poor carbon source, such as raffinose. Raffinose is a trisaccharide composed of galactose, fructose and glucose. Raffinose is not able to induce high levels of *GAL1* expression, which requires galactose. The *GAL1* promoter is exquisitely sensitive to galactose. In the presence of galactose, expression of the *GAL1* gene



Regulation of the *GAL1* promoter. In the presence of glucose, transcription is repressed because repressor proteins bind to regulatory sites in the DNA and to the Gal4p transcriptional activator. Glucose repression is relieved in the presence of raffinose, but Gal4p remains inactive. Gal4p activates transcription in the presence of galactose due to the removal of the Gal80p protein.

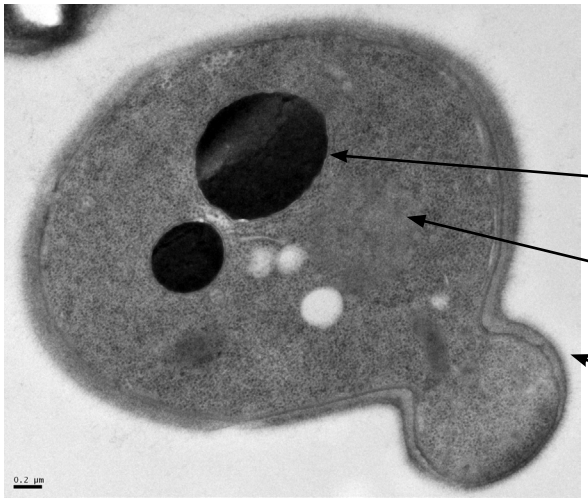
increases ~1000-fold above the level observed in the presence of glucose. This stimulation is primarily due to the activity of Gal4p, which is no longer bound to the inhibitory Gal80p protein. Gal4p acts as a master regulator of galactose metabolism. In addition to activating *GAL1* transcription, Gal4p also binds to the promoters of the *GAL7* and *GAL10* genes, which are situated adjacent to the *GAL1* gene on yeast chromosome 2. Like *GAL1*, the *GAL7* and *GAL10* genes encode proteins involved in galactose metabolism.

Preparing protein extracts from yeast cells

Proteins comprise about half of the dry weight of most cells and include the many structural proteins, catalysts, receptors and signaling proteins responsible for cell function. To understand cell function, scientists often want to analyze the protein composition of cells. Protein analysis begins with the preparation of a cell extract, ideally under conditions that minimize protein degradation. Preparing good cell extracts is something of an art, and many factors need to be considered during the design of an extraction protocol. In this course, we will be analyzing protein function in yeast. An average haploid yeast cell contains ~6 pg protein (Sherman, 2002). Although yeast cells have many advantages for genetic studies, they are notoriously difficult to use for biochemical studies. Nonetheless, scientists have been able to develop procedures for extracting yeast proteins that circumvent these experimental barriers.

Chapter 13

The first consideration in designing an extraction procedure is the compartmentalization of cells. All cells are surrounded by a plasma membrane and eukaryotic cells contain additional membranes that surround organelles. Fungal cells also have cellulose-based cell walls that protect the cells against mechanical and osmotic forces. Cell extraction procedures begin with the disruption of the plasma membrane and cell wall by mechanical and/or chemical treatments. Mechanical disruption of yeast cells must be fairly vigorous because their cell walls are very tough. Mechanical methods commonly used to disrupt yeast include sonication, high pressure, and "beating" with glass beads. These vigorous treatments run the risk of damaging proteins because of the heat and foaming generated during the processes.



Transmission electron micrograph of an *S. cerevisiae* cell. A 60 nm thin section was prepared by the high pressure freezing and freeze-substitution method described in Buser & Drubin (2013).

Vacuole, which stains heavily with this technique, contains many proteases.

Nucleus stains poorly in this technique.

Polysaccharide-based wall surrounds the cell.

Reproduced with permission of Christopher Bruser; www.celllibrary.org (CIL10452)

Chemical treatments offer a gentler alternative to mechanical disruption for preparing extracts. Chemical extraction procedures frequently include detergents that solubilize membrane lipids, thereby allowing proteins to diffuse out of the cell. Most detergents do not discriminate between intracellular and plasma membranes, so a detergent extract usually contains proteins from multiple organelles as well as cytoplasmic proteins. In this experiment, we will use sodium dodecyl sulfate (SDS) as the detergent. SDS is a denaturing detergent that unfolds protein structures by breaking the thousands of weak bonds that normally stabilize protein structures. The proteins are converted to random coils coated along their lengths by negatively charged SDS molecules.

When preparing extracts, care must be taken to protect proteins from degradation by cellular proteases. Cells contain proteases with many different specificities that are responsible for normal turnover of proteins in cells. Cell disruption often releases proteases from compartments such as lysosomes, providing them access to cytoplasmic proteins. Yeast are notoriously rich in proteases. In an intact yeast cell, many of these proteases are located in the yeast vacuole, which is analogous to the mammalian lysosome. The protocol that we will use in this course (Amberg *et al.*, 2005) uses a combination of heat and SDS to rapidly destroy the yeast proteases. Cell suspensions are immediately plunged into a boiling water bath after the addition of SDS. Extracts prepared by this method are suitable for electrophoresis and western blot analysis.

Exercise 1 – Prepare cell cultures for extraction

DAY 1 - First lab session of the week

1. Each student should collect two culture tubes containing 1 mL of YC-Ura medium with 2% raffinose. Each of two tubes should have a **DIFFERENT** color cap. The different colored caps are used to indicate whether the culture will be given glucose or galacatose on Day 2 of the experiment. Label each tube with the name of your strain and the plasmid transformed into it.
2. Inoculate each tube with a single yeast colony. Together, each team will prepare 6 cultures, including 2 cultures of each transformed strain. The two samples of each culture should be given caps of different colors.
3. Place the cultures in the 30˚C shaking incubator. Turn the incubator on.

DAY 2 - The day of your lab class: 4-5 hours before class, the teaching staff will:

- add 1 mL of YC+glucose (repression medium) to the raffinose cultures with one color cap.
- add 1 mL YC + galactose (induction medium) to raffinose cultures with the other color cap.

You will be told which color corresponds to which sugar source

Exercise 2 – Preparing cell extracts

Prepare the cells for extraction

1. Collect your team's 6 cultures and note which cultures contain glucose and which cultures contain galactose.
2. Determine the cell concentration or each culture. Vortex the tubes and transfer 100 µL of each cell culture to 900 µL deionized water. Measure the OD_{600} of cultures in the spectrophotometer. Note the values in your notebook. You will refer to them later when interpreting your SDS-PAGE gels (Chapter 14).
3. Transfer 1.5 mL of your cultures into **labelled** microcentrifuge tubes. Collect the cells by centrifugation for 1 minute using the microcentrifuge at top speed. Remove and discard the supernatant.
4. Rinse the cells. Add 1 mL deionized water to each tube. Resuspend the cell pellets by gently drawing the cells in and out of a micropipette tip, taking care to prevent premature lysis of the cells. This rinse step removes proteins from the culture medium that may contaminate the cell pellet. Centrifuge again for 1 minute at top speed.
5. Discard the supernatant and resuspend the cells in 100 µL deionized water.

Prepare the protein extract

6. Add 100 µL of 0.2N NaOH to each tube, and incubate the cells for 5 minutes at room temperature. (The addition of NaOH does not lyse the cells, but it makes them more permeable and more fragile.)
7. Pellet the cells again in the microcentrifuge and remove the supernatant.
8. Resuspend the cells in 50 µl 2 X SDS-PAGE sample buffer.* Lock the lids in place with the built-in locking mechanism. IMMEDIATELY place the tubes in a boiling water bath. Leave the cells in the water bath for 3 minutes. This treatment effectively denatures the proteins. Yeast cells contain many proteases that could degrade other cellular proteins, so it's important that the proteases are denatured before they have the chance to attack other cellular proteins.

   NOTE: The 2 X SDS-PAGE sample buffer contains 2-mercaptoethanol (also known as β-mercaptoethanol, or BME). Use appropriate caution and work quickly when handling this reagent. BME is a volatile reagent with a strong odor reminiscent of rotten fish.
9. After boiling, store the samples in the freezer for future use.

*2 X SDS-PAGE sample buffer consists of:
120 mM Tris/HCl, pH 6.8
10% glycerol (glycerol is significantly more dense than water)
4% SDS
8% 2-mercaptoethanol
0.004% bromophenol blue (a tracking dye for electrophoresis)

References

Amberg DC, Burke DJ & Strathern JN (2005). *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Buser C & Drubin DG (2013) Ultrastructural imaging of endocytic sites in *Saccharomyces cerevisiae* by transmission electron microscopy and immunolabeling. *Microsc. Micoroanal.* **19**: 381-392.

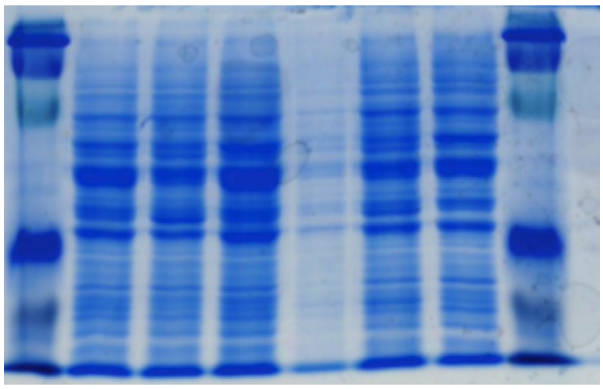
Gelperin DM, White MA, Wilkinson ML *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

Johnston M (1987) A model fungal regulatory mechanism: the *GAL1* genes of *Saccharomyces cerevisiae*. *Microbiol Rev* **51:** 458-476.

Sherman F (2002) Getting started with yeast. *Method Enzymol* **350:** 3-41.

Chapter 14

SDS-PAGE



This lab will introduce you to SDS-PAGE, a simple and inexpensive method for resolving proteins in complex mixtures. SDS-PAGE gels provide the starting materials for western blots and for some proteomic techniques. In this lab, you will use SDS-PAGE to analyze the protein extracts that you prepared from yeast strains overexpressing Met and LacZ fusion proteins.

Objectives

At the end of this lab, students will be able to:

- discuss the principles that govern protein separation on discontinuous SDS-PAGE gels.
- cast and run SDS-PAGE gels.
- analyze the pattern of bands on a stained SDS-PAGE gel
- estimate the molecular weight of a protein from its migration on SDS-PAGE gels

Chapter 14

This lab will introduce you to SDS-PAGE (sodium dodecyl sulfate - polyacrylamide gel electrophoresis), a simple and inexpensive method for resolving proteins in complex mixtures. You will use SDS-PAGE gels to analyze the yeast protein extracts that you prepared in the last lab. Each team will make two gels. One gel will be stained with Coomassie Blue to visualize all the proteins in the extracts. The second gel will be used for western blots (Chapter 14) that will specifically detect Met and LacZ fusion proteins in the extracts.

Background

SDS-PAGE is widely used to analyze the proteins in complex extracts. The most commonly used methods are derived from the discontinuous SDS-PAGE system first described by Laemmli (1970). The system actually consists of two gels - a resolving (aka running) gel in which proteins are resolved on the basis of their molecular weights (MWs) and a stacking gel in which proteins are concentrated prior to entering the resolving gel. Differences in the compositions of the stacking gel, resolving gel and electrophoresis buffer produce a system that is capable of finely resolving proteins according to their MWs.

Gel electrophoresis of macromolecules

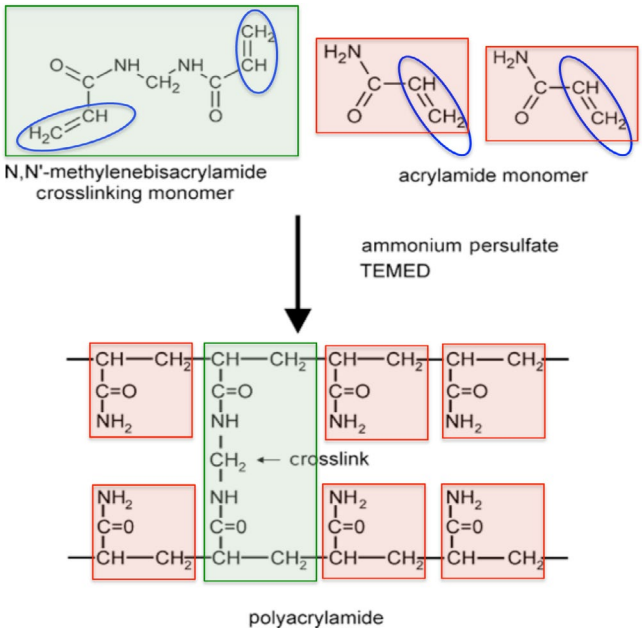
In gel electrophoresis, an electric field is used to move charged molecules through a matrix of a polymerized substance such as agarose or polyacrylamide. The rates at which individual molecules move through the gel depend on the properties of both the separation system and the molecules themselves. Gel matrices are permeated with networks of pores through which the molecules move. The amount of resistance that the matrix presents to the movement of a molecule depends on the diameter of the pore as well as the size and geometry of the molecule. Researchers can control the size of the pore by adjusting the concentration of gel monomer within a certain range. In general, smaller, more highly charged molecules migrate more rapidly through gels than larger or less charged molecules. The mobility of a molecule is also affected by the buffer system and the strength of the electrophoretic field used for the separation.

You have already used agarose gel electrophoresis to separate DNA molecules. Recall that the size of a linear DNA molecule can be estimated from the rate at which it moves through an agarose gel, because DNA molecules have a uniform charge to mass ratio. Protein electrophoresis is somewhat more complicated than DNA electrophoresis. Proteins are much smaller than DNA molecules, so polyacrylamide gels are used for their separation. In addition, proteins are much more structurally diverse than DNA, so chemical treatments (see below) are used to impart a uniform geometry and charge/mass ratio to the proteins.

Chemistry of acrylamide polymerization

The polyacrylamide gels used to separate proteins are formed by the chemical polymerization of acrylamide and a cross-linking reagent, N,N'methylenebisacrylamide (opposite page). Investigators are able to control the size of the pores in the gel by adjusting the concentration of acrylamide, as well as the ratio of acrylamide to bisacrylamide. Raising either the concentration of acrylamide or bisacrylamide, while holding the other concentration

constant, will decrease the pore size of the gel. Polymerization occurs because of free oxygen radicals that react with the vinyl groups in acrylamide and bisacrylamide, as shown in the figure below. The oxygen radicals are generated from the catalyst, ammonium persulfate (APS), when it reacts with a second catalyst, N,N,N',N'-tetramethylethylenediamine (TEMED).



Acrylamide gel polymerization. Ammonium persulfate and TEMED catalyze the polymerization of acrylamide and bis-acrylamide monomers into a crosslinked network.

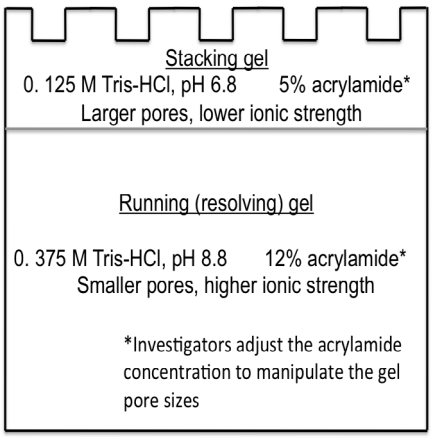
Proteins are denatured prior to electrophoresis

Compared to DNA molecules, proteins are structurally very diverse. Proteins show tremendous variation in their amino acid compositions and in the distribution of amino acids in their folded structures, features with important implications for electrophoresis. Recall that proteins are mixtures of hydrophobic and hydrophilic amino acids and that the primary sequence of the protein determines its final folded form. Because of the hydrophobic effect, the surfaces of proteins proteins have a higher frequency of polar and charged amino acids than the interiors, where hydrophobic residues predominate. Folded proteins assume many different geometries and their surfaces are mosaics with respect to the distribution of R groups with different chemistries. Because proteins are so diverse with respect to their surface charges and geometries, the molecular weights of ***folded*** proteins cannot be simply determined by their migration rate in an electric field. Postively and negatively charged proteins would migrate in different directions!

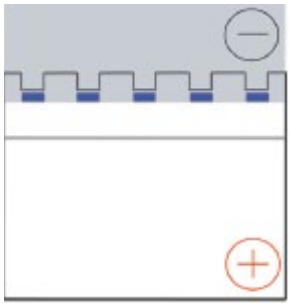
To resolve the proteins in a sample according to their size, investigators must convert the proteins to a uniform geometry and impart a uniform charge/mass ratio to the proteins. In SDS-PAGE, the solution is to denature the proteins by boiling them with the anionic detergent, sodium dodecyl sulfate (SDS) and 2-mercaptoethanol. The combination of heat and detergent is sufficient to break the many noncovalent bonds that stabilize protein folds, and 2-mercaptoethanol breaks any covalent bonds between cysteine residues. Like other detergents, SDS is an amphipathic molecule, consisting of a hydrophobic 12-carbon chain and a hydrophilic sulfate group. The SDS

hydrocarbon chain permeates the protein interior and binds to hydrophobic groups, reducing the protein to a random coil, coated with negatively charged detergent molecules all along its length. Denatured proteins bind quite a lot of SDS, amounting to ~1.4 g SDS/g protein, or ~one SDS molecule for every two amino acids.

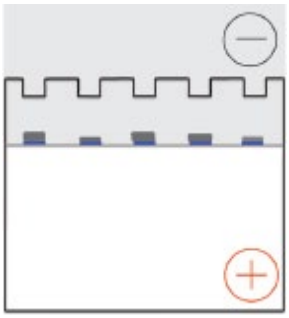
Discontinuities between the stacking and running gels underlie the resolving power of the SDS-PAGE gels



The Laemmli (1970) SDS-PAGE system can be considered a 3-component system. The stacking and running (resolving) gels have different pore sizes, ionic strengths and pHs. The third component is the electrophoresis buffer (25 mM Tris, 192 mM glycine,, 0.1% SDS, pH ~8.3), which contains large amounts of glycine. The ionization state of the glycine is critical to the separation. At neutral pH, glycine is a zwitterion, with a negatively charged carboxyl group and a positively charged amino group. The pK_a of the amino group is 9.6, considerably higher than the pH of the chamber buffer. Consequently, very little glycine has a negative charge in the chamber buffer or stacking gel, and significant ionization does not occur until the glycine enters the more alkaline pH 8.8 environment of the running gel. Let's follow the progress of protein samples during SDS-PAGE to see how differences in the composition of these three components generate the high resolving power of SDS-PAGE gels.

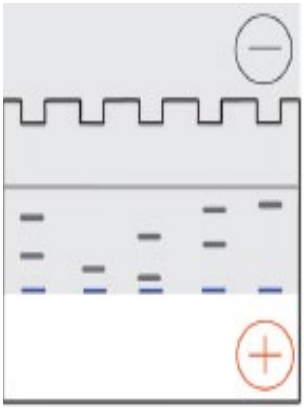


The sample buffer used for SDS-PAGE contains a tracking dye, bromophenol blue (BPB), which will migrate with the leading edge of the proteins being separated on the gel. The sample buffer also contains glycerol, which allows the protein samples to settle into the bottom of the gel wells. The gel is vertically positioned in the electrophoresis apparatus and covered with chamber buffer containing glycine (right, shaded).



Once a voltage is applied, the chloride ions in the sample buffer and stacking gel move rapidly toward the positive pole, forming the leading edge of a moving ion front. Glycine molecules have very little charge in the stacking gel, so they migrate at the rear of the moving ion front. This difference in chloride and glycine mobility sets up a steep voltage gradient in the stacking gel that sweeps along the negatively charged protein-SDS complexes. The large pores of the stacking gel present very little resistance to the movement of protein-SDS complexes, which then "stack up" into a very concentrated region at the interface between the running and stacking gels (right). Protein-SDS complexes remain concentrated at the interface until the slowly migrating glycine molecules reach the boundary between the two gels.

Dramatic changes occur as the glycine ions enter the running gel. The pH of the running gel is closer to the pK_a of the glycine amino groups, so a significant fraction of the glycine molecules assume a negative charge. Negatively charged glycine molecules begin to move at the same rate as the chloride ions, thereby eliminating the voltage difference that controlled protein mobility through the stacking gel. The pores in the running gel are much smaller than those of the stacking gel, so the pores present frictional resistance to the migration of proteins. Proteins begin to migrate at different rates, because of the sieving properties of the gel. Smaller protein-SDS complexes migrate more quickly than larger protein-SDS complexes (right). Within a certain range determined by the porosity of the gel, the migration rate of a protein in the running gel is inversely proportional to the logarithm of its MW.



Proteins are visualized with stains.

With few exceptions, naturally-occurring proteins are invisible on SDS-PAGE gels. Consequently, researchers often use pre-stained protein standards to monitor the approximate positions of proteins during electrophoresis. The pre-stained standards are produced by covalently attaching a large number of chromophores to a protein. The addition of the chromophores increases the MW of the protein and also produces more diffuse bands on the gel. The diffuseness of the bands reflects variation in the number of dye molecules attached to individual protein molecules. We will use prestained standard proteins in our gels, so you will be able to visualize the protein separation that is occurring. Yeast proteins will not be visible, however, because they have not been modified with chromophores.

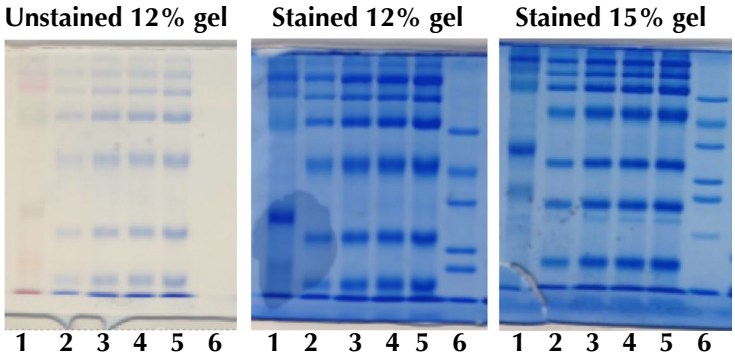
To visualize the positions of proteins after electrophoresis is complete, investigators stain the gels with various dyes that bind noncovalently and with very little specificity to proteins. During the staining process, proteins are also "fixed" in the gel, meaning that proteins become insoluble and unable to diffuse out of the gel. In our experiments, we will use a colloidal suspension of Brilliant Blue G-250, also known as Coomassie Blue G. Brilliant Blue G-250 binds proteins nonspecifically through a large number of ionic and Van der Waals interactions. In this procedure, gels are rinsed with water to remove the buffer salts used for electrophoresis and then treated with the colloidal G-250 suspension. Protein bands appear rapidly, and when necessary, the gels can be destained with deionized water to lower the gel background. Brilliant Blue staining intensity is considered to be a quantitative procedure, because with some exceptions, the intensity of a stained band is directly proportional to the amount of protein in a band.

Protein molecular weights can be calculated from their migration on gels

The sizes of proteins in an extract can be calculated by comparing their migration to a set of standard proteins run on the same gel. Researchers select standard proteins that will be well-resolved on the particular gel that they are running. For example, an investigator using a 7.5% gel will select standards with higher molecular weights (MWs) than an investigator using a 15% gel,

which is better suited to the analysis of small proteins. The principles used to estimate MWs are the same used for agarose gel electrophoresis. A plot of the $\log_{10}$MW of the standard proteins against the distance that each protein migrated on the gel will give a straight line in the region where the gel has good resolving power. (Note: MW is not the same as the mass of a protein. MW is a dimensionless term. For example, myoglobin has a mass of 16.7 kDa and a MW of 16,700.) The sizes of unknown proteins can be estimated by interpolating experimental values on a graph of standard proteins. Proteins whose molecular weights fall outside this range will not be well-resolved on the gel.

The figure below illustrates several of the points discussed above. The same sets of unstained and pre-stained protein standards were separated on either 12% or 15% SDS-PAGE gels. The prestained standards in lanes 1-5 are visible without staining, but they become much more pronounced after staining. The unstained standard in lane 6 requires staining to become visible, but the bands are much more discrete and will give more reliable values when calculating MWs of unknown proteins, because chromophores have not been attached to the proteins. The data in lanes 2-5 also demonstrate that Brilliant Blue staining is a quantitative procedure, because the intensity of bands in each lane increases in direct proportion to the amount of protein in the lane.



Molecular weight standards
1-Kaleidoscope protein standards (BioRad)
2-6.25 µg prestained protein standard
3-12.5 µg prestained protein standard
4-15 µg prestainded protein standard
5-18.75 µg prestainded protein standard
6-Unstained protein standard

When analyzing your experimental data, remember to consider the additional amino acids that have been added during the cloning procedure. The Met proteins that you are working with are fusion proteins with additional amino acids at the C-termini the Met proteins. The BG1805 plasmid encodes HA and His6 epitopes, as well as the ZZ immunoglobin binding domain. Together these sequences add a walloping ~19 kDa to the expected mass of *S. cerevisiae* Met proteins (Gelperin et al., 2005). The pYES2.1 plasmid encodes 33 amino acids that are added to cloned ORFs. The additional sequences include a V5 epitope tag and a $(His)_6$ purification tag at the C-termini of overexpressed proteins. Together, these amino acids add ~5000 kDa to the size of the protein.

NOTE: The MW of the LacZ control protein without the V5 epitope is ~120,000. Because this is such a large protein, it will be very difficult to get an accurate estimate of its MW.

Casting SDS-PAGE gels

These instructions are designed for constructing two 12% SDS-PAGE gels with the Bio-Rad Mini Protean system. One gel will be used for Brilliant Blue G staining in the next lab. The second gel will be used for western blotting.

Assemble the gel casting apparatus

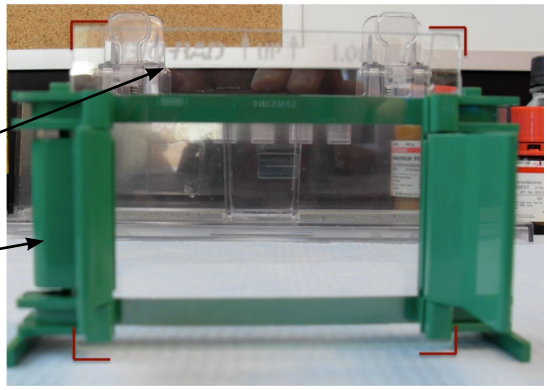
1. Assemble the components that you will need for casting the gel: a tall glass plate with attached 1 mm spacers, a small glass plate, a green casting frame and a casting stand.
2. Place the green casting frame on the bench with the green "feet"resting firmly against the bench and the clamps open (perpendicular to the frame) and facing you.
3. Place the two gel plates in the frame. Insert the taller spacer plate with the "UP" arrows up and the spacers facing toward you into the casting frame (the BioRad logo should be facing you). Insert the short glass plate in the front of the casting frame. ***There should be a space between the plates.***
4. Secure the plates in the casting frame by pushing the two gates of the frame out to the sides. IMPORTANT: the bottom edges of the two plates should be flush with the lab bench before you clamp the frame closed to ensure a watertight seal. T***o do this, rest the frame vertically on the bench BEFORE closing the gates.***
5. Clamp the casting frame with glass plates into the casting stand, with the gates of the casting frame facing you. ***Repeat steps 1-5 to prepare a second gel in the casting frame.***
6. Check to see if the assembled plates in the casting stand are sealed properly by pipetting a small amount of deionized water into the gap between the plates. If the glass plates hold water and don't leak, you are ready to make the gels. Pour the water out by holding the entire casting platform over a liquid waste container or sink. Use paper towels or tissues to absorb any residual water. If the gel leaks, disassemble the frame, dry the plates and go back to step 3.

Assembling plates in the casting frame.

Large spacer plate with spacer goes in rear, oriented so the arrows point up and the logo is readable. Small plate is placed in front. (Corners of the plates are marked with bracket marks.)

Green gates of casting frame are open.

Casting frame "feet" and bottom edges of plates are flush against the benchtop.



Chapter 14

Prepare two resolving gels.

SAFETY NOTE: Acrylamide and bisacrylamide monomers are weak neurotoxins. Gloves and goggles should be used when working with acrylamide.

Assemble the chemicals that you will need to pour the gels. The table below shows the quantities of each chemical that you will need to pour two gels with the Mini-Protean system. Polymerization occurs rapidly, so be sure to follow the step-by-step instructions below.
NOTE: catalysts should NOT be included into the mixture until you are ready to pour the gels!!

Reagent	**Resolving gel**	**Stacking gel**
Deionized water	3.5 mL	2.1 mL
30% acrylamide:bis-acrylamide (29:1)	4.0 mL	0.63 mL
1.5 M Tris-HCl, 0.4% SDS, pH 8.8	2.5 mL	------
0.5 M Tris-HCl, 0.4% SDS, pH 6.8	------	1.0 mL
10% ammonium persulfate (catalyst)	100 µL	30 µL
TEMED (catalyst)	10 µL	7.5 µL

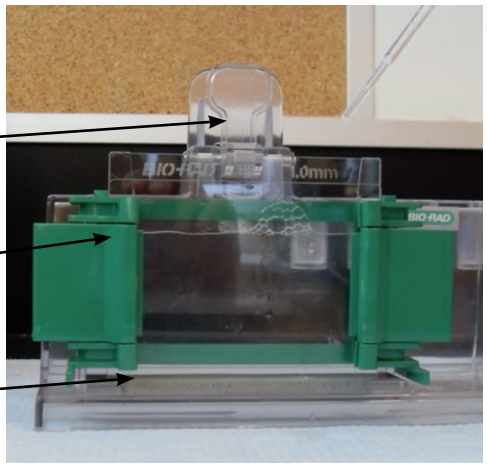
1. Label two 15 mL conical tubes "Resolving gel" and "Stacking gel".
2. Prepare ONLY the resolving gels at this time. Mix the acrylamide solution, **pH 8.8** Tris buffer and water, as shown in the chart above. Mix the ingredients gently, trying not to introduce air. Oxygen inhibits polymerization of acrylamide gels.
3. To the resolving gel mixture, add 100 µL of a 10% ammonium persulfate (APS) solution. Gently mix the solution, trying not to introduce air. Oxygen inhibits acrylamide polymerization.
4. Add 10 µL of TEMED catalyst. Once again, gently mix in the catalyst trying not to introduce air bubbles. CAUTION: TEMED has an unpleasant odor. Cover the tube immediately after you aliquot this reagent.
5. Working quickly, use a plastic transfer pipette to fill the space between the two plates until the resolving gel solution reaches a height just above the green clamps on the gel casting frame. Draw up any remaining acrylamide into the transfer pipet. (You will know that the acrylamide has polymerized when you can no longer push the acrylamide out of the pipet.)
6. Using a transfer pipet, add deionized water so that it gently flows across the surface of the polyacrylamide mixture. The water layer ensures that the polyacrylamide gel will have a level surface once it polymerizes.
7. Allow the gel to polymerize, which takes ~15-20 minutes. You will note that the interface between the polyacrylamide and water overlay disappears temporarily while the gel polymerizes. A sharp new interface then forms between the two layers, indicating that polymerization is complete. (You can also check the remaining polyacrylamide in the transfer pipette to see if it has polymerized.)
8. When polymerization is complete, remove the water from the top of the resolving gel by tilting the gel to the side and using a paper towel or Kimwipe to wick out the water.

Casting the SDS-PAGE gel.

Casting stand and plates are clamped into the casting stand

Fill plates with running gel solution to the top of the closed green gates of the casting frame.

Bottom edges of plates are flush against the grey pad of the casting stand.



Pour the stacking gels

1. Prepare the stacking gels. Mix the acrylamide solution, **pH 6.8** Tris buffer and water, as shown in the chart above.
2. Add 30 µL 10% APS and 7.5 µL TEMED to the stacking gel acrylamide mixture. Mix the contents by gently inverting the tube twice.
3. Use a transfer pipette to pipette the stacking gel on top of the resolving gel between the two glass plates. Add enough stacking solution until it just reaches the top of the small plate.
4. Carefully, but quickly, lower the comb into position, being careful not to introduce air bubbles. (The Bio Rad logo on the comb should be facing you.) Adding the comb will force some solution out of the gel, but this is fine. If air bubbles become trapped below the comb, remove the comb and reposition it.

Save the SDS-PAGE gels

1. Carefully remove the gels from the casting stand and then from their green frames.
2. Keeping the combs in the gel, wrap the gels in a wet paper towel. Then wrap the gels in plastic wrap to be used in later labs.

Chapter 14

Running SDS-PAGE gels

Set up the electrophoresis apparatus

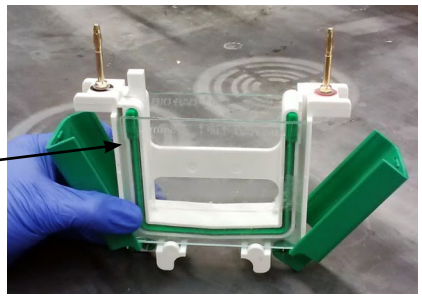
1. Retrieve one of the SDS-PAGE gels from the refrigerator.
2. Carefully remove the comb from the spacer gel.
3. Remove the casting frame from the gel cassette sandwich and place the sandwich against the gasket on one side of the electrode assembly, with the short plate facing inward. Place a second gel cassette or a buffer dam against the gasket in the other side of the electrode assembly.
4. Clamp the green clamps on the sides of the electrode assembly (below).
5. Lower the chamber into the electrophoresis tank.
6. Fill the space between the two gels with Tris-glycine running buffer. This forms the upper chamber for electrophoresis.
7. Add Tris-glycine running buffer to the outer (lower) chamber until the level is high enough to cover the platinum wire in the electrode assembly.

Electrode assembly

One gel is positioned on each side of the electrode assembly

Top of short plate fits snugly into notch in the green gasket of the electrode assembly

Electrode assembly with two gels is lowered into the clamping frame



Load and run samples on the SDS-PAGE gel

1. Retrieve your cell extracts from the freezer. Recall that the samples have already been mixed with a tracking dye and glycerol.
2. Using gel loading micropipette tips (tips have very long, thin points and fit P20s or P200s), load up to 15 µL of sample into each well. Load 5 µL of a molecular weight standard into one lane of the gel. Load samples slowly and allow the samples to settle evenly on the bottom of the well.

   NOTE: Be sure to record the order of samples loaded onto the gel.
3. Connect the tank to the power supply. Fit the tank cover onto the electrodes protruding up from the electrode assembly. Insert the electrical leads into the power supply outlets (connect black to black and red to red).
4. Turn on the power supply. Run the gel at a constant voltage of 120-150 V. Run the gel until the blue dye front nearly reaches the bottom of the gel. This may take between 45-60 min.

Staining SDS-PAGE gels

1. Turn off the power supply.
2. Remove the gel apparatus from the tank. Open the clamping frame and remove the gel cassette sandwich. Carefully, pry the two plates apart with a spatula. With the spatula, remove the lower right or left corner of the gel to serve as an orientation marker. ***Be sure to indicate in your lab notebook whether the notched corner corresponds to lane 1 or lane 10 of the gel.*** You may also remove the stacking gel with the spatula, if you desire.
3. Place the gel in a small plastic tray and label the tray with your initials on a piece of tape. To do this, fill the tray about halfway with deionized water. Gently free the gel from the glass plate, allowing it to slide into the water. The gel should move freely in the water. Place the gel and tray on a rocking platform. Rock the gel for 5 minutes.
4. Drain the water from the gel and add ~20 mL of Simply Blue. Cover the gel container with saran-wrap and rock overnight.
5. In the morning, drain the Simply Blue stain into an appropriately labeled waste container in the hood of the lab room.
6. Destain the gel by filling the container about half full with deionized water. Shake the gel in the water for ~2 minutes. Pour off the water and add new deionized water. Repeat, if necessary, until protein bands become visible.
7. When individual bands are detectable, record your data. You may photograph the gel with your cell phone camera against a white background. Alternatively, place the gel in a clear plastic page protector and scan the gel.
8. After recording the data, dispose of the gel in the Biohazard waste container.

Chapter 14

References

Gelperin DM, White MA, Wilkinson ML *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

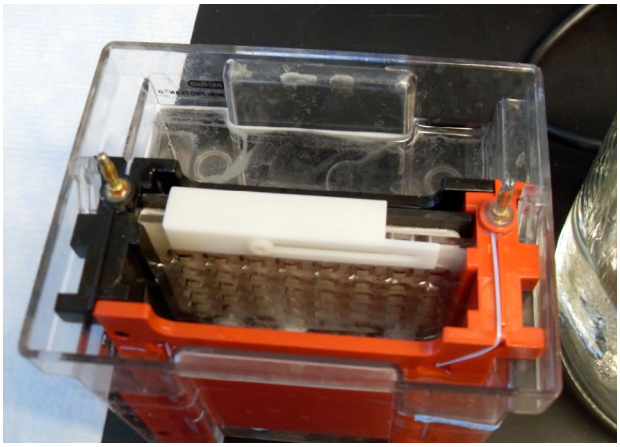
Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680–685.

Shapiro AL, Viñuela E, & Maizel JV, Jr. (1967) Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem Biophys Res Commun* **28:** 815–820.

Steinberg TH (2009). Protein gel staining methods: An introduction and overview. *Method Enzymol* **463:** 542-563.

Chapter 15

Western blots



Western blots are one of the most widely used techniques in cell biology. In a western blot, investigators take advantage of the exquisite sensitivity of antibodies to identify proteins of interest in complex samples. In this lab, you will learn about the different kinds of antibodies used in western blots. You will use western blots to analyze Met and LacZ protein expression in your transformed yeast strains.

Objectives

At the end of this lab, students will be able to:

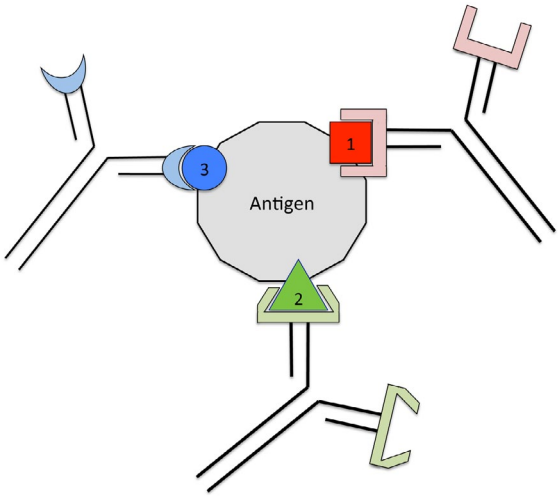
- explain how monoclonal and polyclonal antibodies are produced.
- identify the different functional regions of antibodies and explain how they are used in western blots
- design a strategy that uses antibodies to detect epitope-tagged proteins.
- prepare a western blot to analyze protein expression in cell extracts.

Chapter 15

Western blots provide a method to find the proverbial "needle in a haystack." A typical cell expresses thousands of different proteins, and it is often difficult to detect changes in expression of your favorite protein (Yfp) without a probe that is capable of discriminating the Yfp against a large background of unrelated cellular proteins. Fortunately, antibodies provide highly specific molecular probes that can be used to detect the expression of proteins on western blots. To appreciate the sensitivity of western blots, it's helpful to have some understanding of antibody structure and antibody production during immune respones. (Disclaimer: The following paragraphs provide a highly abbreviated overview of antibodies and one segment of the complex vertebrate immune system. The Department offers an immunology course that will introduce you to the finer details of this fascinating system.)

Antibodies are produced in response to antigens

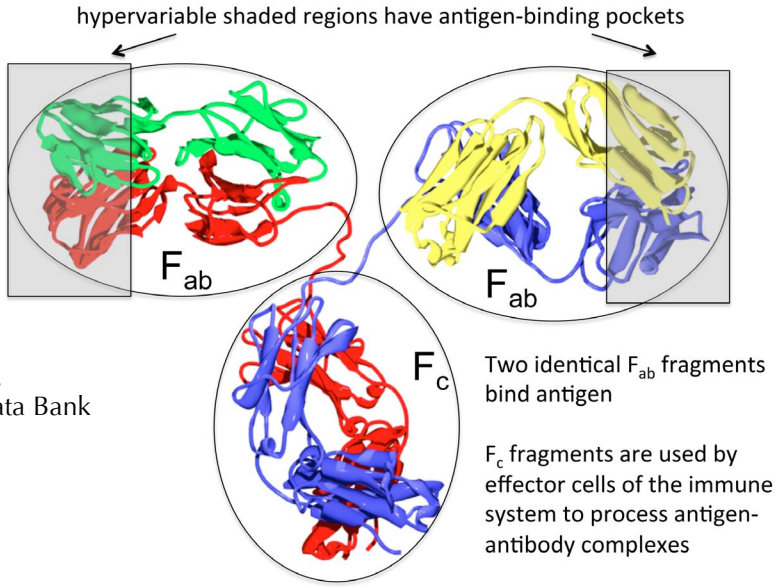
Antibodies are proteins produced by vertebrates with adaptive immune systems capable of responding to foreign antigens. Antigens are defined as substances that stimulate the production of antibodies. Antigens are commonly able to stimulate the production of multiple kinds of antibodies, each of which recognizes a small, distinct region on the surface of the antigen known as an epitope. Antibodies are Y-shaped proteins produced by lymphocytes that bind epitopes with high affinity.



Antibodies binding to an antigen.
An antigen with three different epitopes on its surface is bound by three different antibody molecules, each of which binds a single epitope with high affinity.

The availability of hybridoma cells that secrete large quantities of antibodies with a single specificity has greatly facilitated structural studies on antibodies. Researchers are able to harvest antibody molecules secreted by cultured hybridoma cells and to prepare crystals for X-ray diffraction. Based on a large number of crystallographic studies, we now understand the basic architecture of antibodies, more properly known as immunoglobins. The crystal structures show that immmunoglobins (Igs) are composed of three domains that are readily apparent in the crystal structure (below). The two F_{ab} regions (antigen-binding fragments) that form the arms of the "Y" are hypervariable regions involved in binding antigen. The F_c region (crystallizable fragment) that forms the base of the "Y" is recognized by non-immune effector cells, such as mast cells and macrophages, which process antigen-antibody complexes. Each Ig class has a characteristic heavy

chain, which gives the class its name. We are using antibodies from the IgG class of immunoglobins, which have gamma heavy chains. (IgGs are also known as gamma globulins.) IgA molecules have alpha chains, IgM molecules have mu chains, etc.



Crystal structure of an IgG antibody.
This figure is derived from Protein Data Bank entry 1IGT (Harris *et al.*, 1997).

Antibodies are produced by lymphocytes

In the initial stages of the immune response, small numbers of immature B lymphocytes are able to bind foreign antigen molecules weakly via the antibodies expressed on their surfaces. Antigen binding stimulates the lymphocytes to proliferate and to differentiate into mature lymphocytes that secrete antibodies. An amazing series of transformations occur as B lymphocytes mature in response to antigen. Antigen binding stimulates responding lymphocytes to rearrange segments of their antibody-encoding genes, producing new potential antigen-binding sites. Most rearrangements are unproductive, but some rearrangements generate antibodies with greater affinity for the antigen. Antigens act as selective agents. The lymphocytes that bind the antigen with the highest affinity receive the greatest growth signal and proliferate most rapidly, because a higher fraction of their surface antibodies are bound to antigen at any one time. In the latter stages of differentiation, a hypermutation process further increases the range of potential antibody sequences. Mature B lymphocytes that have survived the selection process are known as plasma cells. ***Each plasma cell secretes a single antibody with high affinity for antigen.*** Plasma cells are virtual antibody factories that can be identified in electron micrographs by their extensive rough endoplasmic reticulum. (See the inside cover for a graphical depiction of this process.) The scope of antibody diversity is immense - vertebrates are capable of producing billions of antibody molecules with distinct specificities.

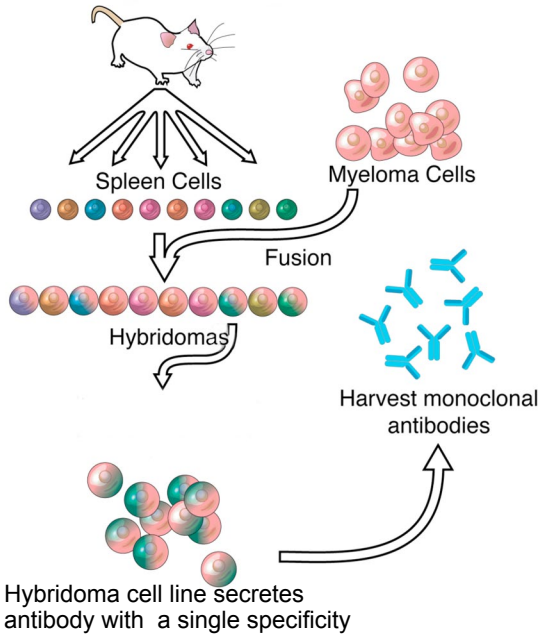
Polyclonal vs. monoclonal antibodies

For our western blots, we will be using both monoclonal and polyclonal antibodies. As their names imply, monoclonal antibodies bind to the same epitope on an antigen. Polyclonal antibodies are actually mixtures of antibodies that bind to different epitopes on an antigen. An

animal's response to antigen is polyclonal, because antigens stimulate the proliferation of multiple lymphocyte clones, each of which produces a different antibody to the antigen. Consequently, the serum collected from an immunized animal contains a mixture of antibodies with different specificities. The polyclonal antibodies used in the lab are purified from the sera of animals that have been inoculated with antigen.

By contrast, monoclonal antibodies are produced in the lab from cultured hybridoma cells. Hybridoma cells are generated by fusing a lymphocyte from an immunized animal, most commonly a mouse, with a cancerous myeloma cell that can divide indefinitely in culture (right). Because the lymphocytes from the spleen of an immunized mouse recognize a range of different epitopes on an antigen, the hybridomas resulting from the fusion secrete a variety of different antibodies. Standard culture techniques are then used to isolate individual hybridoma cell lines, each of which secretes a unique antibody that binds to a single epitope.

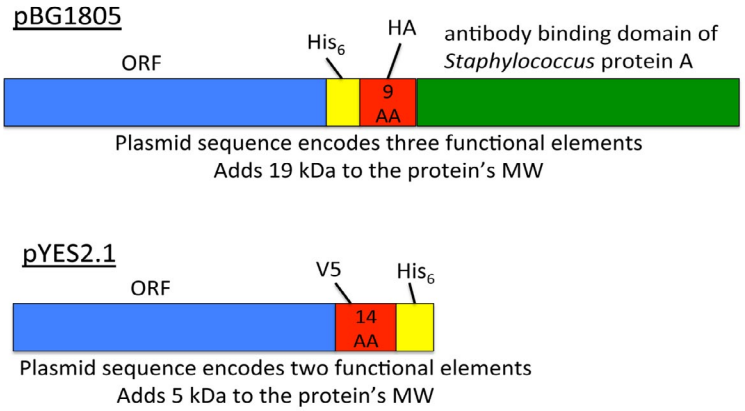
Hybridoma technology has revolutionized biomedical research since its description (Kohler & Milstein, 1975), both because monoclonal antibodies recognize well-defined epitopes and because monoclonal antibodies can be produced indefinitely by cultured hybridoma cells. Investigators often use both monoclonal and polyclonal antibodies at different steps in western blots.



Construction of hybridoma cell lines that secrete monoclonal antibodies. Spleen cells from an immunized mouse are fused with myeloma cells to produce hybridoma cells. *Adapted from http://en.wikipedia.org/wiki/Monoclonal_antibody#mediaviewer/File:Monoclonals.png*

Plasmid-encoded proteins have C-terminal tags

In this lab, we will be using antibodies to detect Met and LacZ fusion proteins expressed in transformed *S. cerevisiae*. These fusion proteins have been engineered to add several functional elements to the C-termini of Met and LacZ proteins. Consequently, the fusion proteins expressed from the pBG1805 (Gelperin *et al.*, 2005) and pYES2.1 plasmids have C-terminal tags that can interact with a variety of reagents. Each of the plasmids encodes a short epitope tag that can be detected with antibodies on western blots. These epitopes correspond to highly immunogenic amino acid sequences on the surfaces of viruses that have been shown to be potent inducers of antibody production. The pBG1805 sequence encodes an 9-amino acid sequence of the human influenza virus hemagglutinin (HA) protein (Sleigh *et al.*, 1981), while the pYES2.1 plasmid encodes a 15-amino acid sequence from the P protein of simian virus V5 (Southern *et al.*, 1991). In our blots, we will use a monoclonal antibody directed against the V5 protein, hereafter referred to as anti-V5, to detect proteins expressed from pYES2.1-based plasmids.



Fusion proteins expressed from pBG1805 and pYES2.1 plasmids have additional functional elements at their C-termini.

Both plasmids also encode a tag consisting of 6 histidines. This His_6-tag binds tightly to metal ions, so it is commonly used to purify overexpressed proteins by passing them through a resin with immobilized zinc or cobalt ions. Unfortunately the His_6-tag is not very immunogenic and is therefore not useful in western blots.

To detect proteins expressed from pBG1805-based plasmids, we will use an unusual strategy that exploits the ZZ domain from *Staphylococcus aureus* protein encoded in the large C-terminal tag of pBG1805. *S. aureus* protein A is an integral membrane protein with multiple extracellular ZZ domains, each of which is able to bind the F_c domain of an IgG. *S. aureus* uses protein A to evade detection by a host's immune system by indiscriminately sequestering host IgGs. The secondary antibody that we are using for the blots was produced in rabbits, and its F_c domain binds tightly to protein A.

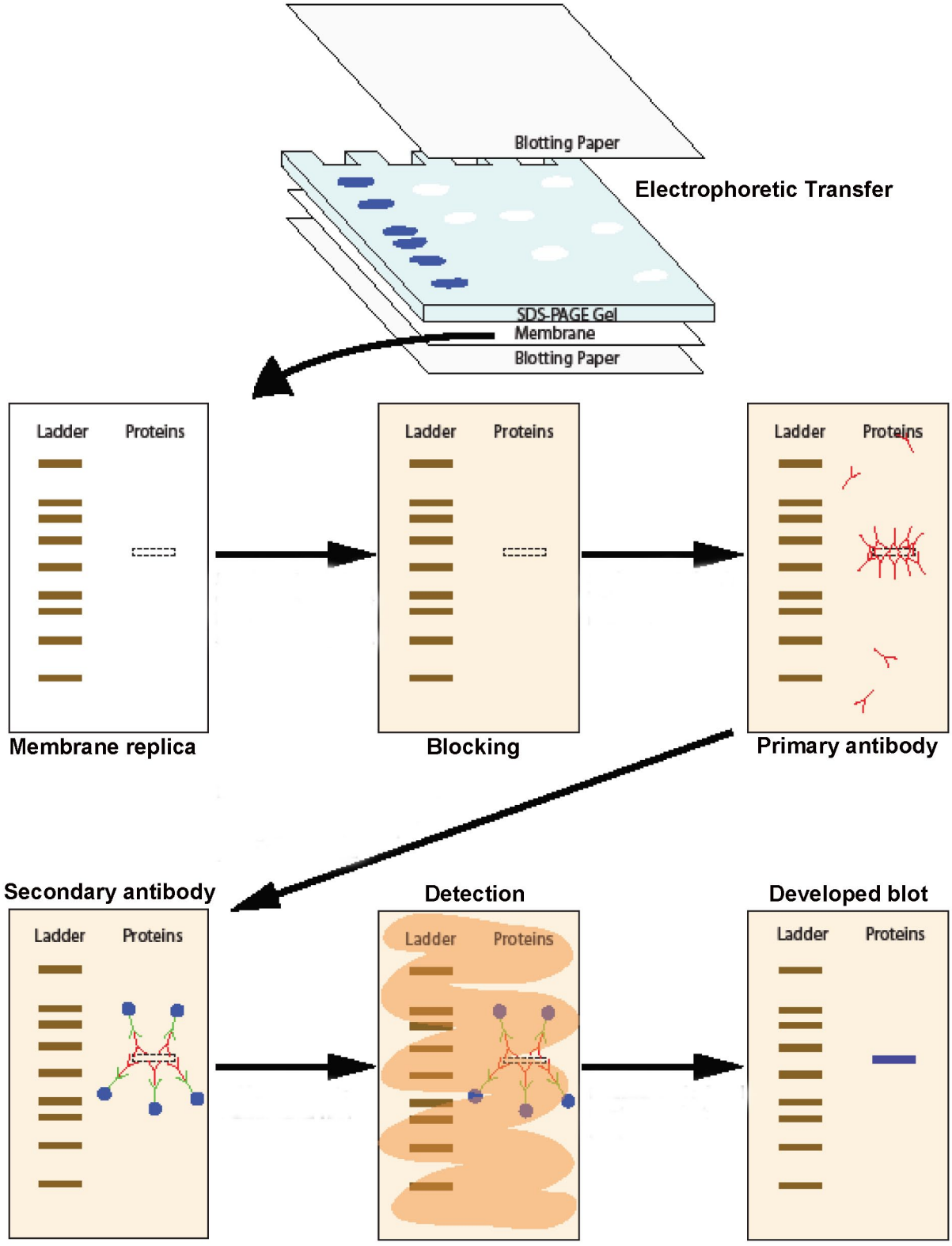
Western blots involve many steps

In a western blot procedure, proteins are first separated on an SDS-PAGE gel and then transferred to a membrane. This membrane replica is treated with antibodies that specifically recognize a protein or epitope of interest. Additional processing steps generate a signal at the position of the bound antibody. Between the steps, various washes are done to increase the signal-to-noise ratio on the final, developed blot. The major steps in a typical western blot are diagrammed on the following page and discussed in greater detail in sections that follow:

- Electrophoretic transfer of proteins from an SDS-PAGE gel to a membrane
- Blocking of nonspecific protein binding sites on transfer membranes
- Incubation of the membrane with a primary antibody specific for the epitope of interest
- Incubation with a secondary antibody that recognizes primary antibodies
- Visualization of bound antibodies

Electrophoretic transfer of proteins from an SDS-PAGE gel to a membrane

The first step in a western blot is to generate a replica of the SDS-PAGE gel by transferring proteins electrophoretically to a synthetic membrane with a high protein binding capacity. In our experiments, we will use membranes made of polyvinylidine fluoride (PVDF), a kind of plastic. PVDF membranes are hydrophobic and the dry membranes do not wet properly with water. Therefore, PVDF membranes are first wet with methanol, then rinsed with deionized water, and


Blotting Paper
Electrophoretic Transfer
SDS-PAGE Gel
Membrane
Blotting Paper
Ladder
Proteins
Membrane replica
Blocking
Primary antibody
Secondary antibody
Detection
Developed blot

finally rinsed with transfer buffer. They must not be allowed to dry out during the transfer and immunoblot procedures. If they do dry out, they must be re-wet with methanol and rinsed with water before proceeding.

During the transfer process, the gel and membrane are placed directly against each other within a "sandwich" of pre-wet filter papers and foam pads, as described in Exercise 1. During the electrophoretic transfer, current should flow evenly across the entire surface area of the gel. It is important, therefore, that air bubbles are not trapped between the gel and membrane. After the electrophoretic transfer, which can be done in a few hours or overnight with reduced voltage, the membrane replica with the transferred proteins can be allowed to dry out and stored for later visualization with antibodies.

Blocking of non-specific protein binding sites on membranes

The transfer membranes used in western blots bind proteins nonspecifically. Before the membranes are incubated with specific (and expensive) antibodies, they must be pretreated with blocking solutions that contain high concentrations of abundant (and cheap) proteins to saturate non-specific binding sites. Think of this step as analogous to an artist priming a canvas with a lower quality paint before the more expensive media is applied. If the transfer membranes are not adequately blocked before the antibody is applied, the nonspecific sites on the membranes will absorb some of the antibodies, reducing the amount of antibody available to bind the target proteins. In our experiments, we will use casein proteins from milk as blocking reagents. Because our experiments do not require high sensitivity, rehydrated non-fat dry milk (direct from the grocery store!) is an adequate source of caseins.

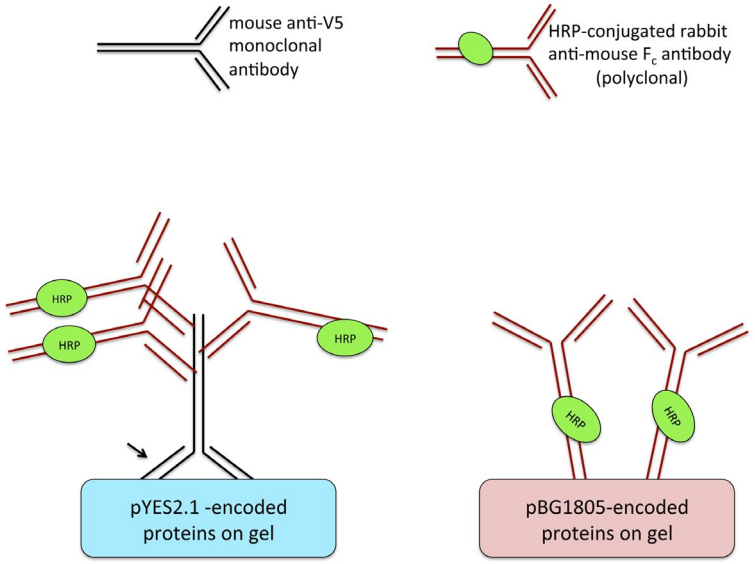
Primary antibody binding

Either polyclonal or monoclonal antibodies can be used as the primary antibody on western blots. Antibodies can be directed toward a naturally-occurring protein or toward an epitope attached to an overexpressed protein (as we are doing). Increasingly, researchers are using epitope-tagged proteins in their experiments, because antibodies against naturally-occurring proteins are expensive and time-consuming to prepare. In addition, an antibody directed against an epitope can be used to detect many different proteins carrying that same epitope. In our western blots, we will use a mouse monoclonal antibody that binds the V5 epitope on Met and LacZ proteins expressed from the pYES2.1 plasmid.

Secondary antibody binding

The secondary antibodies used in western blots are designed to bind the F_C fragments of primary antibodies, taking advantage of cross-species differences in antibody sequences. Secondary antisera are generally prepared by injecting an animal with F_C fragments of IgGs from a second species. The first animal recognizes the F_C fragments as foreign antigens and produces antibodies that bind the F_C fragments. The secondary antibody in our experiment is a rabbit polyclonal antibody prepared against the F_C domains of mouse IgGs. The antibody will bind the F_C domains of the mouse anti-V5 antibodies bound to the pYES2.1-encoded proteins.

In our experiments, we are also able to take advantage of the fact that rabbit F_C fragments also bind tightly to *S. aureus* protein A. Thus, we will also be able to use the secondary antibody to detect primary anitbody bound to the *S. aureus* ZZ domains in pBG1805-encoded sequences.



Visualization of bound antibody

In this final step, the western blot is incubated with substrates for the enzyme that has been conjugated to the secondary antibody. Our secondary antibody has been conjugated to HRP, a hardy enzyme with a high turnover number. (The turnover number is the number of product molecules produced at an enzyme's active site per second.) HRP catalyzes the reaction of hydrogen peroxide and 3,3',5,5' - tetramethylbenzidine (TMB), which generates a dark blue-grey reaction product that precipitates at the reaction site on the western blot. Colored reaction product accumulates with time until the reaction is stopped by washing away unreacted substrate. The reaction should be terminated before nonspecific antibody binding becomes problematic.

Exercise 1 - Preparing the membrane replica

Separate proteins on an SDS-PAGE gel

1. Separate the proteins that will be analyzed on western blots by SDS-PAGE.
2. Remove the electrode apparatus and holder from the tank, and remove the gel from the holder. Do not remove the gel from the plates until you are ready to assemble the transfer cassette (see below).
3. Dispose of the remaining buffer down the sink. Rinse out the buffer tank with deioinized water to remove residual SDS, which can interfere with the transfer process.

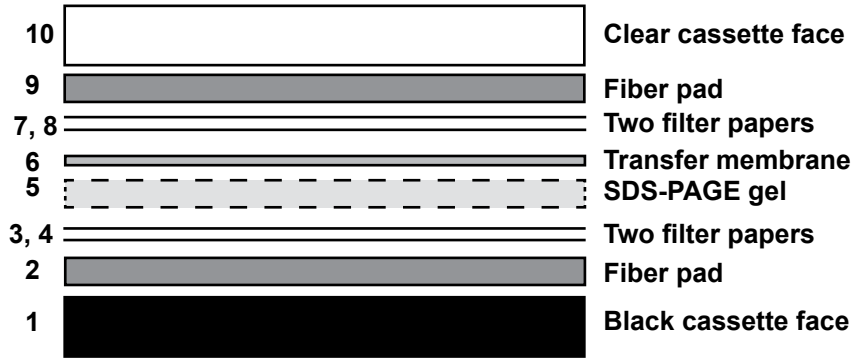
Prepare the transfer membrane

NOTE: DO NOT touch transfer membranes with your fingers. Wear gloves and use filter forceps when you handle transfer membranes.

1. Gather the PVDF membrane and two pieces of thick filter paper, such as Whatman 3MM™. The PVDF membrane and filter papers should be cut to a size that is slightly larger than the SDS-PAGE gel. You will also need a transfer cassette and two fiber pads.
2. Prepare the PVDF membrane. Using pencil, place an orientation mark in a corner of the PVDF membrane for later identification. Wet the membrane by placing it in a small tray containing methanol for ~30-60 seconds with gentle agitation.
3. Dispose of the methanol in the waste container and add deionized water to the tray. Gently agitate for ~1 minute.
4. Replace the deioinized water with transfer buffer. Store the membrane in transfer buffer until you are ready to start the transfer.

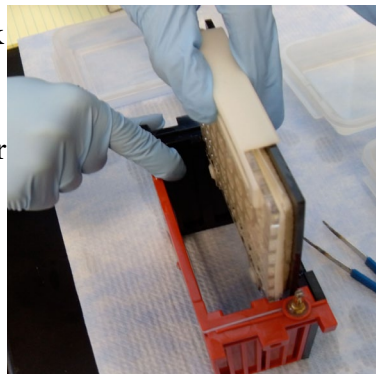
Assemble the transfer cassette

1. Using a spatula or a green plastic wedge, remove the small glass plate from the gel. The gel will remain attached to the large glass plate. With a spatula, remove the lower right corner of the gel to serve as an orientation mark. (This correponds to the first lane of your gel.)
2. Assemble the transfer cassette as shown below. Be sure that all parts of the transfer "sandwich" remain moist at all times.



- Place a wet fiber pad (2) on top of the black cassette face (1).
- Add two pieces of filter paper (3,4).
- Position the gel (5) on top of the filter paper while it is still attached to the glass plate. Use a spatula to carefully release the gel from the plate. You may find it easier to remove the gel by beginning at the bottom edge near the dye front.
- Place the PVDF membrane (6) on top of the gel. Orient the gel so that the pencil mark on the membrane corresponds to the clipped corner of the gel. **Be sure that there are NO air bubbles between the gel and the membrane.**
- Add the remaining filter papers (7,8) and the fiber pad (9).
- Fold the clear cassette face (10) over the gel assembly and carefully slide the clamp into place.

Electrophoretic protein transfer



1. Place the transfer sandwich into the cassette holder with the black face of the transfer cassette aligned with black side of the cassette holder and the clear face aligned with red side of the cassette holder (right). NOTE: Each cassette holder can hold two transfer cassettes.
2. Place the cassette holder and assembled cassettes into the electrophoresis tank. If the transfer will be done at 100V (see step 5 below), add an ice pack to the tank.
3. Fill the electrophoresis tank to the top with transfer buffer.
4. Place lid on tank by aligning black with black and red with red.
5. Transfer proteins at 100 V for 1 hour at room temperature or at 20 V overnight in the cold room.
6. When the transfer is complete, remove the transfer cassette from the tank. Pour the transfer buffer back into its original bottle so that it can be reused.
7. Disassemble the transfer cassette. Depending on your schedule:

- If you will be continuing with the western procedure, skip the rehydration step (step 1) below and continue with the blocking step (step 2). Be careful that the membrane remains moist!
- If you will be processing the membrane at a later time, allow the membrane to dry out. Wrap the membrane in plastic wrap and save it for a later lab period.

Exercise 2 - Immunodetection

This is a multi-day procedure. Timing may vary for different classes.

Membranes are rehydrated and treated with blocking reagents

1. Wearing GLOVES, unwrap the dry blot from the plastic wrap. Use the prestained standards to identify the side of the membrane to which the proteins are bound. Submerge the membrane in methanol with this side facing up. Gently agitate the membrane by hand rocking for 30-60 seconds until the membrane has been uniformly wet with methanol. Decant the methanol into the appropriate container and fill the tray half way full with deionized water. Gently agitate the membrane for an additional minute.
2. Decant the water and replace it with sufficient TBS-T (Tris buffered-saline containing 0.05% Tween 20) to cover the blot. Place the blot on a rocking platform. Equilibrate the blot in TBS-T for 5 minutes with slow rocking. At the end of 5 minutes, drain off the TBS-T.
3. Pour enough blocking solution (5% nonfat milk in TBS-T) onto the blot to cover it.
4. Cover the tray with a small piece of plastic wrap. Place the tray on a rocking platform in the cold room. The blot should float freely in the tray so that both sides are washed. Incubate the blot for at least an hour or up to 24 hours at 4˚C.

Membranes are washed and incubated with primary antibody (~24 hours)

1. Locate your blot in the cold room and bring it back to the lab room.
2. Remove the plastic wrap from the container holding the blot and pour off the blocking solution. SAVE the plastic wrap! You will need it to cover the container again!
3. Add enough TBS-T to cover the blot and place the container on the rocking platform. Rock for 5 minutes.
4. Pour off the TBS-T. Add 15 mL of primary antibody diluted in blocking buffer.
5. Cover the container with the same piece of plastic wrap and place the tray on the rocking platform in the 4˚C cold room. Make sure that the blot floats freely in the tray and that the standards are on the top face of the blot. Incubate overnight at 4˚C with slow rocking. NOTE: The timing of this step is the most critical in the procedure. Shortening the incubation time with primary antibody may reduce the sensitivity of the western blot.

Secondary antibody binding and detection (1.5-2 hours)

1. Locate your blot in the cold room and bring it to your lab classroom.
2. Carefully drain the antibody from the blot into the test tube marked "Used primary antibody". (Antibodies are expensive. Fortunately, the solutions can be re-used.)
3. Fill the tray with the blot about half-full with TBS-T. Place the tray on a rocking platform and wash the membrane for 5 minutes to remove unbound primary antibody. Drain the TBS-T when the wash is complete.
4. Repeat step 3 once more, for a total of two washes.
5. Add enough secondary antibody to cover the blot and incubate the membrane for 1 hour with gentle rocking at room temperature. The secondary antibody, which is conjuated to horseradish peroxidase (HRP), has been diluted in blocking solution.
6. Carefully drain the antibody from the blot into the test tube marked "Used secondary antibody."
7. Wash the membrane 3 times for 5 minutes each with TBS-T, as in step 3.
8. Drain the TBS-T from the blot. Using a P1000 micropipette, cover the blot with 1 mL of 3,3'5,5'-tetramethyl benzidine (TMB), a colorigenic substrate for HRP. Let the color continue to develop until distinct bands are apparent. Bands will probably become apparent within minutes. Do not allow the blot to over-develop, when nonspecific bands become apparent.
9. Stop color development by diluting the substrate with an excess of deionized water. Drain the diluted substrate into the waste container.
10. Allow the blot to dry on a piece of filter paper. Record your data with the scanner.

Chapter 15

References

Gelperin DM, White MA, Wilkinson ML *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

Harris LJ, Larson SB, Hasel KW & McPherson A (1997) Refined structure of an intact IgG2a monoclonal antibody. *Biochemistry* **36:** 1581-1597.

Kohler G & Milstein C (1975) Continuous cultures of fused cells secreting antibody of predefinied specificity. *Nature* **256:** 495-497.

Sleigh MJ, Both GW, Underwood PA & Bender VJ (1981) Antigenic drift in the hemagglutinin of the Hong Kong influenza subtype: Correlation of amino acid changes with alterations in viral antigenicity. *J. Virol.* **37:** 845-853.

Southern JA, Young DF, Heaney F, Baumgartner WK & Randall RE (1971) Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. *J Gen Virol* **72:** 1551-1557.

Write It Up!

A guide to lab reports

The goal of our class this semester has been to determine if *S. pombe* orthologs of the *S. cerevisiae MET* genes are able to complement deficiencies in *S. cerevisiae met* strains. Throughout the semester, you will be collecting data and reporting that data in 5 micro-reports. At the end of the semester, the data will be compiled into a final report and a poster. The final report should be written in the form of a scientific publication, following the format of a *FEMS Yeast Research* paper. This chapter contains general and specific guidelines for preparing microreports, the final report and the poster.

The focus of each micro-report is a figure, sometimes accompanied by a table, with the experimental data. The micro-reports are scheduled so that you will receive feedback on your presentation and interpretation of the data, as well as your scientific writing, throughout the semester. You can expect to see improvement as the semester progresses. The experiments in the five micro-reports are listed below in chronological order. Figures and tables are prepared by the team. Before each microreport is due, teams will post their figure to the class data sharing site and present their figure to the class for discussion and feedback.

1. Identification of *met* deletion strains by their growth phenotypes on defined media containing various sulfur sources
2. Genotype analysis of *met* deletion strains by yeast colony PCR
3. Identification of yeast overexpression plasmids by restriction mapping
4. Transformation of *met* deletion strains with plasmids and complementation analysis of the transformed strains
5. Analysis of Met protein overexpression in transformed *met* deletion strains by SDS-PAGE and western blotting of cell extracts

Before proceeding, consider some general tips for good scientific writing:

- Good scientific writing is **CONCISE**! A rambling report quickly bores the reader and weakens the message of the report. Observe all page limits. Abbreviations are often helpful. With the exception of standard abbreviations, e.g. length and degrees, spell out the abbreviated term the first time that you use it and follow the term with the abbreviation in parentheses.
- Good scientific writing is **PRECISE**! Use the correct conventions for strains, genes and proteins. This class introduces you to a large number of scientific terms, which are defined in the glossary section. Use of the correct term can often prevent the need for extra words.
- Good scientific writing is **EFFECTIVE**! The same skills that produce a great essay on Henry Thoreau's *Walden* are relevant here! Use the active voice when possible. Writing should be smooth, not choppy. Avoid run-on sentences and be sure that antecedents are clear.
- Use sub-headings to divide the text into logical segments.

Micro-reports - general guidelines

Page limit: 2 PAGES, excluding figures, tables and references

Micro-reports have a condensed format. ***You can think of each micro-report as the equivalent of a figure or table from a scientific publication, together with the methods needed to understand the experiment and a brief analysis.***

Before each micro-report is due, teams will present and discuss their data with their section. This can provide value feedback for the report. Teams are also advised to look at the results obtained by students in other sections who worked with the same strains and plasmids. Good science is reproducible. Did other students get the same results as you did? Do **NOT** reproduce the results from other sections in your report (remember the plagiarism and academic integrity rules!), but it is fine to discuss these results with proper attribution.

Micro-report organization. Include the following sections:

1. Heading

Your name
Section number or TA's name (*e.g.* Jen's section)

2. Purpose - state the purpose of the experiment in **one sentence**.

3. Methods and Materials

The Materials and Methods (M&M) section should be written in paragraph format. Methods should NOT be written as lists of steps, as they might appear in your notebook or in a recipe. Avoid excessive detail. For example, DON'T state: "The solution was prepared by adding 5 µL of 200 mM NaCl to 95 µL of deionized water. " Instead, state: "The solution contained 10 mM NaCl." A reader who chooses to repeat your experiment may have his/her own way of preparing the solutions. The final concentrations of components are the important consideration.

If you are using a published technique, you can cite the procedure without reproducing the detailed steps. In this course, you will probably find it convenient to frequently refer to procedures in the lab manual. If you are using a commercial kit, *e.g.* the Zyppy™ plasmid purification kit, you can state that you followed the manufacturer's instructions. In all cases, be sure to include any modifications to the published procedure.

Rule of thumb: A good M&M section provides enough information for a trained professional to reproduce your experiments.

4. Results and Discussion

The section contains a brief narrative that guides your reader through the figures and legends that present your experimental data. **Figures** need to be clearly labeled and to be accompanied by a figure legend. The figure legend should have a title and include explanations of

the different panels or graphs in the figure. The figure legend should be placed below the figure. A well-written legend contains enough information that an expert reader can understand the experiment shown in the figure from its legend (assuming that the reader also looks at the M&M section).

Tables should have a title. Columns and rows should be clearly labeled and the units of measurement (e.g. grams/liter) should be included. When appropriate, include statistical measures of error in both figures and tables.

Report the results of all your experiments, even if you think they are incorrect. Discuss any experimental problems that you encountered in the experiment and speculate how these could have affected the results. Compare your results to those posted by other groups on the data sharing site. If your results agree, you may feel more confident about your results. Propose further experiments to resolve any remaining questions.

State your conclusions in one or two sentences.

5. Thought question

The rubric for each micro-report will include a thought question that requires you to apply your conceptual knowledge to a novel situation. **NOTE: This question is worth a significant portion of your grade for the report.**

6. References

Use the FEMS Yeast Research format for citations and references. Cite the lab manual: O'Connor CM (2014) *Investigations in Molecular Cell Biology*. Hayden-McNeill, Plymouth, MI.

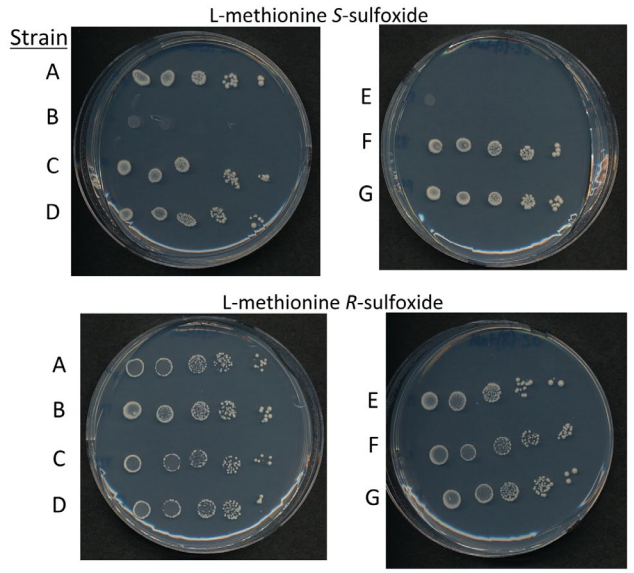
Document format:

- Reports must be typewritten
- Use 1-inch margins on all sides and double line spacing
- Use a font that generates less than 15 characters per inch

The next few pages provide specific guidelines for the five micro-reports of the semester. ***A rubric will be posted before each micro-report is due. The rubric will contain specific details about the grading for each micro-report. The rubric will also contain the thought question for the micro-report. It may also contain some changes and/or additions to the guidelines in this chapter. As such, the rubric should be considered to contain the definitive set of instructions.***

Micro-report 1: Selective plating experiment

The figure is at the heart of every micro-report. This multi-panel figure of a spot plating experiment was prepared by students in an advanced lab class.



Growth on restrictive media containing methionine sulfoxide diastereomers Strains A-G were spot plated on SC media containing either 20 mg/L L-methionine-*S*-sulfoxide or 20 mg/L L-methionine-*R*-sulfoxide in the place of methionine. Strains A-F were constructed from the genotype *met15 his3 leu2 ura3*. Strains carried the additional mutations, as follows: A (none), B (*msra*), C (*msrb*), D (*frmsr*), E (*msra msrb*), F (*msra frmsr*), G (*msrb frmsr*). Plates were incubated for 3 days at 30 degrees C.

Note the following features of the figure:

- Plates are oriented in the same direction so they can be easily compared
- Strains and media are labeled. Strain names can be used in the figure itself or indicated by a code that is defined in the legend.
- Strain names are used when the genotypes of strains are uncertain (your experiment). This figure presents results from an experiment where genotypes were known.
- The reader will need to refer to the M&M section for additional details about the media and experiment.

Specific guidelines for micro-report 1 follow.

Purpose: You have three yeast strains derived from strain, BY4742. In one sentence, what are you trying to do in this experiment?

Materials and Methods: In preparing the M&M, ask yourself "What information will an investigator need to reproduce our experiments?" Provide information on the strains and media that you used, as well as the procedures that you used for spot plating.

Strains: Include the names of your strains as well as the genotype of the BY4742 parent strain.

Media: Identify the culture media you used in the experiments. Decide on a naming convention - the same nomenclature should be used in both the figure and M&M. Reference the manual for the composition of the media, rather than including all the components here.

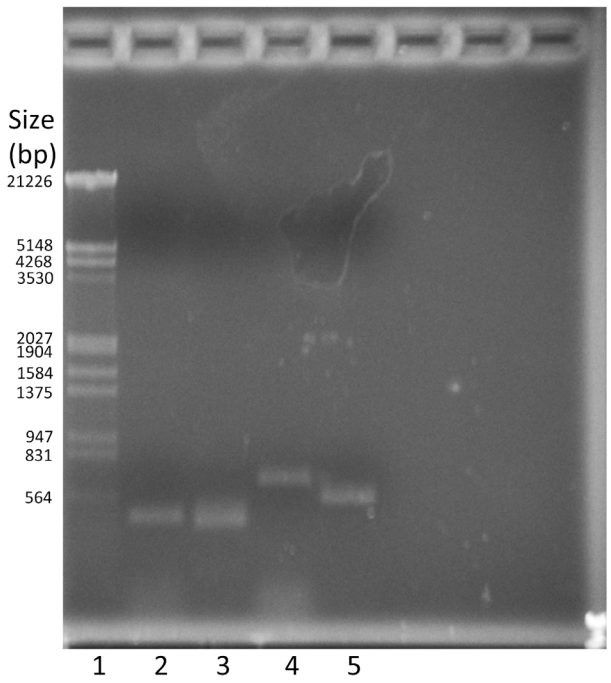
Spot plating: Someone trying to reproduce your results will need to know how your starter cultures were generated (cultures were grown overnight in YPD) and how cultures were diluted for the spot plates (*e.g.* a series of 1:10 dilutions in sterile water). They do NOT need to know that you transferred 10 µL yeast culture to 90 µL water. Readers will need to know how many microliters were used for each spot and the conditions used to incubate (time, temperature) the plates.

Results and Discussion - Your figure with the scanned plates is the focal point of this section. The R&D section tells a story of how you used your plating data to identify the *met* deletions in the YMP strains. Provide a brief narrative that guides your reader through your results and your thinking. How would deletions in your *MET* genes affect the ability of your YMP strains to use different sulfur sources? Do the data allow you to confidently identify the strains?

A **single** summary data table documenting the growth of YMP and BY4742 strains on various culture media is a good way to bring together the experimental data and your conclusions. Describe the growth of the *met* deletion strains on the various media and include your preliminary strain identifications from the experimental data.

Micro-report 2: Yeast colony PCR

In this report, you will describe the results of the yeast colony PCR experiment that you designed to identify the *met* deletions in your three YMP strains. The products of your PCR reaction will be separated on agarose gels similar to this one prepared by students in another class.



Colony PCR of strain BY4742 with gene-specific primers for *MET* genes. Gene-specific primers A and B were used for each PCR reaction. PCR products were separated by electrophoresis on 1% agarose gels. The sizes of the molecular weight standards (lane 1) are shown on the left. Primers used for the reactions correspond to *MET1* (lane 2), *MET3* (lane 3), *MET7* (lane 4) and *MET8* (lane5).

Note the following features of this figure:

- The gel is shown with the loading wells at the **top**.
- Lanes are clearly marked. In this figure, the lanes are numbered. The strains and primers used for each reaction are described in the legend. Alternatively, the lanes could be labeled with the names of strains and primers, taking care that they remain readable. Conclusions are **NOT** included in the legend.
- The sizes of the markers are indicated. (Your markers will be different.)

The reader will need to refer to the M&M for additional details on the strain genotypes and the PCR procedure.

Specific guidelines micro-report 2 follow.

Write it up!

Materials and Methods: Provide information on the specific strains and primers that you used, as well as the procedures for PCR and agarose gel electrophoresis.

Strains: See micro-report 1 guidelines.

Primers: In a publication, authors usually include the sequences of their PCR primers in the text or a table. You should include the names of the *MET* genes, but you do NOT need to include the actual sequences, because these are listed in Chapter 7 of the manual. Be sure to cite the manual, however. Use the correct terminology when you refer to primers. GSP-A primers are sense primers that correspond to 5'-flanking sequences of the *MET* genes. GSP-B primers are antisense primers that are complementary to ORF sequences.

PCR: The PCR methods are described in the manual. "PCR reactions were performed as described (ref.)" is adequate. If you made any modifications, however, you need to describe them *e.g.* ..."with the following modifications"....

Agarose gel electrophoresis: The PCR reaction products were analyzed on agarose gels that were stained with ethidium bromide. Describe the conditions that you used to analyze the PCR products. Include the concentration of agarose, the name of the buffer and the voltage. In very general terms describe this analysis. Refer readers to the manual for additional details.

Results and Discussion: The figure with your gel is the focal point of this section. Briefly describe your experimental design and the logic underlying your choice of primers.

As you discuss your results, remember that the bands you see on the agarose gel are **PCR products**. Size is very important in interpreting the results of a PCR experiment. As you prepared for the experiment, you calculated the sizes of PCR products that would confirm/not confirm your predictions. Did your experimental results confirm your predictions?

If your results did not confirm your predictions, you need to consider if the predictions were wrong OR if the reactions did not work. Did you see bands in any lane? If so, the PCR reagents were probably not at fault. What else might have happened?

A **single** summary data table with the predicted and observed sizes of the PCR products would be helpful. The predicted sizes should very accurate, since they are based on the actual genome sequence, which has nucleotide resolution. Your estimated sizes are much less accurate. You may be able to determine if one PCR product is smaller than another, but you will only be able to place the products within a certain size range.

Micro-report 3: Restriction mapping

Figure: The figure at the heart of this micro-report is an agarose gel containing the restriction digests of your team's plasmids. *Refer to the guidelines for Micro-report 2 to see how to set up figures with agarose gels.* The reader should be able to determine from the gel and its legend which plasmid is shown in each lane and which restriction endonuclease (RE) was used in each digest.

Materials and Methods: To reproduce your experiments, a reader will need to know which plasmids and REs that you used, how you purified the plasmids and some information about the agarose gels used for the restriction maps. When possible, refer to published procedures, noting any modifications to a published procedure. You may want to use subheadings for the M&M section.

Plasmids: The *MET* genes and homologs have been cloned into two different yeast overexpression plasmids. Use the correct nomenclature when referring to plasmids of known genotypes. Plasmid names begin with a lower case "p" and are written in normal font. For much of this report, you will need to refer to the plasmids by their number alone, since you are still trying to identify them.

Plasmid purification: There are many different methods for isolating plasmids, so inform the reader that you used the Zyppy™ Plasmid Mimiprep Kit (Zymo Research) with the modifications described in the lab manual. Include information on how you estimated plasmid DNA concentrations as well as the DNA concentration of each purified plasmid.

Restriction digests: To reproduce your experiments, the reader will want to know how much plasmid DNA you used in the digestions and which REs you used. The ratio of RE units to DNA is important for successful digests. Your reader may be using a different commercial source of REs, so also provide the number of enzyme units (U) in the reaction. Do NOT list the µL of each component used in the reactions. Be sure to include information about the temperature and duration of the incubation.

Agarose gel electrophoresis: See the guidelines for micro-report 2.

Results and Discussion: Bring the reader through the logic for your experimental design. Why did you choose this particular RE? Refer to the bands on the gels as **restriction fragments**. Restriction fragments are the products of **RE digestions**.

Size is very important in interpreting the results of a RE digest. A single summary data table with the predicted and observed sizes of the restriction fragments should be included. Estimate the sizes of the restriction fragments by comparing their migration to the markers. Predicted sizes should very accurate, since they are based on the actual DNA sequences, which has nucleotide resolution. Your estimated sizes are much less accurate. You may be able to determine if one restriction fragment is smaller or larger than another, but you will only be

able to place the fragments within a certain size range. **NOTE:** circular plasmids run anomalously on agarose gels. Supercoiled plasmids migrate more rapidly on gels than a linear DNA molecule of the same size. Conversely, nicked (a single strand of the helix has been cleaved, producing a relaxed circle) plasmids migrate more slowly than linear DNA of the same size. Incomplete digestion can also complicate results. ***Keep in mind that the number of bands on the gel is not as important as the differences in the banding patterns!***

Did your experimental results confirm your predictions? If your results did not confirm your predictions, you need to consider if the predictions were wrong OR if the reactions did not work. The undigested plasmid provides a control for interpreting the latter issue.

Micro-report 4: Complementation analysis

Figure: The figure at the heart of this micro-report is multi-panel figure showing replica plates of strains that have been transformed with overexpression plasmids. *Refer to the guidelines for Micro-report 1 to see how to set up figures with multiple panels.*

Materials and Methods: Provide information on the transformation and replica plating procedures, as well as the media used in the experiments. When possible, reference the lab manual, noting any changes that you made to procedures.

Strains and plasmids: See micro-reports 1 and 3.

Media: See micro-report 1.

Transformation: Someone trying to reproduce your results will need to know details about the transformation procedure, because transformation efficiencies vary widely and show a strong dependence on reagents and incubation conditions. You can reference the lab manual.

Replica plating: This is a standard procedure. You can also reference the lab manual.

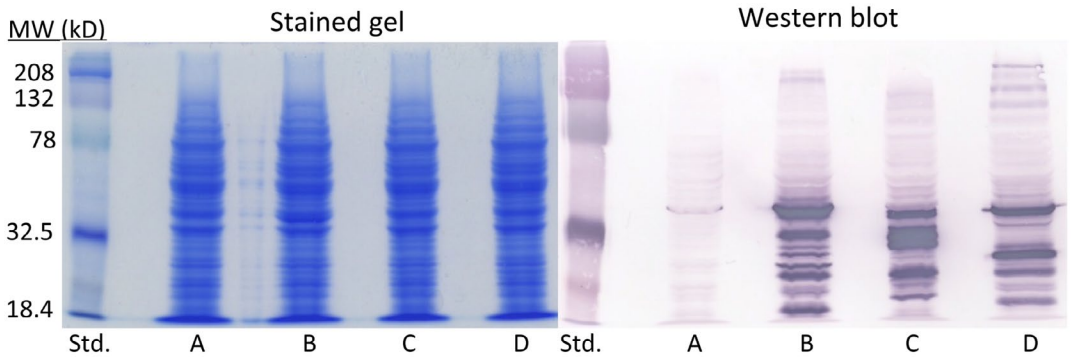
Results and Discussion: Your figure with the scanned plates is the focal point of this section. Include a single data table that with the calculated transformation efficiency with each plasmid and the ability of transformed strains (Y/N or +/-) to grow on the various replica plates. *Transformation efficiencies should be expressed in number of transformed cells/μg plasmid DNA.*

The R&D section should tell a story of how the replica plates allowed you to decide if your Met protein is conserved between the two yeast species. You may or may not have observed complementation. Failure to observe complementation is not necessarily due to experimental error. (Which plates serve as a control against experimental error?) Complementation is a functional assay that depends on both the expression of the fusion protein and the ability of the fusion protein to catalyze a reaction in Met biosynthesis. The fusion proteins have large C-terminal extensions that might affect their normal enzymatic functions. Negative results can be just as important as positive results in advancing scientific understanding.

If you did not observe complementation, discuss possible reasons that this may have happened and propose future experiments that could help to answer these questions. You may want to suggest new plasmid constructs for additional experiments. You may also want to bring in information from your BLASTP analyses. (What is the E-value?) Be sure to include enough justification that your proposed experiments will provide useful data.

Micro-report 5: SDS-PAGE and western blot analysis

In this report, you will describe the results of the SDS-PAGE and western blots that you used to analyze protein expression in your transformed cells under both repressed and induced conditions. Pay special attention to the degree to which these results confirm or contradict the results of the previous transformation/complementation report. The figure should be labeled in such a way that an experienced scientist is able to understand your results from the figure and legend alone. Many experimental details have been relegated to the M&M section. Lanes should be clearly labeled and the molecular weights of the standards should be included. (The example below is from a different class, where students used SDS-PAGE and western blots to analyze overexpression of yeast proteins expressed from pBG1805. Note that, unlike your experiments, a primary antibody to the HA epitope was used to detect proteins on western blots.)



Analysis of MSR overexpression in transformed yeast cells. Extracts were prepared from untransformed BY4741 cells (A) and from BY4741 cells transformed with pBG1805-*MSRA* (B), pBG1805-*MSRB* (C) or pBG1805-*FRMSR* (D), as described in the Materials and Methods. Protein overexpression was induced by incubating cells in media containing 2% galactose for 4 hr at 30 ˚C.. Proteins were separated on a 12% gels, which were stained with Simply Blue® (left) or used for western blots (right). Western blots were incubated with a mouse anti-HA primary antibody and a goat anti-mouse secondary antibody. The molecular weights (MWs) of the protein standards (Std.) are shown on the left.

Materials and Methods: Provide information about transformed strains, incubation conditions, preparation of cell extracts, SDS-PAGE gels and western blots. Reference published procedures when possible, noting any modifications. Subheadings may be helpful.

Transformed strains: Include the names of the strains and plasmids that you used to prepare cell extracts.

Extracts: Include information on the media and incubation times used to manipulate protein overexpression from the strains. Reference the manual for the extraction procedure, noting any modifications.

SDS-PAGE gels: Provide details about the % acrylamide of the gels, running time and voltage used for electrophoresis.

Electrophoretic transfer: Include the time and voltage used to transfer proteins from the SDS-PAGE gel to the PVDF membrane. Refer to the manual for other details.

Western blot: Include the antibodies that you used for the blot and the conditions (time, temperature) that you used for each of the incubations. Include the time that you used to detect overexpressed proteins with TMB. (This gives a some sense of the abundance of the protein in your extracts.) You do NOT need to include all the wash steps - just reference the manual.

Results and Discussion - Begin by discussing the **SDS-PAGE gel**. The SDS-PAGE gel provides a snapshot of cellular proteins and a rough comparison of protein concentrations in different extracts. The staining intensity of a band reflects its abundance in the extract.

- How did the total amount of protein compare between induced and non-induced samples? (What might this indicate about the different carbon sources?)
- Did you see any changes in individual bands on the gel? Is it possible to detect the fusion protein against the background of other proteins in the extract? Recall that cells have thousands of proteins and that a band may consist of more than one protein species.

The **western blot** allows you to detect the fusion protein against the background of other cell proteins. Include a table showing predicted and actual sizes of Met or LacZ proteins detected using the western blot technique. The sizes of the proteins are particularly important. (It is probably not possible to get an exact value of the protein sizes, because of the fuzziness of the standards on the western blots. Nonetheless, you should be able to place the proteins within a certain range.) The plasmid-encoded proteins will be larger than the naturally occurring protein because of the epitope tags encoded by the plasmid. Are the observed sizes what you expected?

Other questions to address in the Discussion:

- Did you see evidence of induction by the *GAL1* promoter on either the gel or blot?
- How did the western blot results relate to your complementation results? How do these results enrich/weaken your complementation findings?
- If you did not detect proteins on your blot, propose some explanations. Compare the samples. Do the other samples provide good positive controls for your technique?

Final lab report

The research goal of our class is to determine if the *S. pombe* orthologs of *S. cerevisiae MET* genes can complement deficiencies in *S. cerevisiae met* strains. During the semester, you have been collecting data and organizing the data into figures and tables. You will assemble these draft versions into a final report. The data hopefully form a nice story that answers your research question and/or suggests additional experiments.

The final report should be in the form of a scientific publication, which includes the numbered sections below. Unlike the mini-reports, the Results and Discussion sections in the final report should be separate. As you write, keep in mind that ***good scientific writing is precise and concise!*** In the professional world, journals provide an economic incentive for brevity by levying page charges on authors- the longer the report, the higher the page costs to the authors! Be careful to observe the page maximum for each section.

Experiments to include in the final report include:

- Growth of your strains on selective media
- Yeast colony PCR
- Restriction map of your overexpression plasmids
- Replica plates of transformed strains
- SDS-PAGE and western blot analysis of cell extracts
- Independent experiment (placed where most appropriate within the text)

You are welcome to cut and paste from your earlier micro-reports. *After all, it is your work and you are welcome to reproduce it!* As you do so, be well-advised to incorporate the suggestions of your TA to improve the presentation. The only data that should be included in the final report is data that your team has generated. You should NOT include data from experiments that has been posted by other students on the data sharing site.

COVER PAGE
Include the title and abstract, as well as your name and section assignment

1. Title - Describe your project in 150 characters or less.

2. Abstract (1 paragraph, 150 words maximum) - ***Single-spaced***
The abstract is a very brief summary of your work that should be comprehensible to the non-expert. The abstract should include the goals of your experiments, some mention (not description) of the methods used in the experiments, a succinct summary of your experimental results, and your overall conclusions. Avoid specific details of the experiments. A reader should be able to read the abstract and understand the overall experiment and results without reading the rest of the report. ***Compose the abstract AFTER the rest of the report has been completed.***

BODY OF THE REPORT - ***Double-spaced (Note that this is different from the abstract.)***

3. Introduction (1.5-2 pages)

The introduction provides the context to your experiments. What is the goal of your experiments? How does your experiment fit in with what's already known about your topic? Has anyone done experiments similar to yours before? To answer these questions, an introduction provides the reader with relevant background information derived from the scientific literature. Discuss the function of the enzyme and its evolutionary conservation. Use the format of *FEMS Yeast Research* to insert citations into your report. ***Include at least 7 citations from the scientific literature.***

In the last paragraph of the introduction, give the reader a preview of your report. Tell the reader what experimental approaches you used to answer your question. End this section with one or two sentences summarizing your conclusions.

4. Materials and Methods (2-3 pages)

During the course of the semester, you have used a variety of techniques, which have been the subject of your micro-reports. In publications, the M&M section is usually divided into sub-sections, often with their own subheadings, reflecting the different experiments and techniques.

The first paragraph in an M&M section often contains details about the organisms used in the experiments and their culture conditions. When yeast and other microorganisms are used for the experiments, a genotype table is often included. ***Include a genotype table with table with strain and plasmid information as Table 1.*** Consult Cordente *et al.* (2009) for a good example. Include the three YMP strains that you worked with at the beginning of the semester in this table. Our YMP strains were derived from strain BY4742, which has the genotype *Matα his3Δ1 leu2Δ0 ura3Δ0*. Each of the YMP strains has a *met* deletion caused by the insertion of a KAN^R gene. The *met* deletion alleles should be referenced as: *met#:: KAN^R*. (See Chapter 6.)

The plasmids for this report are based on pYES2.1 and pBG1805, both of which drive expression using the *GAL1* promoter. To describe these overexpression plasmids, use the format:

pBG1805-*GAL1:MET#* or pYES2.1-*GAL1:Met#*

Briefly summarize the methods used in the experiments with appropriate subheadings. Feel free to cut and paste from previous micro-reports. Add additional information for your independent experiment. Remember - if you are using a published technique (or reagent), you do not need to repeat all the details in the M&M section. Instead, you may refer to the published procedure or formulation. In this course, you can refer to a protocol in the lab manual, but you need to include any modifications that you have made to the procedures or recipes.

Rule of thumb: ***A good M&M section should provide enough information for a trained professional to reproduce your experiments.***

5. Results - (2-3 pages, not including tables and figures)

The Results section tells a story. It provides a brief narrative that guides the reader through the figures and tables with the data. Smooth transitions are important in the Results section. The reader should be able to grasp the reasoning that led from one experiment to the next. Do NOT discuss the results in any depth in this section - there is a separate discussion section in this report. Provide just enough of your conclusions that the transition to the next experiment is a logical one. For example, your selective plating data influenced your choice of primers for the yeast colony PCR experiment. (Your tables actually include some of these conclusions.) Because different kinds of experiments are frequently used in a paper, authors frequently use subheadings in the Results section.

Figures and tables from the micro-reports should be given numbers that correspond to their order of appearance in the text. **PLACE** the figures and table within the Results section ***NEAR*** where they are referenced. (Because of page breaks, it is not always possible to place the figure or table on the same page as it is referenced. Do NOT simply attach the tables and figures to the end of the report. The Cordente et al. (2009) paper provide examples of numbering and placement. Follow the general guidelines for micro-reports. Remember that a trained biologist should be able to understand your results by simply "looking at the pictures."

6. Discussion - 2 pages maximum

Our goal this semester was to determine if Met protein functions were conserved between *S. cerevisiae* and *S. pombe*, using complementation to analyze protein function. As you write this report, address the following issues:

- Did you observe any differences in the abilities of *S. cerevisiae* and *S. pombe* genes to complement the *met* mutation in your strain? Is your complementation data consistent with conservation?
- Reconcile the complementation data with the SDS-PAGE and western blot results. Were Met proteins of the expected size expressed in transformed cells? How does complementation relate to expression of the Met protein?
- If you did not observe complementation, was this because of technical issues or because of protein divergence between the two yeasts? How would you test this?
- How did your independent experiment confirm or extend your results?
- What additional experiments should be done to prove to the scientific community that MetXp function is/is not conserved between the two yeast species?

7. References - Single-spaced

List the references using the *FEMS Yeast Research* format, as you did in the bibliography assignment.

Write it up!

Poster

A good poster communicates data concisely and effectively. This poster presents the same data as your final report. Unlike the final report, which is read carefully by a single reviewer, your audience is a larger number of individuals who "visit" your poster. Your poster audience will be more diverse than your class section and they may have limited familiarity with your project. (Friends and family are welcome!) *The poster will have much less detail than your final report.* You will be able to fill in the missing details during your conversations with your visitors. Since the visitors will spend only a few minutes at your poster, so you need to communicate your results efficiently. Always consider your audience as you prepare your poster. Your visitors may provide helpful ideas and suggestions for your final report.

Follow these general guidelines:

- Posters are 30" wide and 40" high. Each is mounted on an easel. Your material needs to fit within these boundaries.
- The poster presents data and background material in a **series of panels**. Each panel is mounted separately on the poster board with push pins. Panels can be mounted in either landscape or portrait orientation. You may group several panels together. ***Each panel or group of panels needs to have a title in BOLD font.***
- Use large font sizes so that a viewer can read the text from a distance of 1-2 feet. Use BOLD fonts for titles. (Specific suggestions for font sizes are included below.)
- Panels with figures or tables should have a legend. The legend title should be larger or emphasized to set it aside from the text of the legend.
- Prepare the panels in a program(s) of your choice. Print the panels on regular letter paper. Include color if you wish, but this is optional.

Panels to include on your poster:

1. Title - This panel includes the title and the names of the authors in a smaller font. The title summarizes the content of your poster. The title MUST include the name of your gene, *e.g. MET3*. A font size of 48 pt or more for the title and a font size of 28-32 points for the authors' names are recommended.

2. Abstract - The title of this panel is "Abstract". "Abstract" should be in a larger and bolder font (24-28 pt) than the text (18-20 pt) of your abstract.

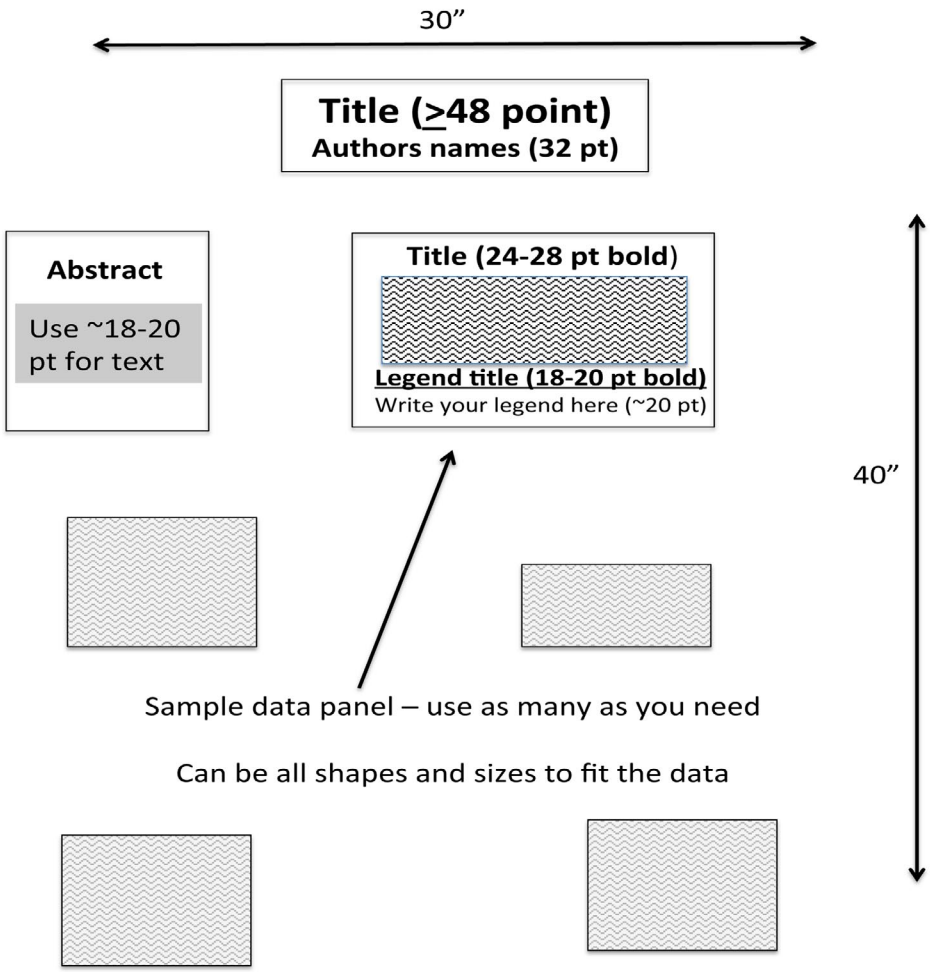
3. Background - Your visitors will not be as familiar with your project as you are. It would be a good idea to provide some information about the methionine biosynthetic pathway, as well as other information on your gene that you consider relevant.

4. Data panels - Use as many panels as you need to tell your story! Each data panel needs a title, figure or table, and legend. The title of a data panel often summarizes the content, *e.g.* "Plasmid identification using restriction digestion." Font sizes decrease as you move through the panel. The title of the figure legend should be smaller (18-20 pt) than the font size of the panel title (24-28pt).

5. Conclusion panel - Include a short list of conclusions from your work.

7. References - Use "References" or "Literature cited" as the title of this panel. List the references in *FEMS Yeast Research* format.

A mockup of a potential poster is shown below.



Organization is critical to an effective poster!

Be creative! Make your poster visually appealing!

References

Cordente AG, Heinrich A, Pretorius IS & Swiegers JH (2009) Isolation of sulfite reductase variants of a commercial wine yeast with significantly reduced hydrogen sulfide production. *FEMS Yeast Res* **9**: 446-459.

Glossary

A

absorbance
unitless spectrophotometer data; values are directly proportional to the concentration of a light-absorbing substance

accession number
unique identifier for a record in a database

agar
algal polysaccharides with gelling properties that are useful in preparing solid growth media for microorganisms

agarose gel electrophoresis
method for separating nucleic acids; because nucleic acids have a uniform negative charge, separation is primarily determined by size

alleles
alternative forms of a gene

amino acid
monomeric subunit of proteins; a central carbon binds a proton, an amino group, a carboxyl group and a variable side chain known as the R (reactive) group

anneal
binding of two single-stranded nucleic acids with complementary sequences; the double-stranded hybrid is stabilized by hydrogen bonds

annotation (genome)
process of analyzing genome sequences to identifying ORFs and functional elements

antibody
an immunoglobulin; protein produced by B lymphocytes that binds a specific antigen with high affinity

antigen
substances that stimulates the production of antibodies by lymphocytes

autoclave
chamber for sterilizing media and equipment using pressurized steam

auxotroph
microorganism that has lost its ability to synthesize an essential nutrient due to a gene mutation

B

bacteriophage
virus that infects bacteria cells

bioinformatics
interdisciplinary research area concerned with the development of computational methods to analyze, store and access DNA and protein sequences

biosynthesis
process by which enzymes in living cells catalyze the synthesis of a biological molecule.

bp
base pair; lengths of DNA molecules are expressed in base pairs

Glossary

C

clone
population of genetically identical cells that are descended from a single cell

cofactor
nonprotein component required by enzymes in order to function

colony
visible population containing hundreds of millions of gentically similar organisms, most likely descended from a single cell

competent
able to take up DNA from the environment

complementation
restoration of a normal phenotype in a mutant cell by introduction of a gene that encodes a functional version of mutant gene product

complementation group
functional definition for a gene derived by studying collections of mutant organisms. Cells in the same complementation group carry mutations in the same gene.

complex (molecular)
association of molecules held together by noncovalent bonds

curator
professional who maintains and organizes records in databases

cuvette
glass or plastic sample holder used with a spectrophotometer

D

database
organized collection of records, each of which has searchable fields

denaturant
treatment or agent, such as excessive heat, extreme pH, or chemical modification, that causes denaturation of proteins or nucleic acids

denaturation
destruction of native protein or nucleic acid structures, usually resulting in the loss of biological function

diploid
organism with two copies of each chromosome

domain
region/part of a protein that folds independently of other parts of the protein; domains are often functional units of a protein

E

enzyme
protein catalyst that accelerates the rate of a chemical reaction by lowering the activation energy

epitope
portion of an antigen that binds an antibody

F

functional genomics
research area that studies the conservation of gene function during evolution

fusion protein
protein expressed from a genetically engineered plasmid that contains non-native functional elements at its N-terminus or C-terminus

G

genome
the entire DNA sequence of an organism

genomics
the study of genomes, as opposed to individual genes

genotype
the complete genetic makeup of an individual organism

H

haploid
organism or cell containing only one set of chromosomes

homolog
gene with high sequence similarity to another gene due to shared ancestry

homologous recombination
genetic recombination that normally occurs during meiosis; identical sequences on different DNA molecules align and initiate crossovers

I

inoculation
deliberate introduction of a microorganism into a medium that supports growth and reproduction

intercalating agent
chemical, such as ethidium bromide, that inserts into the DNA helix

L

lag phase
growth period following inoculation of a liquid media, during which cells acclimate to the new environment and condition the media with their own metabolites

log phase
growth period where the number of cells increases exponentially

lymphocyte
immune cell that produces antibodies

M

media
mixture of chemicals used to support the growth of microorganisms, generally including a carbon source, nitrogen source, salts, vitamins and essential minerals

microorganism
small prokaryote or eukaryote; visualization of individual microorganisms requires a microscope

model organism
organism that has been extensively studied in the laboratory for the insights it provides to higher eukaryotes; methods for genetic maniopulation are usually available; common model organisms include yeast, nematodes, fruit flies and zebrafish

monoclonal antibody
antibody binds to a defined epitope on an antigen. Monoclonal antibodies are produced by hybrid cell lines derived by fusing a lymphocyte with a meyloma cell.

mutant
organism with a mutation in a gene of interest

mutation
permanent change in a DNA sequence

O

oligonucleotide
short, single-stranded polymer of nucleotides

optical density
spectrophotometer measurement of light scattering caused by suspensions of particles or cells

ORF (open reading frame)
potential protein coding sequence; stop codons are underrepresented in ORFs

ortholog
similar DNA sequences that have arisen from a common ancestral gene and are now found in different species

P

paralog
similar sequences within the same genome; paralogs arise by gene duplicatiion events

PCR (polymerase chain reaction)
technique used to produce millions of copies of a DNA sequence of interst from a small number of template molecules

phenotype
observable physical properties of an organism

plasmid
small, circular DNA that replicates independently of thehost cell chromosomal DNA

R

recombinant DNA
DNA molecule formed by combining segments of DNA from different sources; constructed by molecular biologists using enzymes such as polymerases, ligases and restriction endonucleases

restriction digest
incubation of DNA with a restriction endonuclease, which cleaves the DNA into a collection of fragments

restriction endonuclease
enzyme that cleaves DNA at a defined sequence of bases; also known as a restriction enzyme

restriction fragments
products of a restriction digest generated by endonucleolytic cleavage of a large piece of DNA

restriction map
map of the restriction endonuclease recognition sites in a length of DNA

restriction site
nucleotide sequence in a DNA molecule that is specifically recognized and cleaved by a restriction endonuclease

R group
side chain that distinguishes an amino acid; also known as the reactive group

S

SDS-PAGE
Sodium **D**odecyl **S**ulfate-**P**olyacryamide **G**el **E**lectrophoresis; the separation of proteins through a polyacrylamide gel matrix using an electic field

selective medium
growth medium that enhances the growth of some organisms, while preventing the growth of others, due to an added or absent media component

species
group of organisms capable of breeding with each other and producing fertile descendants

spectrophotometer
instrument that measures the amount of light absorbed by a sample at a particular wavelength

stationary phase
growth period that follows log phase and can be seen as a plateau in the growth curve, as cells begin to deplete available nutrients and growth rate slows

strain (genetic)
microorganisms of the same species that are descended from a single cell; members of the same strain have identical genotypes

supernatant
the liquid above the pellet after sedimentation or centrifugation

T

transformation
uptake of DNA from the environment, causing a change in a cell's phenotype

transformant
cell that has taken up DNA and expresses genes from the DNA, affecting its phenotype

V

vector
in molecular biology, a plasmid or virus used to transfer genes into a cell; usually a recombinant DNA constructed by molecular cloning

Y

YPD
rich, but undefined, media which supports yeast growth, consisting of **Y**east extract, **P**eptone and **D**extrose

Glossary