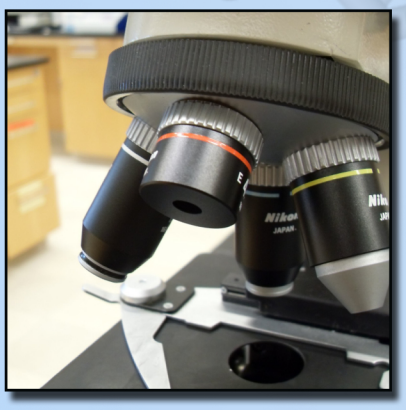
Investigations in Molecular Cell Biology

Author: Clare O'Connor

2013



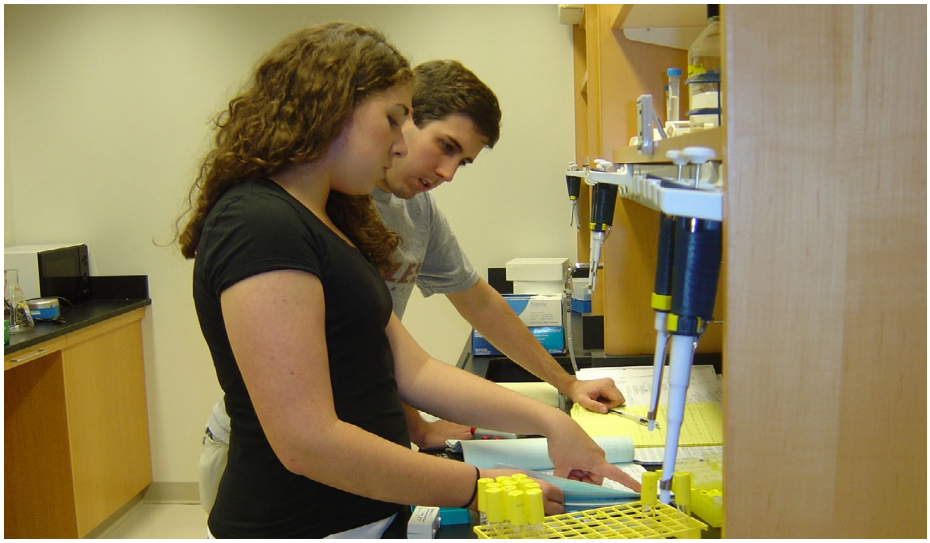
Investigations in

Molecular Cell Biology

Boston College | Clare O'Connor

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 BI204, we will use some of the strain and clone resources generated by the yeast genome project to investigate the evolution of genes involved in the synthesis of methionine and cysteine, essential sulfur-containing amino acids.

BI 204 - Investigations in Molecular Cell Biology

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

Chapter 1

Welcome to BI204 - Investigations in Molecular Cell Biology. BI204 is a new kind of introductory lab course, which has been designed to incorporate an authentic research project. 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. Many of the "genes" annotated by the genome projects have been identified by their similarity to known genes in other organisms, but their functions have not been tested directly (Goffeau *et al.*, 1996). These loose ends provide exciting opportunities for undergraduate students to participate in authentic functional genomics research.

This course is designed as a research project in which students study the evolution 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.

The BI204 research project focuses on the evolutionary conservation of the genes involved in synthesizing Met and Cys. The experiments in the project explore 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.

Introduction

During the course of the semester, students will acquire the core competencies for professional biologists.

- Working in teams, students will propose hypotheses and design experiments to test their hypotheses.
- Students will learn basic skills of molecular cell biology as they conduct their experiments.
- Students will collect, organize and interpret experimental data
- Students will find and use information from the primary scientific literature and online databases as they develop their experimental design and interpret their experiments.
- Students will communicate their scientific results in a series of short oral presentations and written reports.
- Using feedback from their peers and the teaching staff, students will compile these short reports at the end of the semester into a poster and a final report written in the format of a scientific presentation.

Pathways over time: our research project

We will use the budding yeast, *Saccharomyces cerevisiae*, to study the evolution of the genes involved in sulfur amino acid synthesis. *S. cerevisiae* is a unicellular eukaryote that has been widely used as a model organism for over 50 years (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. *S. cerevisiae* is also inexpensive and simple to culture in the laboratory. For these reasons, the *S. cerevisiae* genome was the first to be sequenced in its entirety. Completion of the yeast genome sequence (Goffeau *et al.*, 1996) allowed the yeast community to prepare genome-wide collections of mutant strains (Winzeler *et al.*, 1999) and plasmids (Gelperin *et al.*, 2005), some of which you will use this semester. We will use *S. cerevisiae* strains with defined defects in methionine and cysteine biosynthesis as the hosts for homologous genes from other organisms. If the foreign gene restores the ability to synthesize methionine to its *S. cerevisiae* host, in a process known as **complementation**, we will know that gene function has been conserved over the evolutionary time frame that separates the two species.

Chapter 1

During the 2013-2014 academic year, we will explore the conservation of Met and Cys biosynthetic enzymes between *S. cerevisiae* and the fission yeast, *Schizosaccharomyces pombe*. Both *S. pombe* and *S. cerevisiae* are ascus-forming yeast from the phylum Ascomycota in the kingdom Fungi. 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. Today, the size of the *S. cerevisiae* genome (Kellis *et al.*,2004), ~12.5 Mbp, is similar to that of *S. pombe*. *S. pombe* is considered to be much closer to ancestral members of the phylum. Our results during this coming year will most likely identify both genes that have and have not been functionally conserved between the two yeast species. Results from this year will also guide the future direction of the project into other branches of life. Over time, we hope that student results will allow us to construct evolutionary trees for many of the genes involved in Met and Cys synthesis.

Course overview

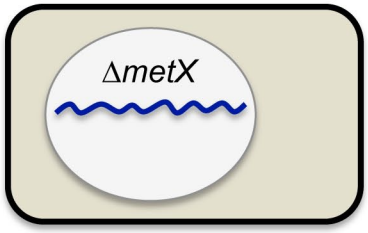
The course can be viewed as a series of related modules, as show in the figure on the opposite page.

1. In the first few laboratories, 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. In the next set of experiments, student teams will characterize three *S. cerevisiae* mutants, each of which is deficient in a gene involved in Met or Cys synthesis (Winzeler *et al.*, 1999). Students will use selective growth media and the polymerase chain reaction (PCR) to distinguish the three strains from one another. Teams will then use one of the strains for the remaining experiments of the semester.
3. Teams will receive three plasmids that will be used for transformation and complementation experiments. The plasmids have been engineered to overexpress protein coding sequences (CDS) under the control of the powerful yeast *GAL1* promoter. One plasmid contains the *S. cerevisiae MET* gene that is deficient in their yeast strain (Gelperin *et al.*, 2005). The second plasmid carries the *S. pombe* homolog for the *S. cerevisiae MET* gene that is missing in their yeast strain. The third plasmid is a control plasmid to ensure that overexpression is occurring in transformed cells. Students will isolate the plasmids from bacterial stocks and identify the plasmids using restriction endonucleases.
4. Teams will transform (Chapter 12) their yeast deficiency strain with the three plasmids. Selective plates will be used to determine if transformation and/or complementation has occurred.
5. Teams will study expression of the plasmid-encoded genes using western blots. The proteins expressed in transformed cells are fusion proteins carrying epitopes at their C-termini that can be recognized by antibodies.
6. In the last segment of the semester, 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)
Meet the yeast (Ch. 2)
Micropipettes and measurement (Ch. 3)
Yeast culture techniques (Ch. 4)
Databases (Ch. 5)

2. Identify yeast deletion strains

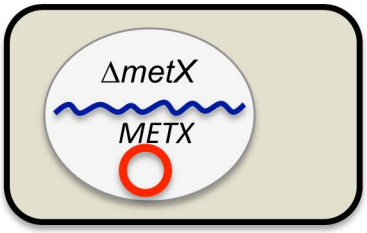


Identify missing genes in Δ*met* strains by:
ability to grow on sulfur sources (Ch. 6)
yeast colony PCR (Chs. 7-8)

3. Characterize *MET* expression plasmids

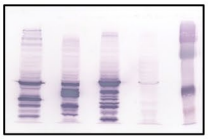
Identify *S. pombe* homolog (Ch. 9)
Isolate plasmids (Ch. 10)
Restriction mapping (Ch. 11)

4. Transform strains 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

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.

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.

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

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.

Acknowledgments

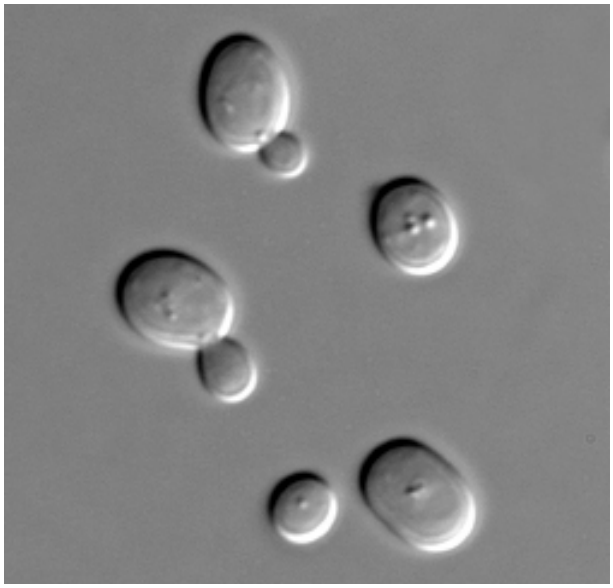
It has been a pleasure to work with many Boston College colleagues and students on this third edition of the BI204 laboratory manual. In particular, Dr. Douglas Warner has helped to design the experiments and has contributed many revisions to the text. Many thanks to Holli Rowedder for her careful editing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

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

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

Chapter 2

Meet the yeast



Since the budding yeast (left), *Saccharomyces cerevesiae*, and the fission yeast, *Schizosaccharomyces pombe* diverged from a common ancestor, they have evolved distinctive morphologies and controls on cell division. In this lab, you will use the light microscope to compare cultures of *S. cerevisiae* and *S. pombe*.

Objectives

In this laboratory, students will:

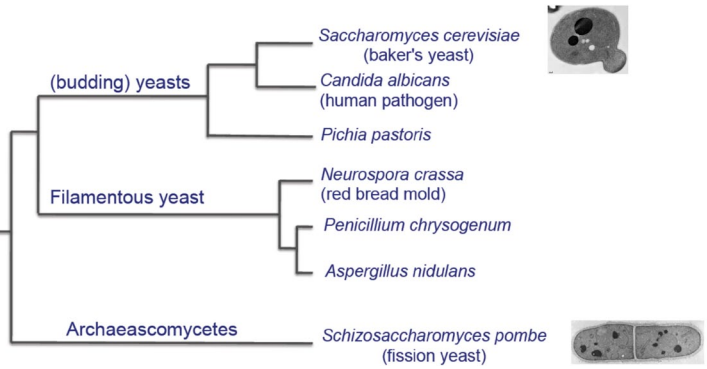
- identify the components of a compound light microscope
- adjust a light microscope to observe to different yeast specimens
- stain yeast samples with iodine to improve contrast
- identify the morphological characteristics of *S. cerevisiae* and *S. pombe*
- distinguish log phase and stationary phase yeast cultures

Chapter 2

Two very different yeast

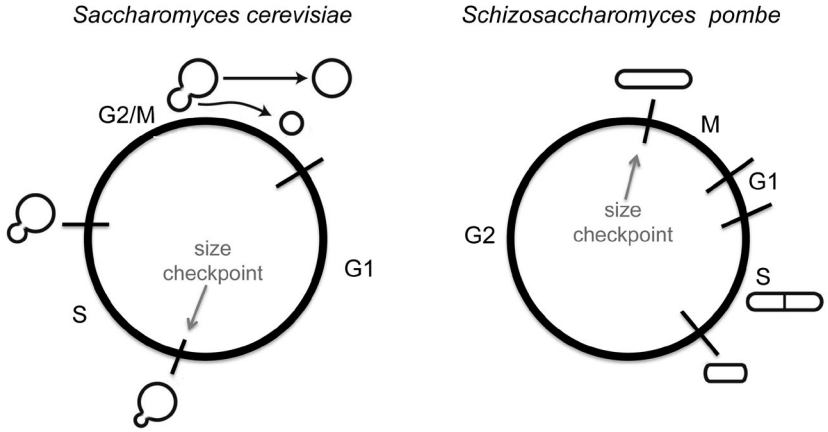
As their names imply, the budding yeast, *Saccharomyces cerevisiae*, and the fission yeast, *Schizosaccharomyces pombe*, are sugar-loving fungi that were originally isolated from beer. *S. cerevisiae* has been important to human civilization for millenia, because of its various roles in the preparation of wine, bread and beer. Over the past century, scientists have worked with genetically pure strains in the laboratory (Mortimer, 2000). *S. cerevisiae* and *S. pombe* are members of the ascomycota phylum, which can exist in both diploid and haploid forms. In response to various stresses, haploid strains of opposite mating type are induced to mate and undergo meiosis. The four spores generated from meiosis are contained within a resistant structure known as the ascus, which gives the phylum its name, ascomycota.

Diversification of yeast in the Phylum Ascomycota.
Most phylogenetic trees predict that the budding and fission yeast diverged from a common ancestor ~1 billion years ago. *S. pombe* is considered more similar to the common ancestor. 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.)



In this lab, you will use a compound light microscope to observe *S. cerevisiae* and *S. pombe* from both rapidly-growing and nutrient-rich and nutrient-depleted cultures. Our microscopes do not have the sophisticated optics used to obtain the images shown above, but you will be able to distinguish cells at various points in the cell cycle and to distinguish *S. cerevisiae* and *S. pombe*. In nutrient-rich media, cells grow exponentially, and this period of rapid growth is often referred to as log phase. Log phase yeast pass continuously through the cell cycle, and a log phase culture will have cells in the G1, S, G2 and M phases of the cell cycle.

As shown in the figure on the opposite page, however, the proportion of cells in each phase varies signicantly in *S. cerevisiae* and *S. pombe* cultures, because the principal cell cycle checkpoint occurs at a different place in the cycle (Turner *et al.*, 2012). In *S. cerevisiae*, 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. At the time that 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. By contrast, *S. pombe* divides by medial fission. Cells grow in length until they are 12-15 μm, at which point 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. One would expect that haploid yeast to be more susceptible to adverse effects of spontaneous mutations than diploid yeast, in which deleterious mutations may be masked by a functional second allele.



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.

When nutrients are depleted, however, cells need to down-regulate their metabolism and enter a stress-resistant state. 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). Cells can survive in stationary phase for extended periods of time, resuming growth when conditions are favorable. *S. pombe* enters stationary phase more rapidly than *S. cerevisiae* which passes through an intermediate metabolic stage in which the rate of cell division is sharply reduced before it enters true stationary phase. The "stationary" phase *S. cerevisiae* cells in this experiment are in this intermediate state.

Observing microorganisms with light microscopy

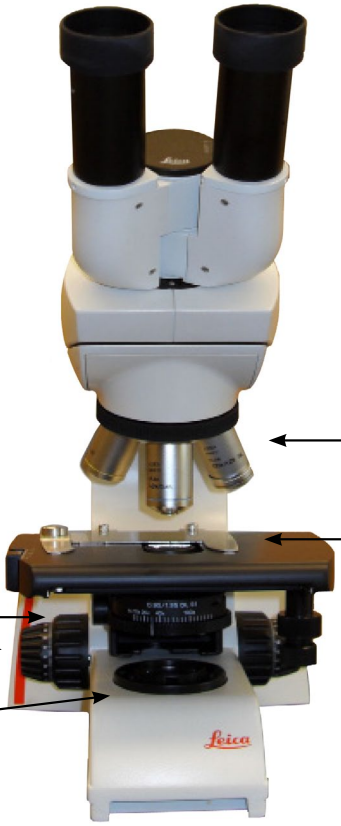
Microscopes are essential for viewing microorganisms. The first person to observe yeast and bacteria, was Anton van Leeuwenhoek, who called them *animalcules*. Yeast cells typically have diameters of ~10 μm, while bacteria have diameters of ~1 μm, both of which are far too small to be seen without considerable magnification. Light microscopes have a maximum resolution of ~0.2 μm, which is sufficient to resolve individual yeast cells and provide rough infomation about 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 present in yeast.

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

Chapter 2

Leica DM500 Light microscope

Eyepieces have 10X magnification
Interpupillary distance is adjustable



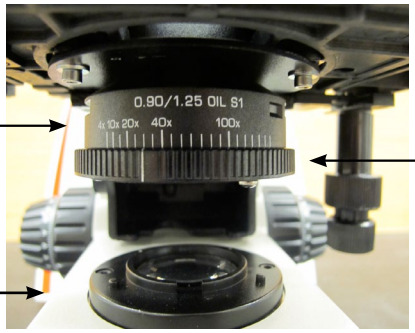
Parfocal **objective lenses**
4X, 10X, 40X and 100X

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)

that specimens remain reasonably well-focused when the lenses are changed. (**We will not be using the 100X lens, which requires immersion oil.**) When working with the microscope, ***always begin with the lowest power objective***, which is easiest to focus, and work your way to the higher power objectives.

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.

Students should work in groups of three.

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 2

Exercise 2 - observing yeast cultures with the microscope

Each student will prepare one slide. Students should observe the specimens on each other's slides.

1. Prepare three slides: In this experiment, you will prepare three slides, each of which contains two different samples for easy 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 space provided on the opposite page to record your data.

Slide 1: compare log phase cultures of *S. pombe* and *S. cerevisiae*.
Slide 2: compare log and stationary phase cultures of *S. cerevisiae*.
Slide 3: compare log and stationary phase cultures of *S. pombe*.

2. Prepare concentrated cell suspensions.

- Concentrate the cells in your log phase yeast cultures by centrifuging the culture tube 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.
- Use a transfer pipet to remove most of the culture medium, until the medium just covers the cell pellet.
- Resuspend the cells with the vortex mixer.
- Note: It is not necessary to centrifuge the stationary phase cultures, which are more concentrated than the log phase cultures.

3. Transfer and stain the cell samples

- Transfer a very small drop (the size of this o) of each cell suspension to the slide, using a disposable pipet.
- Stain the cells by adding a drop of Gram's Iodine to each cell suspension. The drop of iodine should be about **three times greater than the drop of cells**.
- Cover each sample with a coverslip.

4. Observe the cells and record your observations

- Use the same sequence of microscope adjustments that you used in the ruler exercise to visualize the cultures. Start at low magnification and gradually increase the magnification, making changes in the condenser diaphragm as needed. (Play with the position of the aperture diaphragm a bit to maximize the quality of the image.)
- In the space provided, draw some examples of the forms that you see in the cultures and the relative proportions of each form. Comment on both the sizes and shapes of the cells.

What differences did you observe between the two species?
What differences did you observe between stationary and log phase cells ?
Were there any differences in iodine staining between species or growth phases? What would that infer about glycogen storage?

Slide 1: Compare *S. cerevisiae* and *S. pombe* log phase cultures

Slide 2: Compare *S. cerevisiae* log and stationary phase cultures

Slide 3: Compare *S. pombe* log and stationary phase cultures

When you are finished with your observations, dispose of the slides in the glass waste.

Chapter 2

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.

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

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

Werner-Washburne, M, Braun, E, Johnston, GC & Singer, RA (1993) Stationary phase in the yeast *Saccharomyces cerevisiae*. *Microbiol Rev* **57:** 383-401.

Chapter 3

Mastering the micropipette

Welcome to the microworld! In this class, you are working with microorganisms, including 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

- Learn to select and adjust micropipettes
- Learn to accurately transfer microliter volumes
- Use the spectrophotometer to measure absorbance
- Understand experimental errors in measurements

Chapter 3

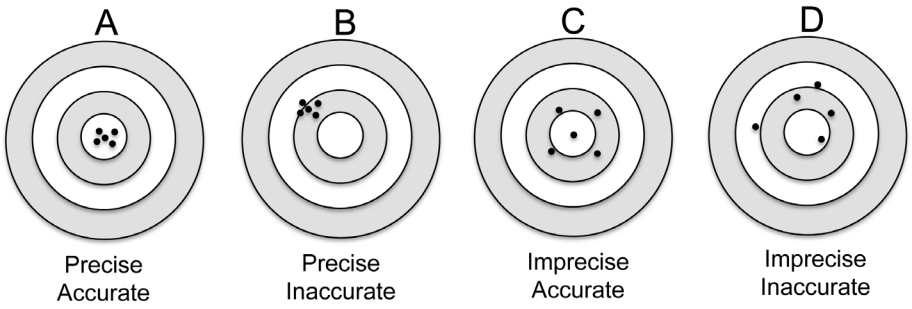
Using micropipettes correctly

Arguably, the most important scientific equipment that you will use this semester are adjustable micropipettes. Micropipettes are precision instruments that are designed to accurately 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 microliter (abbreviated μL) = 10^{-3} milliliter (mL) or 10^{-6} liter (L)
(A useful tip for Mac users: The keyboard shortcut for the Greek letter μ is Alt-m)

Accuracy and precision

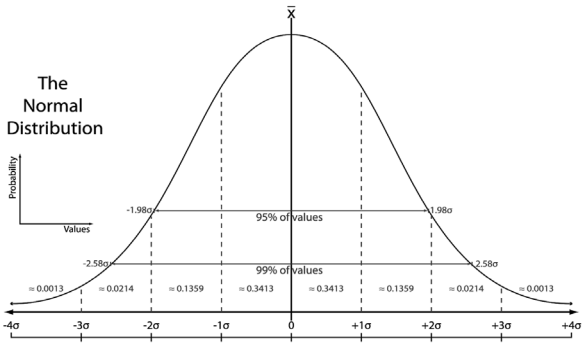
Ideally, micropipettes will deliver liquids with accuracy and precision. Accuracy depends on the micropipette delivering the correct volume. Precise results are reproducible. Let's use a target analogy to demonstrate the difference between accurate and precise results. Imagine that four students try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



The best way to determine the accuracy and precision of micropipettes is to use them to weigh set volumes of distilled water 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 selected (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. In general, micropipettes are designed to operate with accuracies within a few percent (generally <3%) of the intended value. The accuracy of a micropipette decreases somewhat, however, 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. Precision is expressed as the standard deviation (σ)of a set of measurements. Assuming random error, ~2/3 of measurements will fall within one standard deviation of the mean, and 95% of measurements will fall within two standard deviations of the mean.



Choosing the micropipette

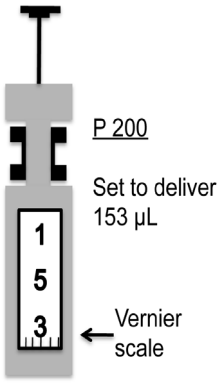
There are three different sizes of micropipettes in the laboratory, which we will refer to as the P20, P200 and P1000. Our micropipettes have been purchased from several different manufacturers, but the principles of operation are the same. The numbers after the "P" refer to the maximum number of microliters that the micropipette is designed to transfer. Use the chart below to select the correct micropipette for an operation. 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.

Micropipette	Recommended range (µL)	Smallest increment (µL)
P20	1 - 20	0.02
P200	20 - 200	0.2
P1000	100 - 1000	2.0

Micropipettes use disposable plastic tips. The P1000 tips are larger than those used with P200s and P20s. P1000 tips may be either natural/clear or blue, depending on the vendor, while P20 and P200 tips may be either yellow or natural/clear in color.

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!

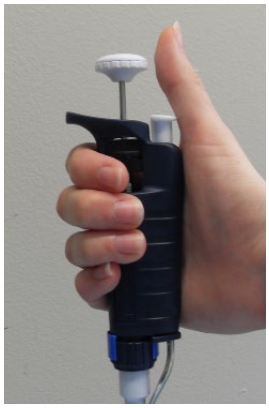


NEVER turn the indicator dial beyond the upper or lower volume limits of the micropipette! This could damage the piston.

Chapter 3

Transferring volumes accurately

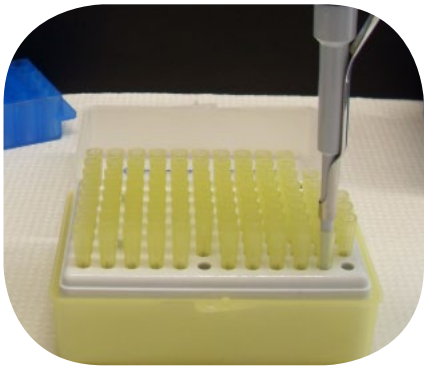
Micropipettes work by air displacement. The operator depresses a plunger that moves an internal piston to one of two different positions. The first stop is used to fill the micropipette tip, and the second stop is used to dispense the contents of the tip. As the operator depresses the plunger to the 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 micropipette tips.
- 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.

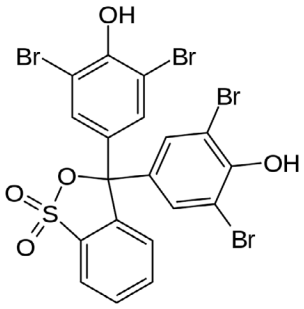
WARNING: The most common - and serious - operator error is depressing the plunger to the second stop before filling the micropipette tip.

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.



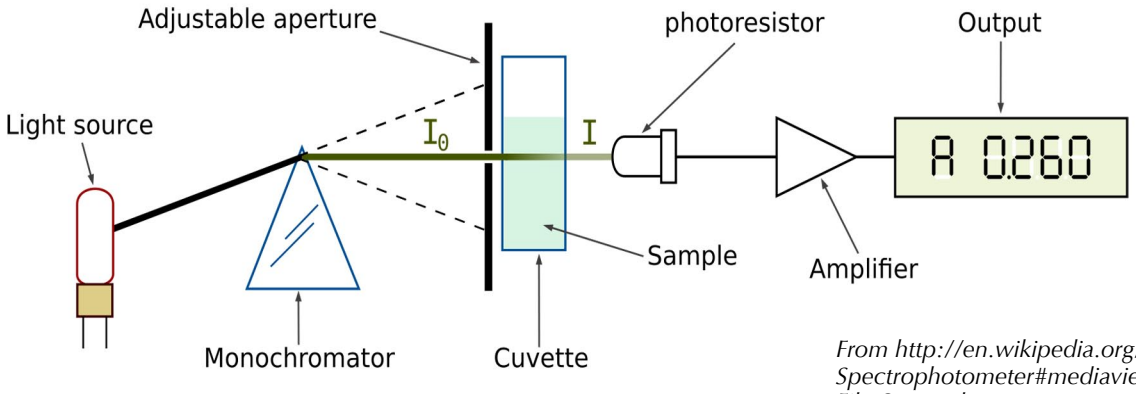
The spectrophotometer readings provide an indirect measurement of pipette performance. The proper way to calibrate the micropipettes would be to weigh out volumes of water, which has a specific gravity of 1.0 g/mL. Unfortunately, we do not have enough balances with sufficient accuracy for the class to 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 3

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:

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

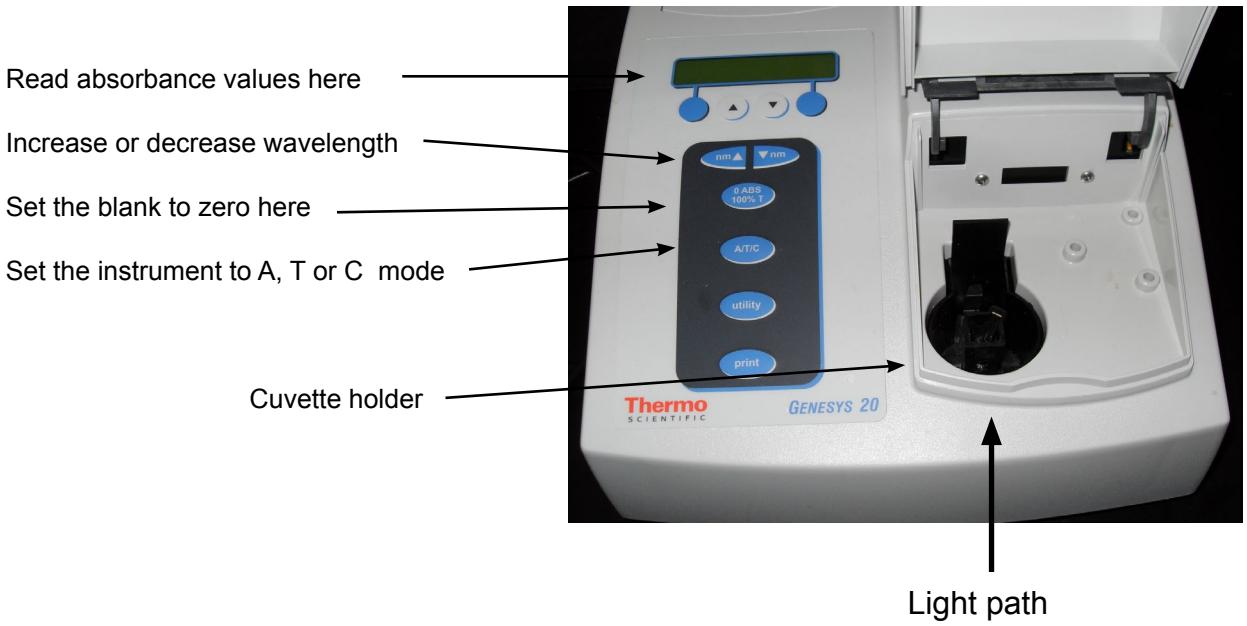
Spectrophotometers can express data as either % transmittance or absorbance. Most investigators prefer to collect absorbance values, because the absorbance of a compound is directly proportional to its concentration. Recall the Lambert-Beer Law, traditionally expressed as:

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

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. Set the mode to absorbance.
2. 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 cassette 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 3

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.

Lab 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.

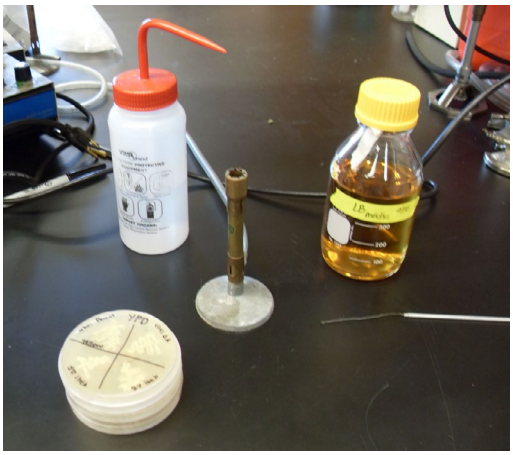
Applicant A: 0.161 g, 0.147 g, 0.142 g
Applicant B: 0.158 g, 0.156 g, 0.157 g
Applicant C: 0.143 g, 0.153 g, 0.150 g

Which applicant makes the most precise measurements?

Which applicant makes the most accurate measurements?

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 strains in the lab, as well as methods for culturing cells and estimating cell numbers.

Objectives

- Understand the phases of microbial growth
- Learn and practice sterile techniques used to culture yeast and other microorganisms
- Learn how to use spectrophotometry and spot plating to estimate the number of cells in yeast cultures

Chapter 4

Throughout the semester, you will be working with cultures of yeast. The techniques used for yeast and bacteria are similar, except that the media composition and culture temperature are optimized for individual organisms. In general, culture media contain a carbon source, nitrogen source, salts, vitamins and essential minerals. During the course of the semester, you will use a number of different strains of yeast and different kinds of media. Your success in the lab will depend on your ability to use sterile technique, which is needed to maintain the integrity of the different strains. An equally important element in laboratory success is careful bookkeeping!

In this lab, you will prepare stock cultures of *S. cerevisiae met* mutant strains on streak plates, quantify the number of viable cells in cultures of *S. cerevisiae* and S. *pombe* using spot plates, and *u*se light scattering to quickly estimate cell densities of these same yeast cultures

Sterile technique

Sterile technique is ESSENTIAL when working with microorganisms! This semester, we will be working with several different strains of yeast and bacteria. The strains have defined genotypes that have been generated by careful planning and experimentation. It is important to protect strains from contamination with other laboratory 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 of valuable laboratory strains.

Some simple precautions will reduce the possibility of contamination:

- Before working with strains, wipe down a small working area on the lab bench with 70% ethanol.
- 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.
- Minimize the time that the caps are removed from vessels containing microorganisms or sterile media. Caps should always be kept right-side up to prevent contamination from airborne microbes falling into the caps.

The culture media and reagents that we will use have been sterilized by either autoclaving or filtration. An autoclave (opposite page) is a chamber that uses pressurized steam

to kill cells on surfaces and in solutions, using temperatures near 121˚C and pressures from 30-40 psi. (For comparison, atmospheric pressure is ~15 psi.) Filtration is used in the place of autoclaving when solutions contain temperature-sensitive compounds. The pores in the filters used to remove microorganisms are typically 0.2 or 0.45 µm, which are sufficiently small to prevent the passage of bacteria. It is not difficult to keep stocks of media and reagents sterile as long as you work quickly and follow the directions above.

An autoclave.

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).

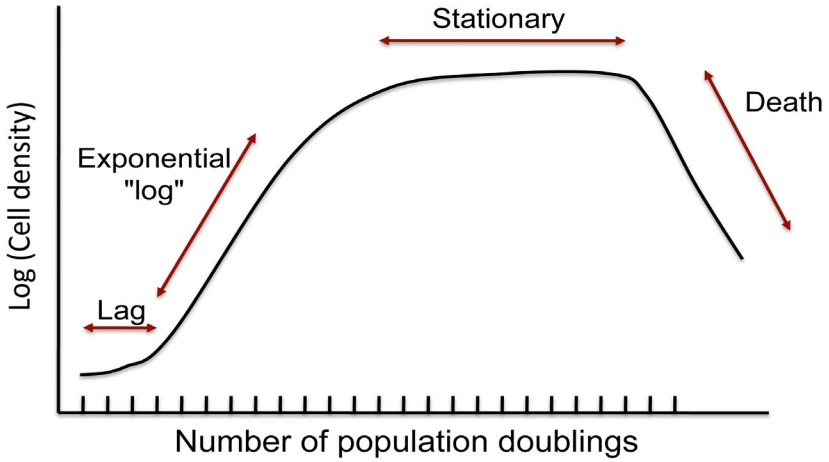
Because YPD is composed largely of crude extracts, its composition may show significant batch-to-batch variation. This variation is rarely a problem, however, because YPD contains more than enough essential nutrients to satisfy the metabolic requirements of cells. Many experiments, however, require media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

Yeast can be grown in liquid media or on the surface of plates containing solid media. Agar is usually added to liquid media, causing it to solidify. When cells are grown in liquid media, it is impossible to distinguish cells that have different genotypes from one another. By contrast, cells grow in colonies on solid media. Each cell in a colony is derived from a common ancestor and the cells are therefore genetically very similar, if not identical, to each other. For most of our experiments this semester, we will use solid media, because we need to distinguish cells with different genotypes from one another.

Chapter 4

Yeast growth phases

When yeast are grown in liquid medium, the culture follows a well-established pattern for microbial growth. 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), N_0 represents the number of cells at time zero. 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, which we will use for our experiments. At 30˚C, wild-type yeast have a doubling time of ~90 minutes in YPD. We will also make sure that the cultures are well-aerated by growing them on either rotary shakers or a rotating wheel.

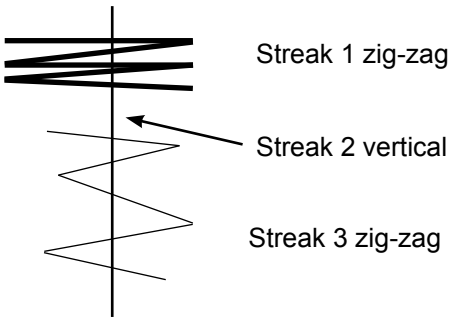
After a few doubling times, cells begin to deplete the nutrients in the culture, their growth rate slows, and the cells enter ***stationary phase.*** Cells in stationary phase have a different transcriptional program that allows them to survive in the less favorable environment, sometimes for considerable lengths of time. 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. 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've undergone 9-10 divisions. A colony, which has hundreds of millions of cells, is therefore a population of genetically very similar, but not necessarily 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** loop, wooden applicator stick, toothpick or pipette 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 freshly sterilized loop or toothpick 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 be used as stocks for future experiments. As you streak your cultures, 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.



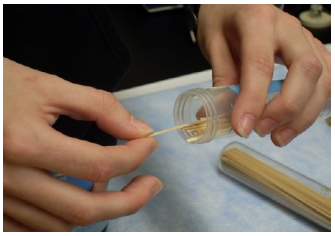
Chapter 4

Preparing a streak plate

1. Your team will be assigned three different *S. cerevisiae met* strains to culture. Gather the parent strains to be propagated, sterile toothpicks, and agar plate(s) with the appropriate fresh media.

2. Divide the plates with fresh media into sectors by marking the bottom of each 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. Carefully draw it out of the container. (Note: toothpicks and applicator sticks will have been sterilized in the autoclave, so 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.

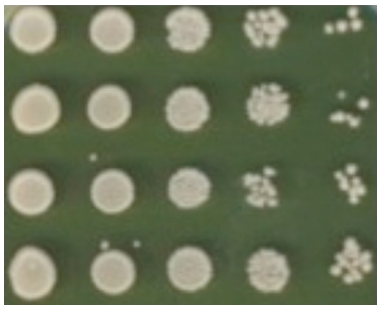
5. Transfer cells to the plate with fresh media. Make several zigzags across the outside edge of one sector with the toothpick or loop. *LIGHTLY* touch the agar surface as you move the toothpick. Think of pushing a hockey puck across an ice rink, rather than digging a ditch. Replace the lid and discard 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. Invert the plate and incubate it at 30˚C until individual colonies are visible, which is usually 24-48 hours.

Exercise 2 - Spot plates

You will use spot plates to estimate the cell densities of log phase and stationary phase cultures of *S. cerevisiae* and *S. pombe*. 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.



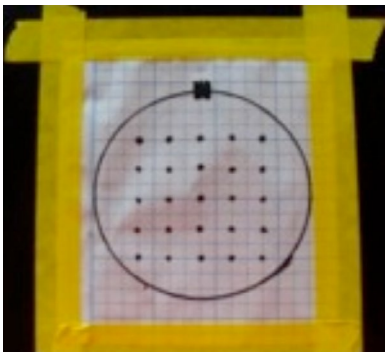
Spot plate.
Each row on the plate represents a series of 1:10 dilutions of a liquid culture of *S. cerevisiae*. Five µL of each dilution was spotted on the plate. The plate was incubated for two days at 30˚C. Individual colonies are apparent at the highest dilution of each extract.

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's 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 100,000-fold. The original culture would have contained 400,000 cells in 5 µL, which corresponds to 80 million cells per mL (8 x 10^7 cells/mL).

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

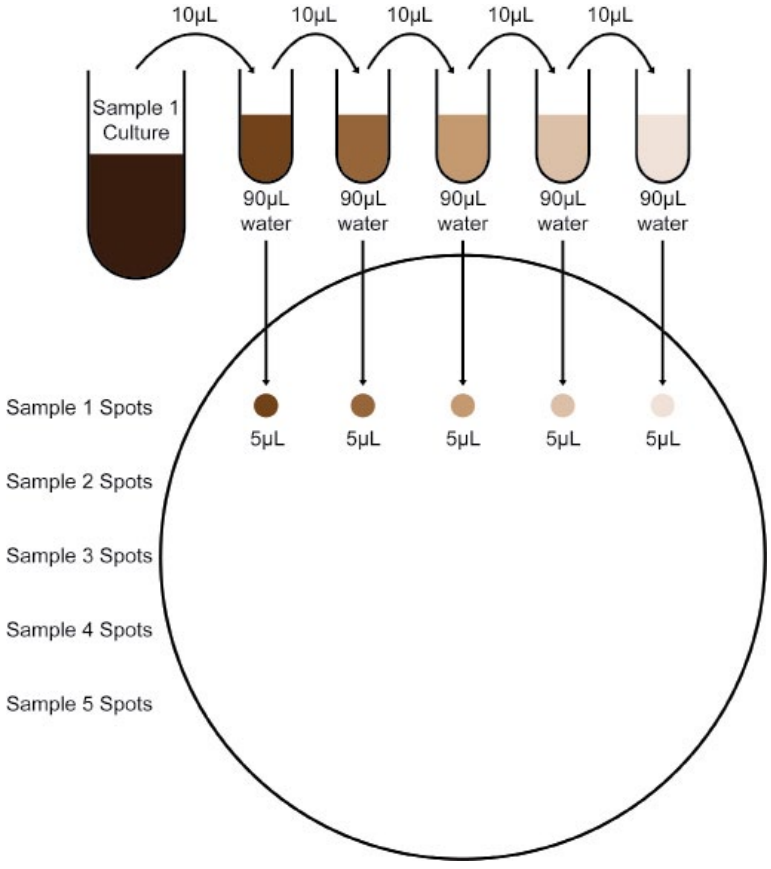


3. Prepare a series of five 1:10 dilutions from each culture using sterile distilled water. (Diagrams in your lab notebook are often helpful in designing dilution series.) To prepare a serial dilution, first pipette 90 µL sterile water into five microcentrifuge tubes. Next, use a P20 to transfer 10 µL 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 P20 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. 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 dilutions to settle and adsorb to the medium.



7. 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 and placed in the refrigerator or cold room for your analysis later.

9. Record your data with the scanner. *To do this, remove the top from the plate and invert both the plate and the 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 that you used and the volume of the spot.

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 to determine cell number is based on light scattering. 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.

Chapter 4

Estimating cell densities with the spectrophotometer
Follow the directions in Chapter 3 (p. 21) 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 and orient the cuvette in the holder so that the flat side of the cuvette faces the front of the instrument. (Note: it is not necessary to use culture medium as the blank.)

3. Close the lid and press the "0 Abs/100%T" button to establish a baseline value for further measurements.

4. Remove the cuvette from the instrument and replace the water with 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.

5. Repeat step 4 with each of your cultures.

6. 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.

References

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

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

- Learn how information is processed in databases
- Understand how the *Saccharomyces* Genome Project provided the reference sequences for *S. cerevisiae* genes
- Use the NCBI databases to find DNA and protein sequence information about a *MET*/*CYS* gene
- Use the *Saccharomyces* Genome Database to find information about the protein encoded by a *MET* gene and the protein's function in metabolism

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. The results 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 assigned a *MET* gene that you will be studying for the semester. In the next few labs, you will identify which of your three unknown yeast strains carries a deletion in that *MET* gene. To distinguish between the three yeast strains, you will need to obtain information about the DNA sequences that have been disrupted in the mutant strains. You will also need to find information about the roles that the encoded proteins normally play in metabolism. In this lab, you will search for gene-specific information in several online databases. As you progress through this lab, you may feel like you're going in circles at points, 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. Each database assigns its own unique accession numbers and organizes information into fields that will be useful to researchers who search the database. Once a record is accepted into a 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 that are designed to fit the needs of particular research communities. For example, scientists studying histones or protein kinases can access information in highly specialized databases.

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 exchanges data daily with smaller counterparts in Europe and Japan, providing multiple entry points into the international network of databases.

It's 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!

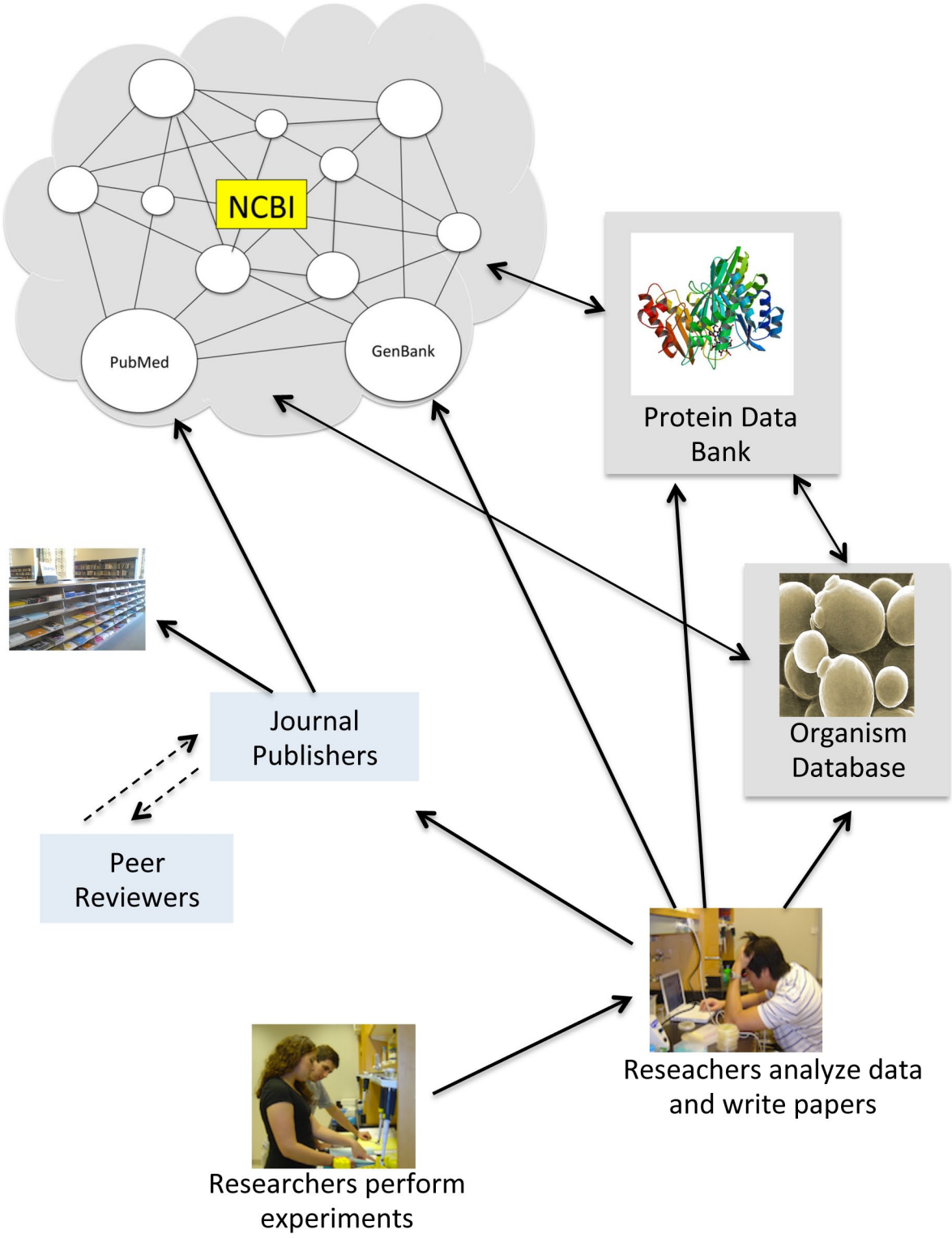
From the research bench to the database

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 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. In this course, we will be using both primary and derivative databases. Let's look at a few databases.

The information in databases originates in experiments. The figure on the following page summarizes information flow from the bench to databases. 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. Citations submitted to PubMed are assigned a PMID accession number. PMID numbers are assigned sequentially and the numbers have grown quite large. Today, PubMed currently contains over 21 million records! PubMed users can restrict their searches to fields such as title work, author, journal, publication year, reviews, and more. The usability of PubMed continues to grow. Users are able to use a clipboard, save their searches, and arrange for RSS feeds



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.

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 powerful "Find It" button to access the actual articles. ***Later in the semester, you will use PubMed to search for articles that will provide background information for your final semester report.***

Investigators submit experimental data to specialized research databases

Depending on the experiment, researchers may elect to submit their data to additional databases. Databases that accept direct submissions from researchers are considered primary 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, and 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 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. 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. The number of GenBank submissions has risen to over 135 records, including sequence information for whole genomes, individual genes, transcripts, plasmids, and more. Not surprisingly, there is considerable redundancy in GenBank records. GenBank is now part of NCBI's Nucleotide database, together with smaller, more specialized nucleotide databases prepared by curators. Among these databases is RefSeq, a source of nonredundant records prepared by NCBI curators. ***In this lab, you will search the Nucleotide database for the reference sequences for S. cerevisiae MET and CYS genes.***

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. 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 the novel properties of the protein encoded by the MET gene variant to the Saccharomyces Genome Database. The Saccharomyces Genome Database (SGD), which serves as a central resource for the S. cerevisiae research community (which now includes you). The SGD uses the yeast genome sequence as its organizing structure.

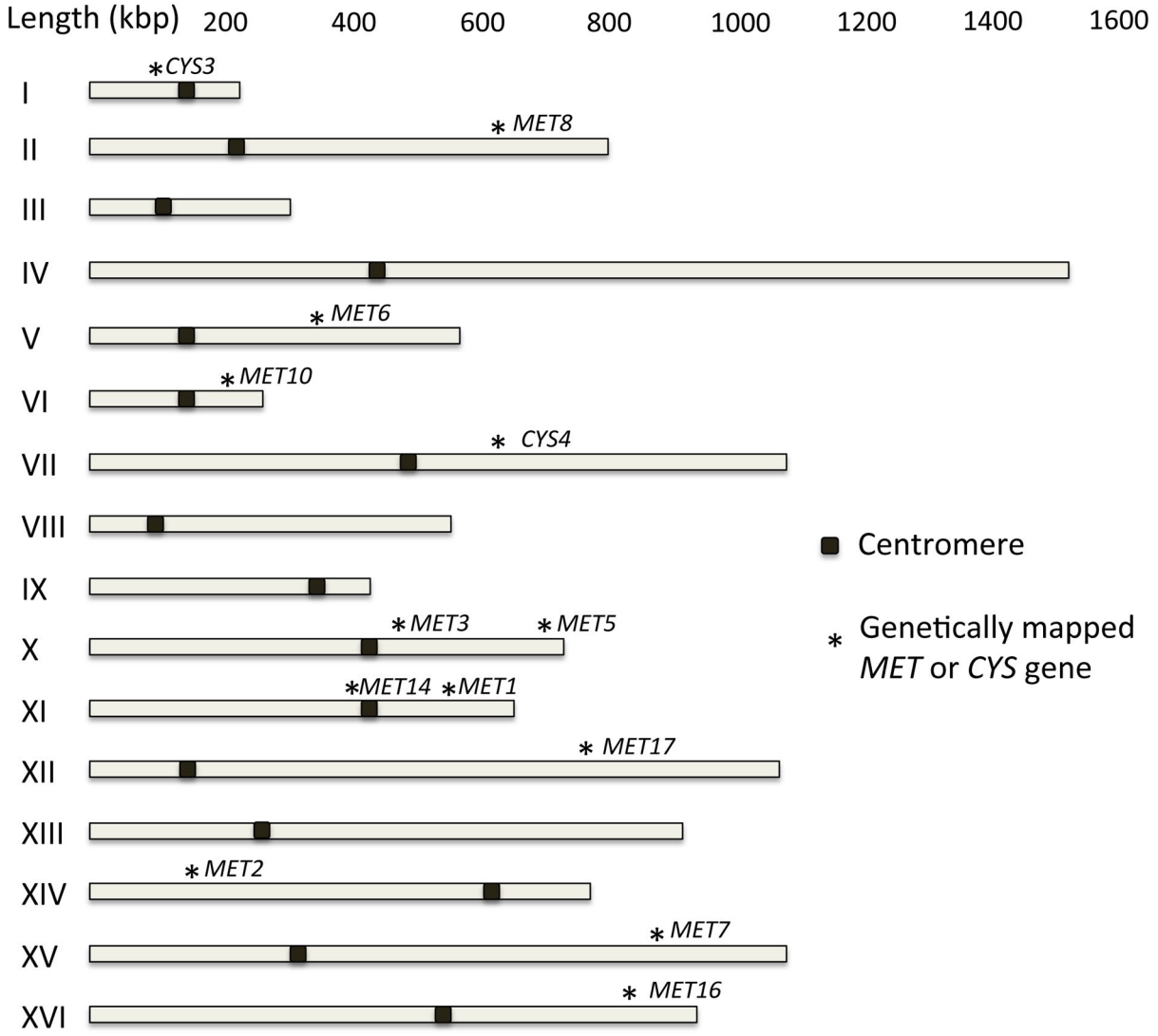
Saccharomyces genome project provided the reference sequence

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 their behavior 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 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 with approximately 1000 known genes. 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 sequence information was available for many chromosomal regions, including many *MET* genes. The detailed information collected by the yeast research community greatly facilitated interpretation of the yeast genome sequence, which was the first eukaryotic sequence to be decoded.

The yeast genome project was an impressive example of collaboration within the yeast research community. Over 600 researchers in 92 laboratories determined the complete DNA sequence of strain 288C with a high degree of accuracy (Goffeau *et al.*, 1996). A single strain was chosen for sequencing, because *S. cerevisiae* laboratory strains naturally accumulate mutations over time and can rapidly diverge from each other (Mortimer, 2000). The deletion strains that we are using in this class are derived from strain 288C.

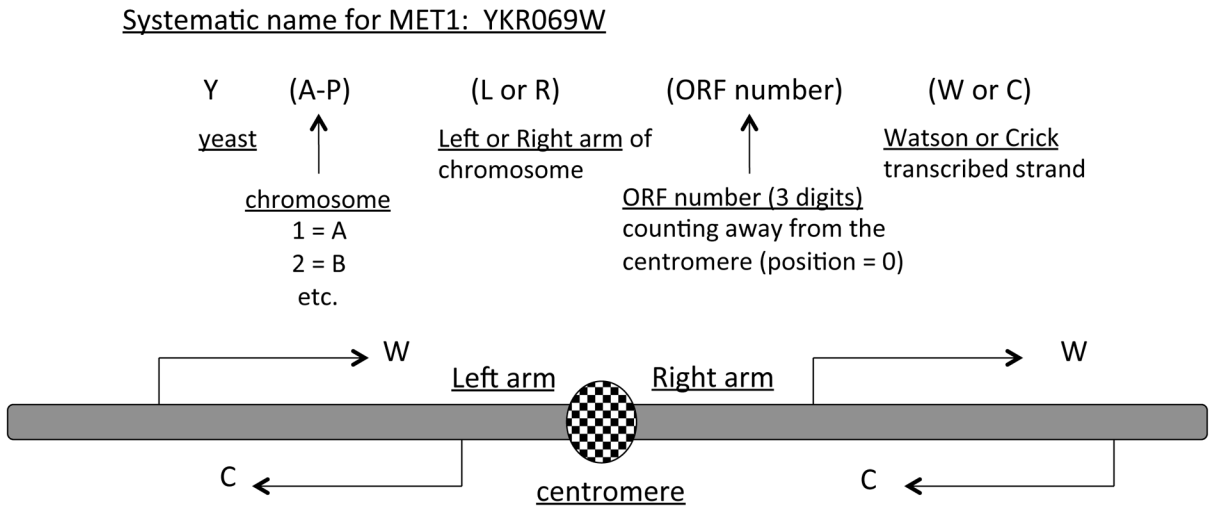
The ~12 million base pair (Mbp) yeast genome provides the definitive physical map of the 16 yeast chromosomes. The sequence generally confirmed the gene order predicted by the earlier genetic maps, but provided more accurate spacing for the distances separating individual yeast genes. The figure on the opposite page depicts the 16 yeast chromosomes, with the position of the centromere indicated on each chromosome. The figure also contains the positions of *MET* and *CYS* genes encoding enzymes involved in Met and Cys synthesis.



The genome project data provided the organizing structure for the *Saccharomyces* Genome Database, which systematically assigned accession numbers to ORFs based on their location and orientation on yeast chromosomes, as well as the direction of gene transcription. As shown at the top of the following page, 7-character systematic names begin 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.

Test yourself: The *S. cerevisiae* genome contains genes, *SAM1* and *SAM2*, 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 information below to determine their chromosomal locations, and place the two genes on the map above.

Chapter 5



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. NCBI curators assigned an NC_____ to each of the 16 chromosome sequences, indicating that the sequences are non-redundant chromosome sequences. The reference sequences were collected into the smaller RefSeq database, which is part of NCBI Nucleotide.

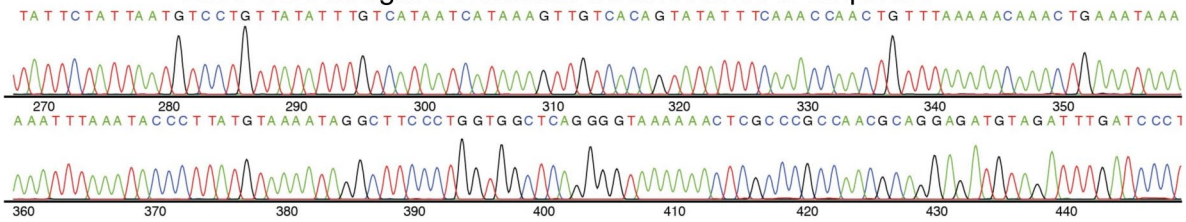
The figure on the opposite page outlines the process used to decode and annotate the *S. cerevisiae* sequence. During the annotation of the genome sequence, researchers used computational methods (see below) to identify approximately 6000 open reading frames (ORFs), or potential protein coding sequences, in the genome. ORFs were identified as 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. Each potential ORF identified in the project was assigned an NM_____ accession number, consistent with a transcript sequence, or potential mRNA. Note that the NM_____ transcripts are unlike real mRNAs, because they lack untranslated sequences at both their 5'- and 3'-ends.

Computational methods were also used to predict the amino acid sequences of the proteins encoding by transcriptsand the resulting NP_____ records were deposited in NCBI's Protein database. It is important to keep in mind that very few protein sequences in the Protein Database have been determined by chemical sequencing of a protein, which is a much more laborious task than DNA sequencing. Although the protein sequences are not experimentally validated, the transcript-derived NP_____ sequences are thought to be correct, since the sequences are frequently used in overexpression plasmids that produce functional proteins. Later in the semester, you will use plasmids that overexpress *MET* coding sequences.

Saccharomyces Genome Project

1. DNA Sequencing

Labs were assigned individual chromosomes to sequence



Thousands of sequencing reactions were performed

2. Alignment

Sequences were aligned to generate completechromosome sequences

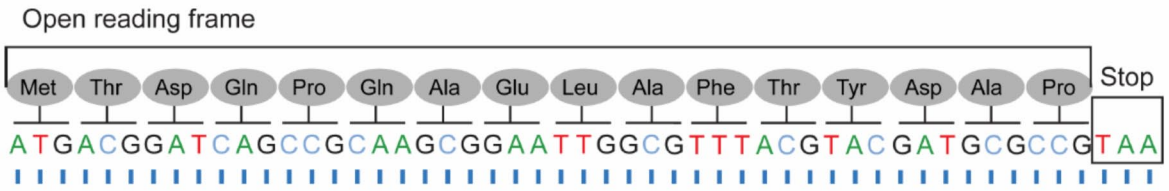
Computer algorithms identify areas of sequence overlap

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 is a large collection of databases. Clicking on the dropdown box brings up a list of individual 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. In your notebook, record the number of records in the PubMed, Nucleotide, Protein, and Structure categories.

- **Modify the search term by adding "AND Saccharomyces cerevisiae" to the search box.** Record the number of records in each of the categories used 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 to open the NC_ record. Near the top, you will see a link(s) to articles in the primary literature. Scroll down a bit in this very long record and look at a few genes. 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 CDS. You may see some genes where introns are predicted to occur. Intron-containing genes can be identified by the word "join" in the first line of the mRNA and CDS descriptions.

- Now let's take a look at the NM_ record for your gene. Use the back button on your browser to return to your search results. Open the NM_ record.

  Record the accession number _____

  How many nucleotides are in the coding sequence (bp)? _____

  Is there an intron in your gene? _____
  (Hint: Check the CDS to see if there are interruptions in nucleotide numbers)

- Find the predicted amino acid sequence near the bottom of the record. The NP____ record should be a few lines above the translation.

  What is the accession (NP) number for the protein sequence? _____

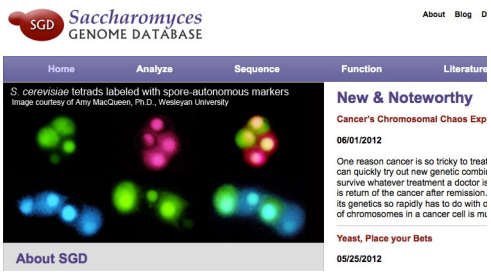
Think: is the NM_ record the actual sequence of the mRNA for your gene? (Do mRNAs all begin with AUG and end with a stop codon?)

Use the back button on your browser and click "All" to find all the *S. cerevisiae* records. Do any of the records look like they may be an mRNA sequence? Pay special attention to early GenBank submissions, which are near the bottom of the list. The earliest records were often submitted by investigators studying individual *MET* transcripts. If you do find an mRNA record, write its accession number in your notebook for future reference. How does its length (bp) compare to that of the NM_ sequence?

Chapter 5

Exercise 2: Using the Saccharomyces Genome Database

Direct your browser to yeastgenome.org



In addition to summarizing information about yeast genes, the SGD serves as a meeting place for the yeast community. Note the meeting announcements and summaries of newsworthy articles. The Community tab at the top of the page brings you to a variety of community resources, including a wiki.

- Type the name of your *MET* gene in the search box. This brings up the summary page for your gene.

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

  Record the systematic name for your gene _____

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

Now that we've looked at the structure of your gene, let's consider its function. The *MET* genes that we are studying all encode enzymes. Return to the SGD summary page for your gene and travel down to "Pathways" in the sidebar. You will see several links to the right, which bring you to the MetaCyc database's information for *S. cerevisiae* enzymes. MetaCyc, the online Encyclopedia of Metabolic Pathways, is a nonredundant database that catalogs pathways supported by experimental evidence. MetaCyc has extensive collections about reactions in humans, other model organisms and some more obscure organisms.

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

- Click on one of the pathway links. What is 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.

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's also important to see which genes show similar changes in expression. Microarray technology has allows investigators to study the simultaneous expression of hundreds of genes. Since the completion of the SGP, expression of *S. cerevisiae* genes has been studied in hundreds of high-throughput experiments.

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 summary. This brings you to a graph showing the number of experiments in log2 units (0 indicates no change, 1 a 2-fold change, 2 a 4-fold change, etc.) where the level of gene expression changed. Were both decreases and increases observed? How large were most of the changes observed in experiments? The *MET* genes encode enzymes that catalyze discrete steps in the synthesis of methionine. Consider:

  Would large changes in the concentration of a Met protein be required to alter the concentration of methionine in a cell? Would a change in the concentration of a single catalyst in the methionine biosynthetic pathway be expected to change methionine concentrations? Would changes in the concentrations of multiple enzymes be required to change methionine concentrations?

- Click on the SPELL link. This will bring you to a summary of many high throughput gene expression studies under a variety of environmental conditions. SPELL gives a list as well as graphical output of the genes that are most similar to your gene in their expression patterns. Green labels indicate increased expression, and red labels indicate reduced expression. Scroll right to see the results of experiments that don't fit on the screen. Note that SPELL gives you the literature citations for the experiments that produced the experimental data.

- Record the list of the six genes that were most frequently co-expressed with your gene. Look up the reactions catalyzed by their gene products. What do you notice about genes that are coordinately expressed with your *MET* gene?

Protein information

- Click on the protein tab at the top of the gene summary page. This page gives you a lot of useful information about your protein, which we'll use later in the course.

- Investigators in one study used a microscopic method to estimate how many molecules of each protein were present in a log phase yeast cell (Ghaemmaghami *et al.*, 2003). Numbers varied from less than 50 to a~1 million molecules/cell.

  How many molecules of your enzyme are present in a log phase cell?

  Is your gene highly expressed? moderately expressed? Are you surprised?

- Note the physical characteristics of your enzyme.
  How many amino acids are encoded in its sequence?
  Is your enzyme acidic or basic? (Check the pI).

- Click on the PDB Homologs under Homologs to access structural information on your enzyme.

  Has the *S. cerevisiae* enzyme been crystallized? What is its PDB accession number?

  Has a homolog from another species been crystallized? If so, note the species and the 4-character PDB accession number(s).

  Structural homologs do not always perform the same enzymatic reaction. Check the PDB number to make sure that the E.C. number is the same as that of your Met protein.

Chapter 5

References

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

Ghaemmaghami, S, Huh, WK, Bower, K *et al.* (2003) Global analysis of protein expression in yeast. *Nature* **425:** 737-741.

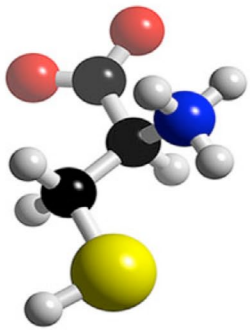
Goffeau, A, Barrell, BG, Bussey, H *et al.* (1996) Life with 6000 genes. *Science* **274:** 563-567.

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

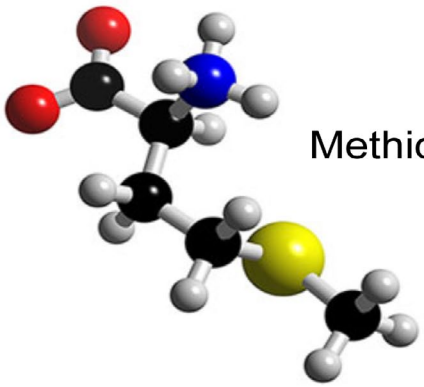
Winzeler, EA, Shoemaker, DD, Astromoff, A *et al.* (1999) Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. *Science* **285:** 901-906.

Chapter 6

Analysis of mutant strains



Cysteine



The growth properties of mutant strains can often provide information about the gene products involved in biochemical pathways within cells. In this experiment, you will use selective and differential media to identify which *MET* or *CYS* genes have been inactivated in yeast deletion strains.

Objectives

- Become familiar with the biochemical pathways involved in methionine and cysteine synthesis
- Understand how selective and differential media are used in genetics
- Design a strategy to distinguish *met* and *cys* mutants by their use of sulfur sources
- Spot plate yeast *met* and *cys* strains on selective media containing different sources of organic sulfur

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 imparts 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. Yeast *met* and *cys* mutants, however, are unable to synthesize Met and Cys from sulfate. Depending on the exact mutation, they require alternative sulfur sources that they are able to 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* or *cys* mutants. The mutant strains that we are using were developed as part of the yeast genome project. In each of our mutant strains, a *MET* or *CYS* gene has been replaced with a bacterial kanamycin resistance (KAN^R) gene by homologous recombination (Winzeler *et al.*, 1999). In the next lab, you will use the polymerase chain reaction (PCR) to more conclusively identify the mutant strains.

Genetic nomenclature

When referring to strains, it is important to use the correct genetic nomenclature. We will be looking for correct usage in your reports! Pay close attention to italics and capital letters. 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 almost all other 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*. *MET6* is one of over twenty different *MET* genes. The *MET* prefix is used because most of the *MET* genes were isolated in genetic screens based on the inability of mutant strains to live in the absence of methionine. 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 mutation 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 on p. 54, except methionine.

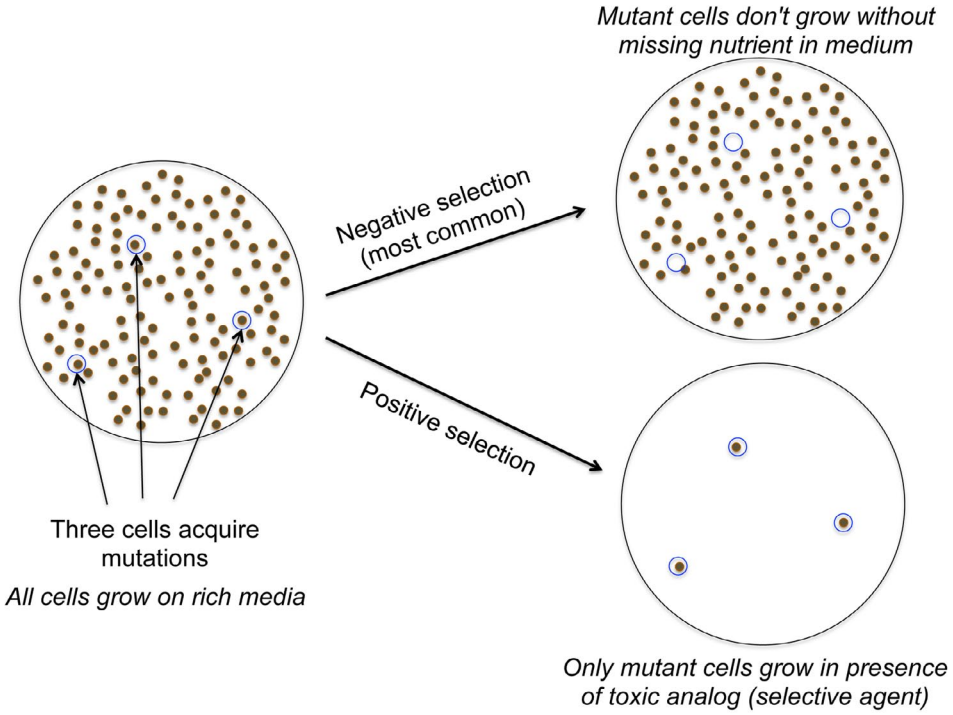
Composition of Yeast Complete (YC) Medium					
Component	**grams/liter**	**Component**	**mg/liter**	**Component**	**mg/liter**
YNB*	1.7	arginine	100	tyrosine	50
$(NH_4)_2SO_4$	5	aspartic acid	50	lysine	100
D-glucose	20	isoleucine	50	methionine	50
		phenylalanine	50	tryptophan	100
		proline	50	leucine	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 in yeast on page 57, you may wonder how the gene numbers became associated with the various 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, ~10^{-8}/base/generation, is much too low for a practical genetic screen, so investigators 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,



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.

they would probably use a negative selection scheme. In a negative scheme, cells are first plated 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.

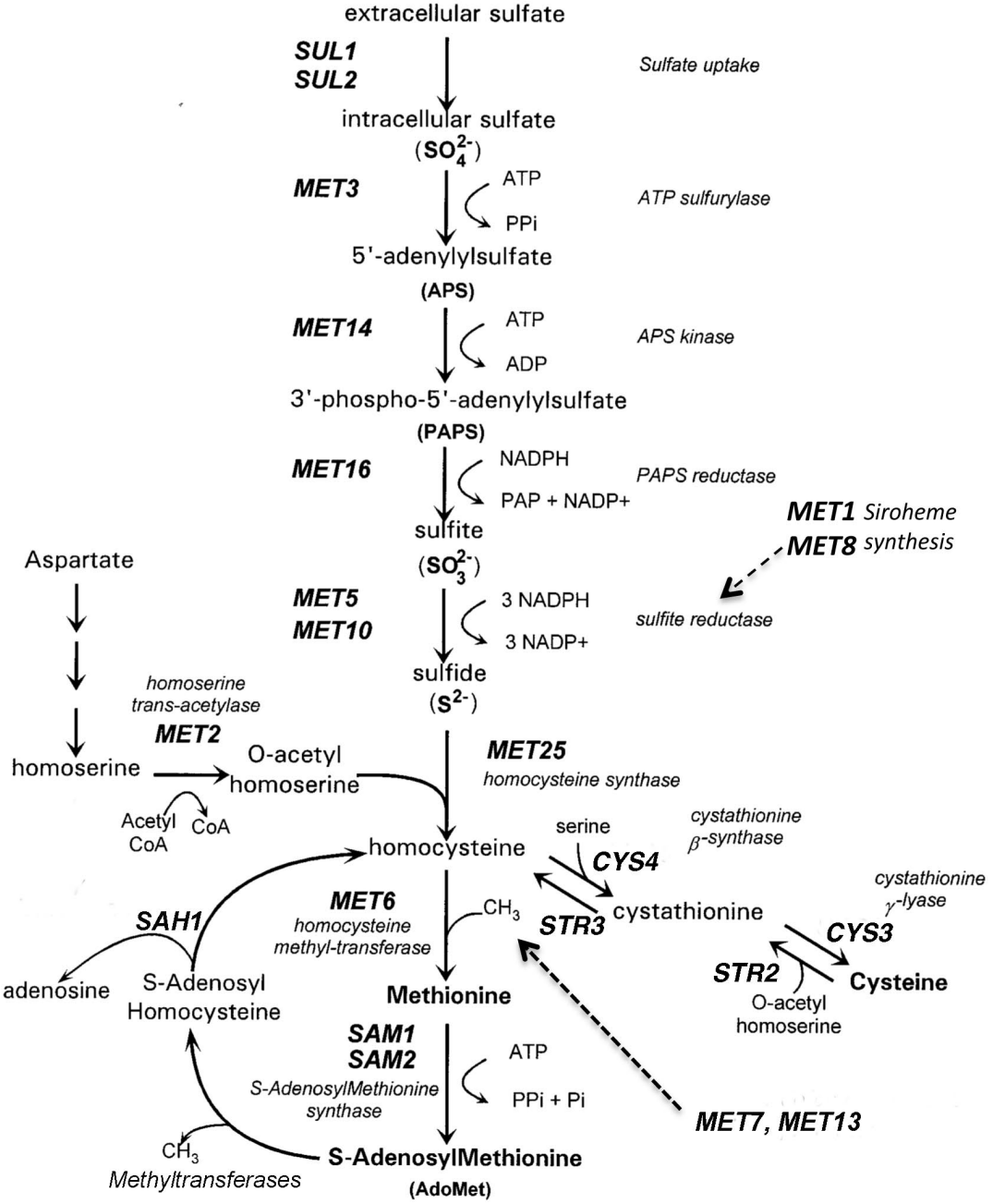
The number and spectrum of mutants obtained in a genetic screen are unpredictable, because of the random nature of mutation. As you might expect, a screen might produce multiple mutants in one gene and no mutations in other genes involved in the phenotype. After completing a screen, investigators must next determine if the mutations are in the same or different genes. For this, geneticists rely on genetic mapping (Chapter 5, p. 40) and 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 various *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* and *cys* mutants. You will also use a differential medium, BiGGY agar, that distinguishes yeast strains by their production of hydrogen sulfide.

NOTE: The *met* and *cys* 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 (Winzeler *et al.*, 1999). 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



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)

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 are less likely to harbor secondary mutations from mutagen treatment and spontaneous reversion to a wild type phenotype is not possible.

Biochemistry of methionine and cysteine synthesis

The consensus view of yeast Met and Cys synthesis 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* and *CYS* genes, which were transformed into yeast *met* and *cys* mutants. Plasmids could only restore Met and Cys synthesis when the plasmid contained the wild type allele of the inactivated gene in the mutant. (You will use plasmid complementation in this class as functional proof that you have correctly identified your strains and plasmids.)

The previous page shows the superpathway for the biosynthesis of sulfur amino acids. In the database lab, you found that each of the *MET* genes is classified as belonging to smaller biochemical pathways, as well as this superpathway. (The only genes that are not included in the superpathway are *MET1* and *MET8*, which belong to the siroheme biosynthesis pathway. Nonetheless, these genes are essential for Met and Cys synthesis.) Some familiarity with the pathways on the previous page will be useful in this lab, where you will distinguish three strains by their utilization of sulfur compounds. In the brief description below, we will follow the progress of the sulfur atom from its transport into the yeast cell to its conversion to Met, Cys and the high-energy methyl donor, S-adenosylmethionine (AdoMet).

Sulfate assimilation involves sulfur activation and reduction to sulfide

The initial portion of the super-pathway, which encompasses the reactions involved in the conversion of sulfate to sulfide, comprises 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 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,

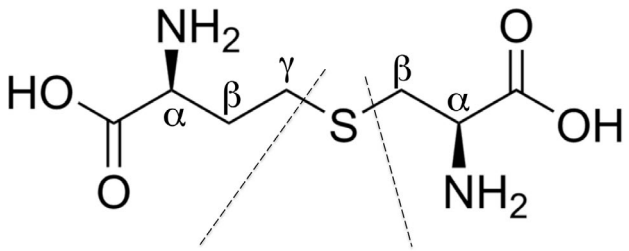
proteins and a variety of small molecules. (Interestingly, APS kinase is the only yeast enzyme involved in sulfate assimilation with homologs in mammals.)

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. The four genes 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_γ bond. Str3p and Str4p (Cys4p) catalyze the synthesis and hydrolysis, respectively, of the cystathionine S-C_β bond.

Methionine and AdoMet are formed during the methyl cycle

Hcy is also the starting point of a cycle that produces Met and S-adenosylmethionine (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.) The substrates for AdoMet-dependent transmethylation reactions are quite diverse and number in the hundreds: nucleotide bases and sugars in DNA and protein, 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* and *cys* mutants that you are analyzing are unable to catalyze one of the reactions shown on p. 57. 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 superpathway. The selective media contain a variety of sulfur sources. Find the position of the sulfur source in the pathway and determine if it is possible to synthesize Met and Cys from that sulfur source if a particular gene-of-interest is mutated.

You will also be using the differential medium, BiGGY agar to distinguish yeast strains by the quantity of hydrogen sulfide that they produce. All strains are expected to grow on BiGGY, since it contains glucose and yeast extract. BiGGY also contains bismuth, which reacts with sulfide to form a brownish to black precipitate. Locate the positions of your mutated genes in the superpathway relative to sulfide. Mutations in genes that lead to sulfide should produce lighter colonies, since less sulfide will be produced. Mutations in genes that lead away from sulfide should produce darker colonies, because the strains will be unable to metabolize sulfide.

Exercise 1 - Predicting growth properties of mutant strains

The *met* and *cys* mutants that we are using for our experiments were derived from the parent strain BY4742, which has the genotype *MATa his3-Δ1 leu2Δ0 lys2Δ0 ura3Δ0*. Each strain contains a additional mutation in a *MET* or *CYS* gene. The defined media that we are using for this experiment is based on YC (p. 54), which contains histidine, leucine, lysine and uracil as supplements. In addition to complete YC medium, which contains methionine, we will use YC-Met dropout media, which lacks methionine. Alternative sulfur sources have been used in the place of Met in some plates. Predict the ability of *met* and *cys* 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.

You will also be using BiGGY agar plates to distinguish mutant strains. Use upward- and downward-facing arrows to predict colonies that are darker or lighter, respectively, than the control BY4742 strain.

	YPD	**YC Complete**	**YC - Met**	**YC-Met +Cys**	**YC-Met +SO_3**	**BiGGY**
met1	+					
met2	+					
met3						
met5						
met6						
met7						
met8						
met10						
met13						
met14						
met16						
met25						
cys3						
cys4						
str2						
str3						
sam1						

Chapter 6

Exercise 2 - Identifying strains by nutritional requirements

Doug - How do you want to do this? We could have students prepare inoculate liquid cultures from their own streak plates in lab 4 (backups available, of course). The text below is from last year. Send me revisions!

Your team will be given three strains, each of which carries a different *met* or *cys* mutation. Prepare spot plates to distinguish the three strains that your team has been given, following the procedure in Chapter 4.

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 series on one plate before proceding to the second plate.

2. Use the same pattern of strains/rows on each of the different selective plates. ***Make sure that the plates are properly labeled so that you can identify which strain corresponds to which row!***

3. Incubate the plates at 30 °C for at least 3 days, when colonies should become apparent. Note the size of the colonies. Rapidly growing strains produce larger colonies than slowly growing strains. Record any observations in your notebook. Note that it may take longer than 3 days for some of the strains to produce single colonies.

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.

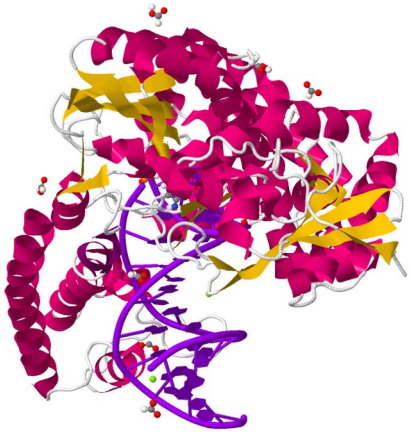
Masselot, M & DeRobichon-Szulmajster H (1975) Methionine biosynthesis in *Saccharomyces cerevisiae*. I. Genetical analysis of auxotrophic mutants. *Mol Gen Genet* **139:** 121-132.

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 as part of the *Saccharomyces* Gene Deletion Project. In the project, yeast investigators systematically replaced each ORF in the yeast genome with an ORF-specific cassette containing a bacterial kanamycin resistance (KAN^R) gene. In this lab, you'll design and carry out PCR to identify which *MET* or *CYS* genes have been replaced by KAN^R genes in your yeast strains. Thermostable DNA polymerases (above) play a key role in PCR.

Objectives

- Understand the reactions occuring at the different temperatures used in PCR
- Understand the principles of primer design and hybrization
- Understand how to optimize PCR conditions for specific products
- Use PCR to distinguish between yeast colonies of different genotypes

Chapter 7

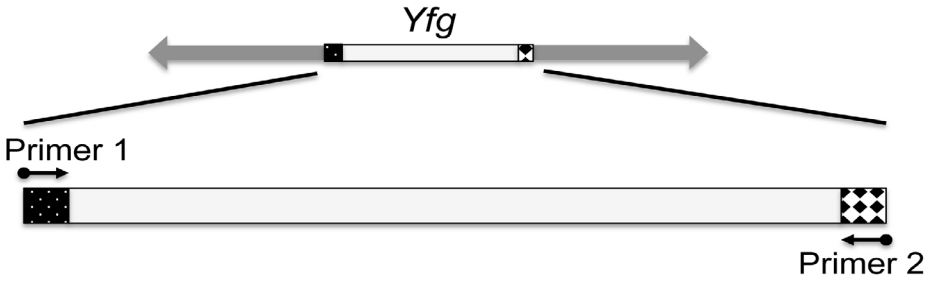
In a previous lab, you used different culture media to distinguish between different *S. cerevisiae met* and *cys* mutants. Your results may have allowed you to tentatively identify your strains. In this lab, you will use the polymerase chain reaction (PCR) to more conclusively identify the mutant strains.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

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 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**

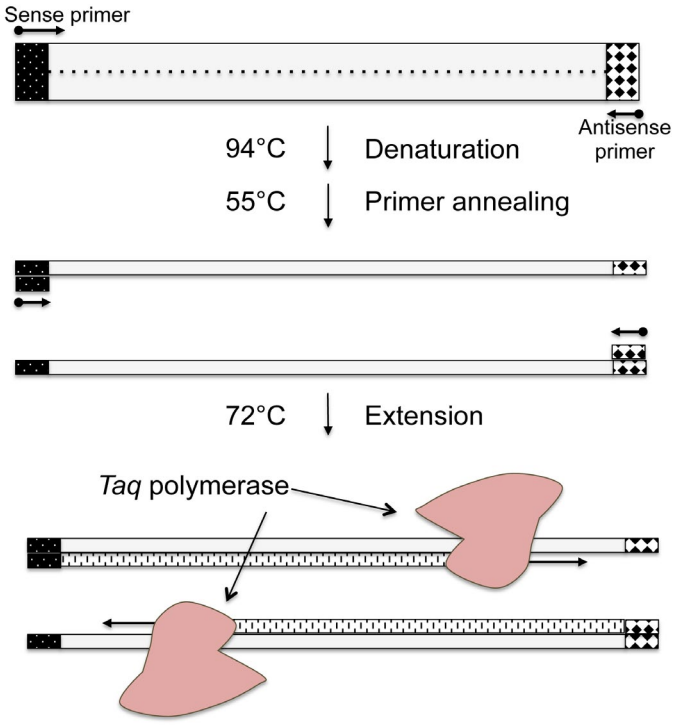


need accurate sequence information for the primer binding sites in the target DNA. (Note: Sequence information is not needed for the entire sequence that will be amplified. PCR is often used to identify sequences that occur between two known primer binding sites.) Two primers are required for PCR. One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

PCR typically 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 of DNA is critical, because DNA polymerase requires single-stranded DNA for as a template. The first cycle of denaturation is longer than subsequent denaturation steps, because PCR templates are often long, complex molecules held together by many hydrogen bonds. In subsequent cycles, shorter templates will predominate (see below).

Following the initial denaturation, PCR involves a series of 30-35 cycles with three segments, as outlined and discussed in greater detail below.

- A denaturation step - commonly 94˚C
- A primer annealing step - commonly 55˚C
- An extension step - commonly 72˚C



PCR reactions include multiple cycles at three different temperatures

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

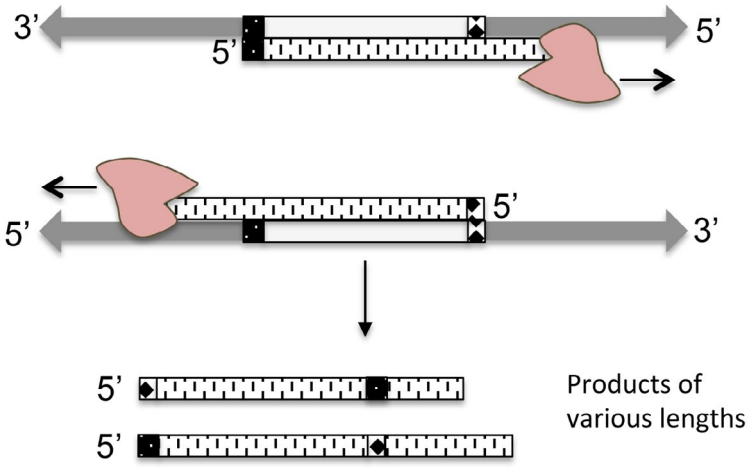
Chapter 7

In the early days of PCR, investigators (or their students!) manually transferred tubes between three water baths programmed for one of the three temperatures used in the procedure. Today, PCR is performed in automatic thermocyclers that rapidly adjust temperatures in a block containing the reactions.

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.

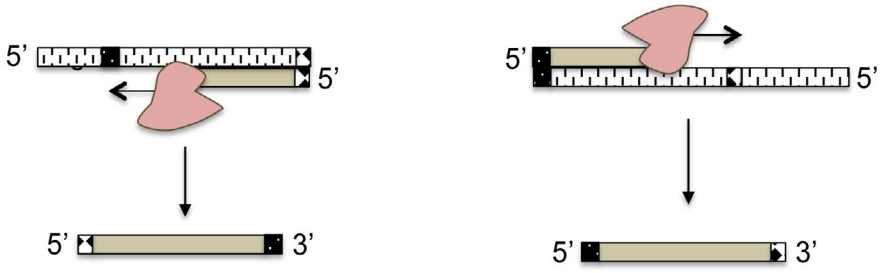
During the first cycle, the thermostable DNA polymerases synthesize DNA, extending the 3' ends of the primers. DNA polymerases are processive enzymes that will continue to synthesize DNA until they literally fall off the DNA. Consequently, the complementary DNA molecules synthesized in the first cycle 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



First cycle of PCR
During the first PCR cycle, 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.

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.

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 fragments of the intended length first appear in the second cycle.
The "anchored" fragments generated during the first cycle of replication begin with either the primer 1 or primer 2 sequence. During the second cycle, replication begins at the other primer site, generating a PCR fragment 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 a PCR reaction. 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 partially complementary, causing DNA polymerase to copy the wrong DNA sequences. Fortunately, investigators are able to adjust experimental parameters to maximize the probability that primers will hybridize with the correct targets on the template DNA.

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. 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. 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 temperature.

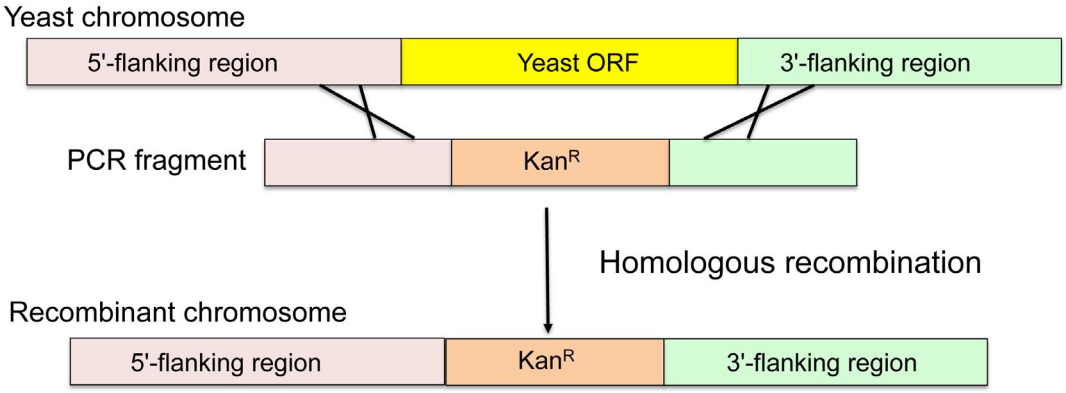
When designing a PCR reaction, investigators also consider the nature of the template DNA. A variety of DNA templates can be used for PCR. Depending on the purpose of the experiment, investigators could choose to use genomic DNA, a plasmid or a cDNA (complementary DNA generated by a reverse transcriptase 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 20-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 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 take advantage of the high frequency with which yeast exchange genes using 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, the DNA strands break and strand exchange occurs when the breaks are healed. Investigators who want to use homologous recombination for gene replacement use molecular cloning to construct a replacement cassette in which a marker gene is flanked on either side by sequences that flank the gene to be replaced in the target chromosome.

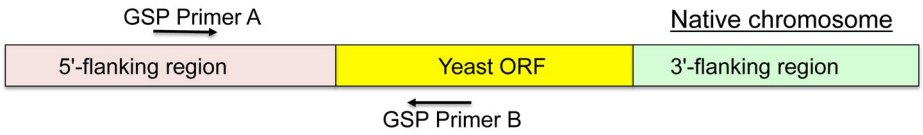
The strains that we are using were generated by this approach during the *Saccharomyces* Genome Deletion Project (Winzeler *et al.*, 1999). After the publication of the DNA sequence, yeast researchers began to systematically replace all of the predicted ORFs in the *S. cerevisiae* genome with a kanamycin resistance (*KAN*^{R}) gene. The figure below outlines the process used to generate the deletion strains. For each ORF, researchers used a series of PCR reactions to construct cassettes in which the *KAN*^{R} gene was flanked at either site by short DNA sequences upstream and downstream of the targeted ORF on the *S. cerevisiae* chromosome. The PCR fragments were then used to transform the BY4742 strain (Brachmann *et al.*, 1998) of *S. cerevisiae*. Strains that had incorporated the *KAN*^{R} gene were selected on plates containing analogs of kanamycin.



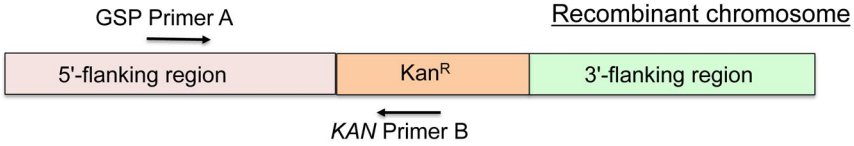
All of the deletion strains were analyzed by PCR to confirm that the targeted ORF was replaced by the *KAN*^{R} cassette. Thus, we can be reasonably certain about the genotypes of our *met* mutants. The Deletion Project also verified that strains were able to grow and mate, showed the appropriate auxotrophies and had functional mitochondria. (See the Deletion Project website for additional details: http://www-sequence.stanford.edu/group/yeast_deletion_project/deletions3.html.) With the large number of strains generated in the project, it was not possible to do a thorough phenotypic analysis of individual strains. Those analyses are left for you to do!

Chapter 7

In this laboratory, we will use the primers designed by the *Saccharomyces* Deletion Project (SGDP) to analyze the *MET* gene disruptions in your yeast strains. As shown in the figure below, we have obtained two gene-specific primers (GSP) for each of the *MET* genes that we are working with. 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). The sequences of the primers are shown in the table at the bottom of the page.



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

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. Courier is a nonproportional font that is often used for nucleic acids, because it assigns all letters the same width. At a glance, it is possible to see that most of the primers are the same length.

Exercise 1 - Physical properties of PCR primers

The table on the opposite page 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 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), so investigators were not always able to design primers with the ideal 40-60% GC-content. Your group will be working with three deletion strains. Complete the table below by adding the length, GC-content and T_m for the primers designed for to analyze your deletion strains. 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 calculate the sizes of the PCR products generated in the reaction, you will next align the primer sequences for each gene against the genome sequence and prepare a simple map of the genomic region.

The PCR pieces that will be generated in the reaction will contain part of the *MET* or *CYS* coding sequence, as well as sequence from the 5'-flanking region. You will recall from Chapter 5 that NC_____ records contain the complete sequence of the chromosome where your gene is located. The NC____ record could be used in the alignment, but we can take advantage of other sequence resources that SGD curators have prepared. For this exercise, we will use records that contain the CDS for your gene as well as 1 kb of upstream sequence and 1 kb of downstream sequence. 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. Because the *S. cerevisiae* genome is compact (Goffeau *et al.* 1996), these elements are usually located within 1 kb from either end of the CDS.

Chapter 7

You will need two browser windows for this exercise. Each member of the group may want to 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 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.

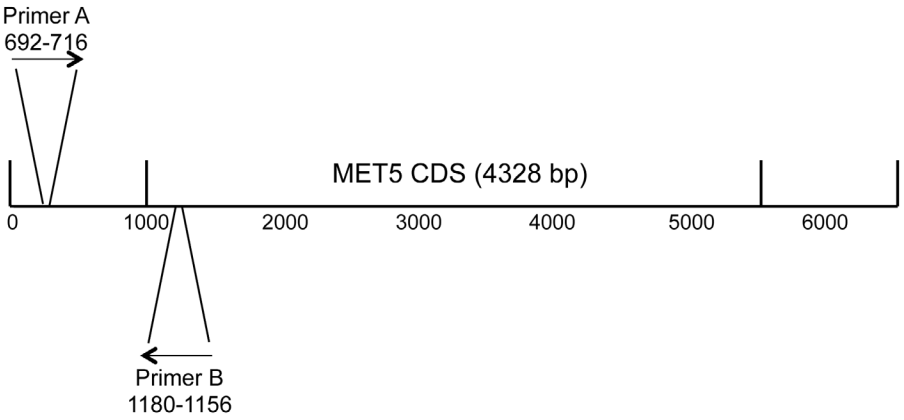
To find the position on the primers in the genomic sequence, we will use NCBI's BLAST tool. BLAST stands for Basic Local Alignment Search Tool and can be used to align either protein or nucleic acid sequences. You will learn more about the BLAST algorithms in Chapter 9.

- 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. Click BLAST. The BLAST results bring up a table with 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!) Do you see any other matches? Record the starting and ending points where the primer matches the genomic 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 your other primer. Click BLAST and record the alignment results. Primer B is the reverse complement the genomic sequence.

Use the information that you collected from the BLAST searches to prepare a map of the genomic region roughly to scale. Include the primer binding sites and the coding region on the genomic DNA. An sample map is shown below for the *MET5* gene, which has a CDS sequence of 4398 bp. Primer A binds to nucleotides 692 to 716. Primer B binds to nucleotides 1180 to 1156. You can see how straightforward it is to calculate the size of your PCR product from this kind of map. The PCR product produced with the *MET5* primers A and B will be 488 bp, of which 180 bp represent CDS, and 308 bp represent the 5'-flanking region.

Disrupted genes containing the *KAN*^{R} cassette will give PCR products when the *KAN*^{R} Primer B is used with Primer A. The *KAN*^{R} Primer B binding site is located 250 bp into the *KAN*^{R} gene. The PCR product generated with *MET5* primer A and the *KAN*^{R} primer B would be 558 bp, including 308 bp from the 5'-flanking region and the *KAN*^{R} sequence.



Draw a map for your gene in the space below.

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

Primer A and Primer B

Primer A and KAN^R primer B

Chapter 7

Exercise 3 - Design yeast colony PCR

In this lab, each team will be able to perform **six** PCR reactions to identify your three yeast strains. You will have the option of using any of the primers in the Table on p. 58. 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 to devise a strategy that will allow you to positively identify your met yeast strains.

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

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.

- ***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.)

- ***Prepare the primer mixture.*** 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 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 primers that don't belong in a reaction. Change tips between every primer transfer that you do.

- ***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 appropriate amount of yeast would fit on the tip of a pin.

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

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

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

- ***Store the reactions at 4°C for analysis in the next lab.***

Chapter 7

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.

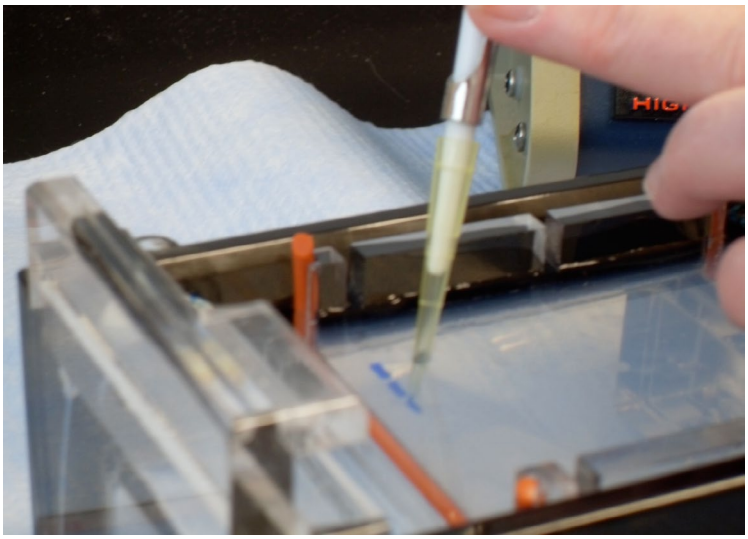
Sakai, RK, Scharf, S, Faloona, F, Mullis, KB, Horn, GT, Erlich, HA & Amheim, N (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230:** 1350-1354.

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'll analyze the products of the PCR reactions from the previous lab.

Objectives

- Prepare agarose gels for separating DNA molecules
- Separate DNA molecules by electrophoresis
- Visualize DNA molecules on gels using intercalating dyes
- Calculate the size of DNA molecules using molecular weight standards

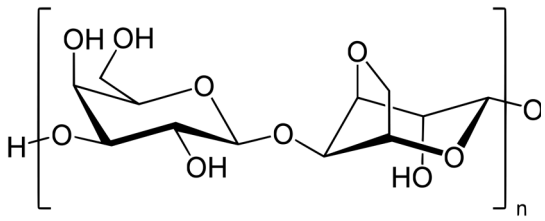
Agarose gel electrophoresis

Background

Agarose gels provide a simple method for analyzing preparations of DNA. Although the base compositions of individual DNA molecules vary, the basic chemical structure of DNA is the same for all DNA molecules. DNA molecules share the same charge/mass ratio, which imparts similar electrophoretic properties to DNAs of widely varying lengths.

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 the repeating disaccharide (1 3)-β-D-galactopyranose-(1 4)-3,6-anhydro-α-L-galactopyranose (right). A typical agarose molecule contains over one hundred monomers. 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 uniformity of DNA structure gives molecules the same charge/mass ratio, which makes it straightforward to calculate the sizes of DNA fragments. (Molecular weights are not always very meaningful with DNA fragments, since a single molecule can contain the sequences of multiple genes.) In general, the migration of DNA molecules is inversely proportional to the logarithm of their lengths, or number of base pairs.

Although the relationship between migration rate and DNA length holds true in most cases, 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. Because of this variability, DNA standards with known

sizes should always be run on the same gels with samples being analyzed. Importantly, the 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.

Size and conformation of DNA

The migration rate of a DNA molecule depends on both its size and its conformation. 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 circular plasmids, is much less predictable. The migration rates of these more structured forms are influenced by the density of coils, the presence of nicks, and other structural features.

Buffer systems

The migration rates of DNA molecules are also affected by the composition of the gel and running buffer. The migration rate of a DNA molecule decreases as the concentration of agarose in the gel increases. The migration rate of a DNA molecule is somewhat higher in a gel made with TAE (Tris: acetate: EDTA) buffer than with TBE (Tris: borate: EDTA) buffer.

Fluorescent intercalating agents are used to visualize DNA fragments 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 remains low. The most commonly used dyes are ethidium bromide and variants of SYBR Green.

SAFETY NOTE: All DNA intercalating agents are potential mutagens. Be sure to wear gloves when handling solutions with intercalating dyes.

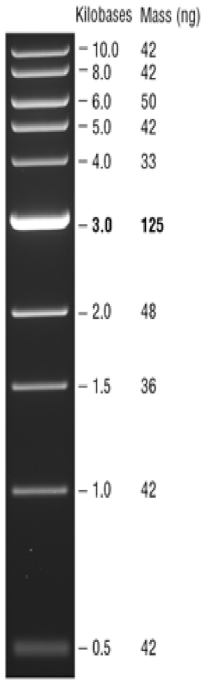
DNA intercalating agents have an absorbance maximum in the long ultraviolet range, so they are viewed with transilluminators that emit light with wavelengths close to the absorbance maximum of the dyes.

SAFETY NOTE: Be careful to wear protective eyewear if you look at an unshielded transilluminator. UV light is damaging to the eye!

Agarose gel electrophoresis

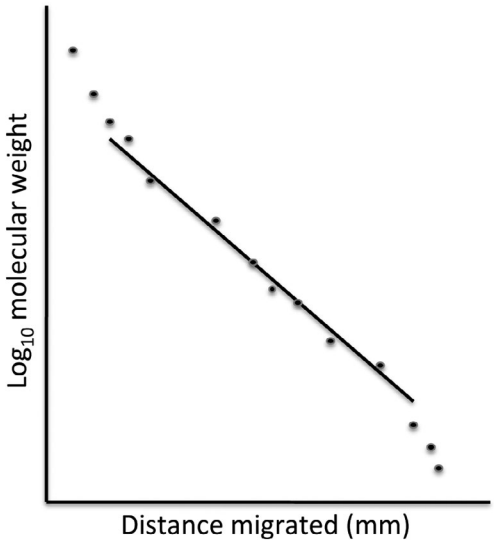
The sizes of DNA fragments can be calculated using standards

Agarose gels provide information about the sizes of DNA fragments, if they are calibrated with appropriate size standards. Agarose gels also provide information about the quantity of DNA in a particular fragment, because the amount of intercalating agent that a DNA molecule binds is proportional to its mass. Consequently, the intensity of a band reflects the amount of DNA in the band.



The grainy picture on the right shows the 1Kb ladder, which you will use as a standard in this course. The ladder is a proprietary preparation of DNA fragments produced by the New England Biolabs company. The 3Kb fragment has ~2.5 more DNA than the other fragments, so that it appears brighter in a stained gel. Using the ng values to the right, it is possible to use this standard to visually estimate the quantity of DNA in an unknown sample.

Many different standards are available for DNA molecules. Investigators choose standards with a similar size range to the molecules that they are studying.



To calculate the sizes of bands in other lanes of the gel, you would first construct a standard curve like the one at the left from the migration of the DNA fragments in the 1 kb ladder.

Plot the log $_{10}$(number of base pairs) for each fragment on the y-axis against the distance that each fragment migrated on the x-axis. You will be able to estimate the sizes of fragments in your sample by interpolating on the standard curve.

Note: Always visually confirm that your estimated sizes are correct!!

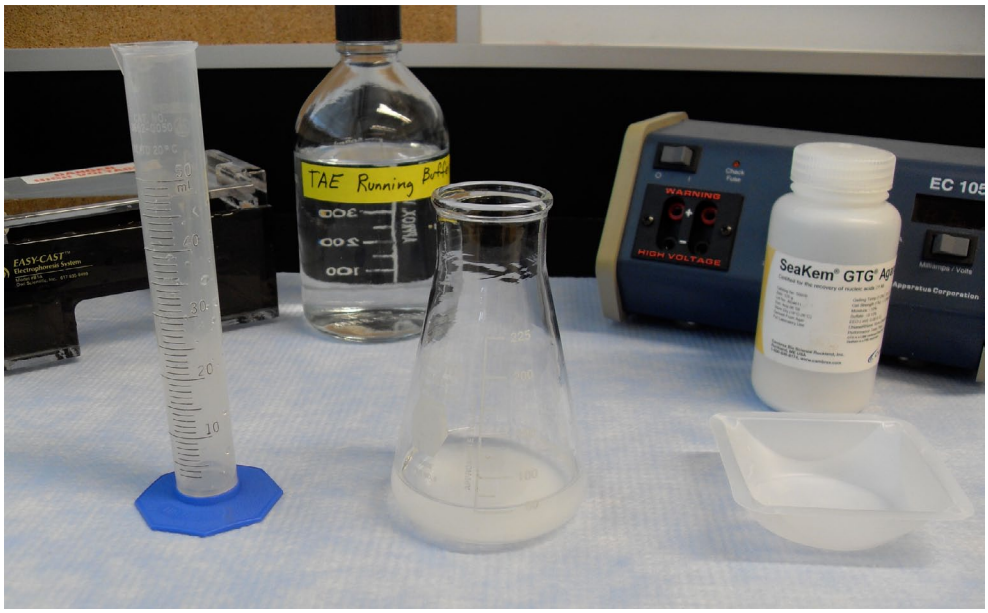
Prepare the agarose gel

In this course, we will use agarose gels to separate DNA molecules produced in PCR and restriction digests. These molecules are well-resolved on 1.25% agarose gels prepared in TAE buffer, which provides good separation of molecules ranging in size from 500 bp - 10 kbp. 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) agarose gel that fits your casting platform. Most of the gel apparatuses in the lab are the BioRad Mini-Sub GT systems, which have a 7 cm x 7 cm casting tray. These apparatuses accommodate 30-40 mL gels. 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! Please 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. Add 1 volume (vol) of 6X loading dye to 5 vol of each sample. (You would add 1 µL of loading dye to 5 µL of sample.) The loading dye contains two dyes, bromophenol blue and xylene cyanol, which migrate with "apparent" molecular weights of ~5 kb and ~0.5 kb, respectively. Loading dye also contains glycerol, which makes the sample dense enough to sink to the bottom of the sample well.) Briefly, centrifuge each sample to mix the dye and samples, if necessary.

Add the loading dye DIRECTLY into the tube containing your PCR products.

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. The DNA fragments 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. The wells should hold up to 20 µL of sample. For PCR reactions, load 5-10 µL of the original PCR reaction (6-12 µL after sample buffer has been added). 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.

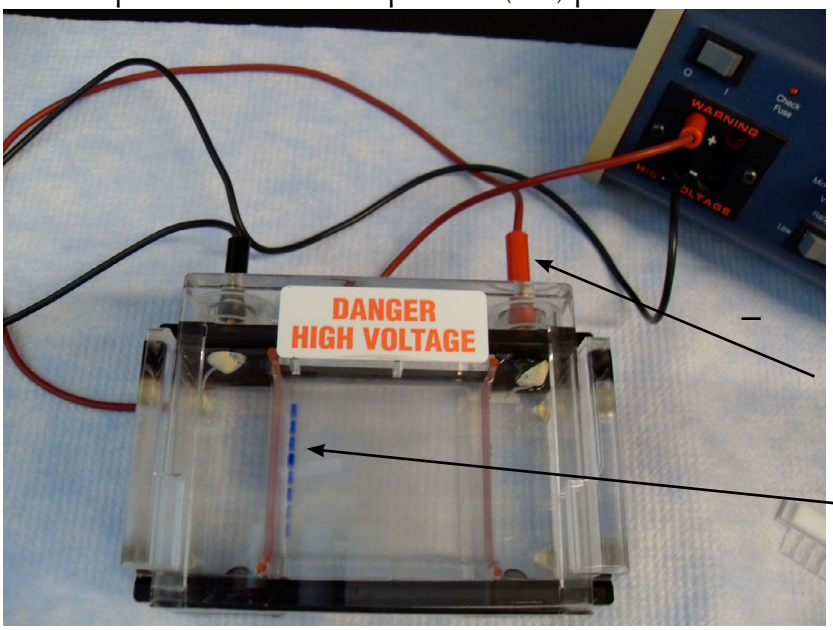
5. Place the lid on the electrophoresis tank and connect the electrodes to the power supply (black-to-black and red-to-red). ***Make sure that the polarity is correct before continuing!***

6. Turn on the power and apply a constant voltage of 75 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 elecctrophorese toward the postivie (red) pole.



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

Samples are loaded into wells.

Agarose gel electrophoresis

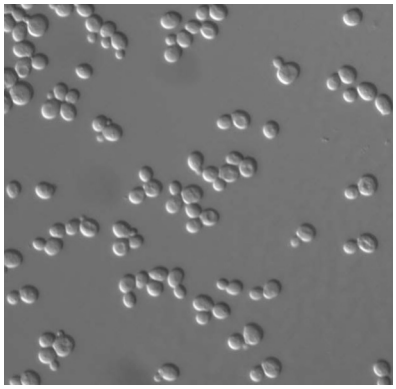
Stain and analyze the agarose gel

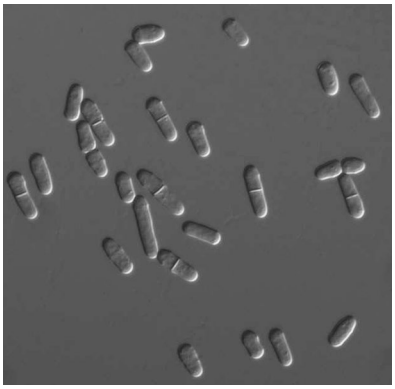
SAFETY NOTE: Wear disposable gloves when staining gels. Gloves are important when working with intercalating dyes, which are potential mutagens.

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 ethidium bromide (EtBr) solution (10 mg/mL), following your TAs instructions. What is the approximate concentration of EtBr in the staining solution? Note: EtBr stocks are light sensitive, so they are stored in the dark.
2. Place the tray on a rocking platform and rock gently for 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 2 minutes.
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. Turn off the transilluminator immediately after you photograph the gel. Save the picture and email a copy to yourself.
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 fragments in your samples by comparing their migration to that of the standards. Are the sizes consistent with your expectations? (In this experiment, the PCR products will have similar sizes, so a standard curve will not be useful.)
9. To construct a standard curve, measure the distance that each fragment migrated from the sample well. Make a table in your notebook with the size of each DNA standard and the distance that each fragment migrated on the gel. Plot the $\log_{10}$(fragment length) of each standard on the y-axis and the distance that each standard migrated on the x-axis.

Chapter 9

Protein conservation





The budding yeast, *Saccharomyces cerevisiae*, and the fission yeast, *Schizosaccharomyces pombe*, are separated by close to a billion years of evolution. Our goal for the semester is to determine if the proteins involved in methionine and cysteine biosynthesis have been conserved between these evolutionary divergent species. In this lab, you will use bionformatics tools to study the evolution of these protein sequences.

Objectives

- Learn the one letter code for amino acids
- Understand how amino acid side chain chemistry is reflected in the BLOSUM62 matrix
- Use the BLASTP algorithm to compare protein sequences
- Prepare a multiple sequence alignment to identify conserved regions in a protein

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**.

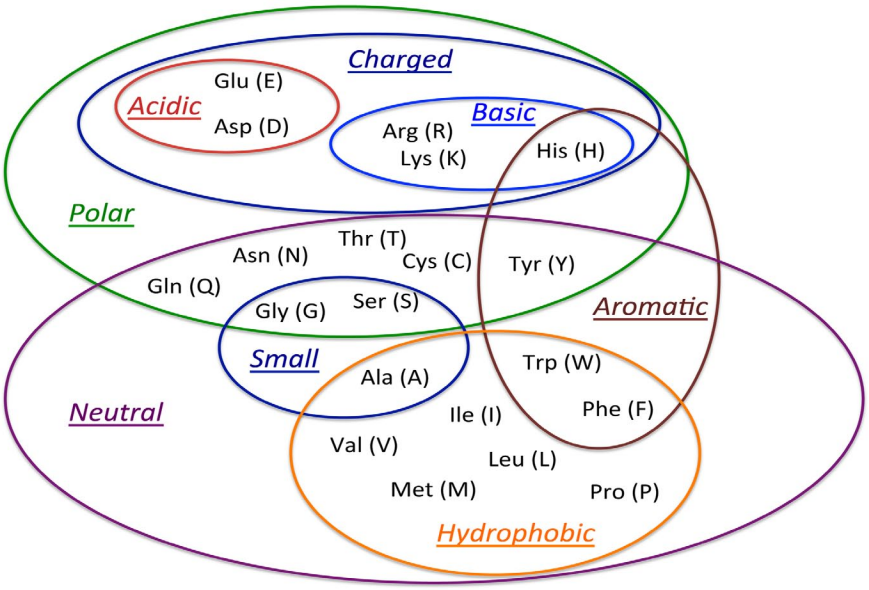
Homologs within the same genome are referred to as paralogs. **Paralogs** arise by gene duplication, but diversify over time and assume distinct functions. Although a whole genome duplication occurred during the evolution of *S. cerevisiae* (Kellis *et al.*, 2004), only a few genes in the methionine superpathway have paralogs. Interestingly, *MET25* 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.

Protein function is intimately related to its structure. You'll 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 molecular biologists.

Amino acid R groups have distinct chemistries

Each of the 20 amino acids commonly found in proteins has an R 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. Size is also important. R groups vary considerably in the bulkiness of their chains. 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. How many hydrophobic amino acids are found in this peptide?

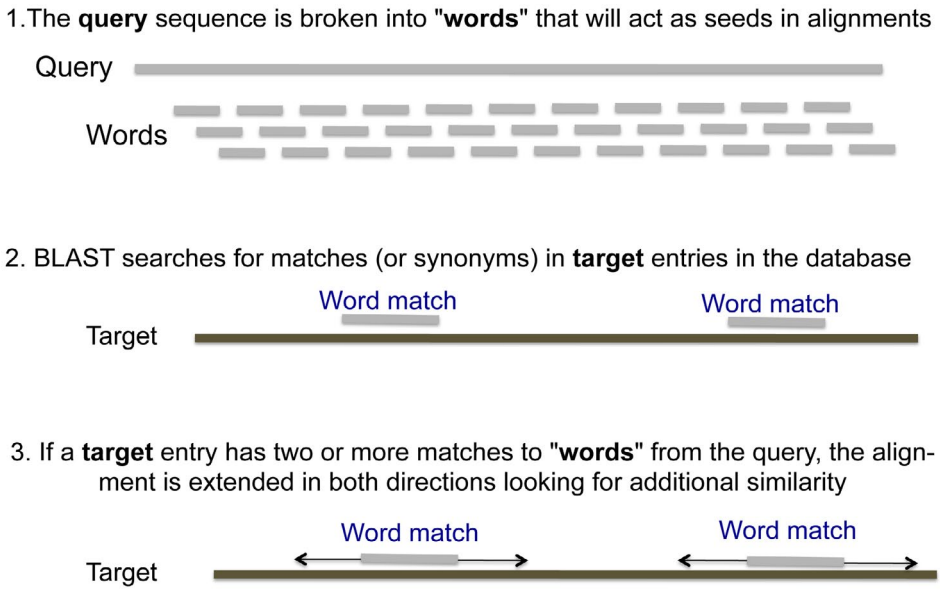
4. Write the name of a music group that you enjoy. Then transpose the name into an amino acid sequence written with the 3-letter code. Pass the amino acid sequence to a friend and have him/her decode it. (Note: the 1-letter code uses all of the alphabet, except B, J, O, U, X and Z).

Chapter 9

BLAST algorithms are used to search databases

There are many different algorithms for searching databases, but BLAST algorithms are some of the most popular. BLAST is an acronym for Basic Local Alignment Search Tool, which aptly describes its function. The original BLAST algorithms were developed and published by a group of investigators in 1990 (Altschul *et al.*, 1990). Since that time, investigators have refined and diversified the original BLAST programs so that a whole family of BLAST programs is currently available. BLAST programs are widely used because they are able to quickly search through large amounts of sequence data. The key to BLAST's speed is its use of local alignments that serve as seeds for more extensive alignments. Other algorithms perform global sequence alignments that match the entire sequence of a query sequence against database information. Global sequence alignments are more sensitive than BLAST in finding related sequences, but they require considerably more processing time. BLAST searches are quite adequate for our purposes, and we will use the online BLAST tools available at the NCBI site throughout this course.

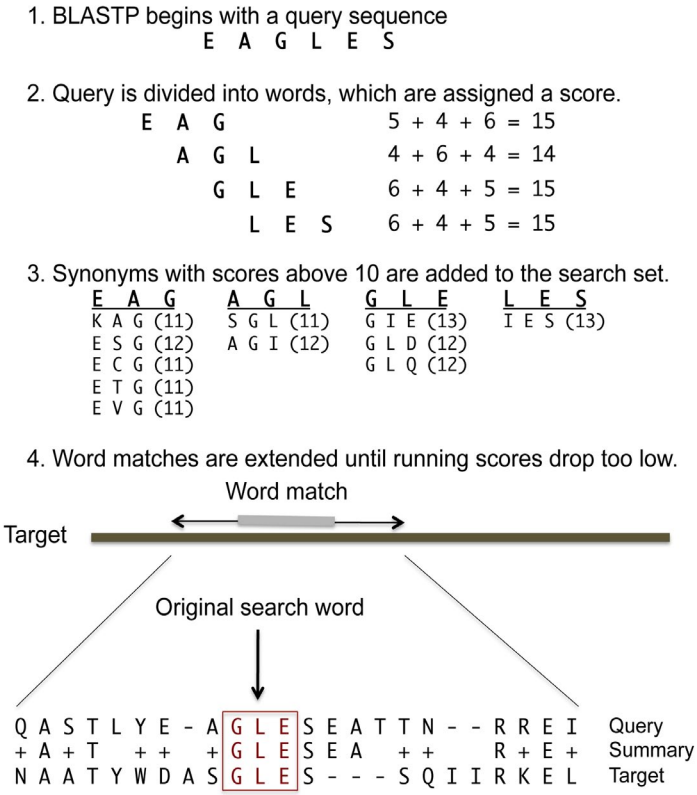
The BLASTN and BLASTP algorithms are designed for searching nucleotide and protein databases, respectively, and they use different scoring matrices and search parameters. BLAST algorithms begin by breaking down query sequence into short "words" and assigning numerical values to the words. A scoring matrix is used to assign numerical values to the words. Words and synonyms above a threshold value are then used to search databases. The default word size for BLASTN is 28 nucleotides, while the default word size for BLASTP is 3 amino acids. This difference in word size is because nucleic acid sequences are written in 4 letters (ACGT), while protein



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.

sequence is written with 20 different letters. The probability of obtaining random, but irrelevant, matches are much greater with nucleotide sequenes. A stretch of 3 amino acids should arise randomly once in every 8000 tripeptides, compared to once in every 64 trinucleotides. In this lab, we will use the BLASTP algorithm, which is more useful than BLASTN for studying protein evolution, because it overlooks synonymous gene mutations that do not change an amino acid.

In BLASTP, a sequence is broken into all possible 3-letter words using a moving window. A numerical score for each word is assigned by adding up values for amino acids from a scoring matrix. Using the BLOSUM62 matrix (below), BLASTP first finds word with a score of 12 of more in a query sequence. BLASTP then identifies potential synonyms that differ from the word at one position, and synonyms above a defined threshold value are included in the search set. NCBI BLASTP uses a default threshold of 10 for synonyms, which can adjusted by the user. Using this search set, BLAST rapidly scans a database and identifies protein sequences that contain at two words/synonyms from the search set. These sequences are set aside for the next part of the BLASTP process, when 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. Again, users can adjust the gap penalties, which are usually quite severe. In the NCBI default settings, the presence of a



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.

gap brings a 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.

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. An E-value of 1 does not necessarily imply, however, that the match has no biological relevance. More nuanced analysis may simply be needed to evaluate the match.

BLOSUM62 scoring matrix

The results obtained in a BLASTP search depend on the scoring matrix used to assign numerical values to different words. A variety of 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 a large number of protein sequences that were more than 62% identical to each other. Investigators computationally determined the frequency of all 2^{10} possible amino acid substitutions that had occurred in these conserved blocks of proteins, and 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 (below) 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

Ala	**A**	4																		
Arg	**R**	-1	5																	
Asn	**N**	-2	0	6																
Asp	**D**	-2	-2	1	6															
Cys	**C**	0	-3	-3	-3	9														
Gln	**Q**	-1	1	0	0	-3	5													
Glu	**E**	-1	0	0	2	-4	2	5												
Gly	**G**	0	-2	0	-1	-3	-2	-2	6											
His	**H**	-2	0	1	-1	-3	0	0	-2	8										
Ile	**I**	-1	-3	-3	-3	-1	-3	-3	-4	-3	4									
Leu	**L**	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4								
Lys	**K**	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5							
Met	**M**	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5						
Phe	**F**	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6					
Pro	**P**	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7				
Ser	**S**	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4			
Thr	**T**	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5		
Trp	**W**	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11	
Tyr	**Y**	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
Val	**V**	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4
		A	**R**	**N**	**D**	**C**	**Q**	**E**	**G**	**H**	**I**	**L**	**K**	**M**	**F**	**P**	**S**	**T**	**W**	**Y**	**V**

Henikoff, S. and Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**:10915-10919.

Exercise 2 - The BLOSUM62 matrix

1. Use the Venn diagram on p. 87 to hypothesize why relatively high values (2 and above) occur for the substitutions:

   Asp to Asn

   Asp to Glu

   His to Tyr

   Ile to Val

   Lys to Glu

2. Calculate the word score for the following tripeptides. Which would you expect to occur most frequently in a protein database?

   Leu - Cys - Pro

   Gln - Ala - Met

   Asp - Glu - His

Chapter 9

Exercise 3 - BLASTP

In this exercise, you will use BLASTP to find a homolog for your protein in *S. pombe*. Direct your browser to the BLAST tools at NCBI:

http://blast.ncbi.nlm.nih.gov

- Enter the NP_ number for the S. *cerevisiae* Met protein (p. 45) in the query box
- For the database, select reference proteins
- Enter *S. pombe* for the organism
- Click BLAST

- On the results page, note the graphic summary at the top which gives you an instant idea about the extent and strength of the match with *S. pombe* sequences.

  Does the match extend for the entire length of the protein?

  How many homologs are present in the S. pombe genome?

Cursor down to the alignment statistics.

Record in the table on p. 94:

NP__ record number of the *S. pombe* match
total score
how much of the query sequence was aligned (coverage)
E-value for the alignment

Cursor down to see the actual alignment of the S. cerevisiae and S. pombe sequences.

- Take note of the center row in the alignment which summarizes the homology between the protein sequences. If an amino acid is conserved between the two species, its 1-letter code name is shown in that row.

  What do the pluses in the center row indicate?

  Are some regions of the sequences more conserved than others?

  Did BLASTP add gaps to one or both of the sequences in the alignment? If so, are they more common in less conserved or highly conserved regions?

Click on the link to the NP_ record for the S. pombe ortholog. Record the EC number for the protein. Is it identical to that for the *S. cerevisiae* enzyme?

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. Model organisms are easily maintained in large numbers in the laboratory. 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

1. Collect the sequence 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 set up 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. 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 automatically detects homologs in 20 sequenced eukaryotic genomes.

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

Note the species that have been searched for homologs. 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 distribution of homologs in eukaryotes. 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 records page has links on the left to gene summaries, which you may want to investigate. Links on the right side of the page bring you to the NP___ record. The Homologene output always provides a graphical depiction of protein domains in each of the homologs.

- *How many domains are found in your protein?*

- *Are the domains equally well-conserved between species?*

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. Record the NP___ numbers for the species listed on the Homologene page. Note: Not all the species listed on the previous page may have homologs! Add the NP_ numbers for *E. coli* and *B. subtilis* 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.

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.***

2. Prepare the multiple sequence alignment.

Examine the BLASTP results that you recorded for your gene. You will next construct a multiple sequence alignment and phylogenetic tree, using the Phylogeny suite of programs. 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 a browser window or tab to www.phylogeny.fr . Click the "Your workspace" tab and set up an account to store your work.
- Under the heading Phylogeny analysis tab, select One Click. 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, but we will ignore them. *We'll let Phylogeny make the decisions for us!*
- Construct a FASTA file containing all the sequences that you would like to compare. You will be pasting files from NCBI records directly into the Phylogeny data entry box. The title line of a FASTA file begins with a ">" character and end 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 NP_xxxxxx name.*** (You will see why later!) **IMPORTANT: Do NOT use a text editor or Word to work with sequences.** These programs introduce hidden punctuation that will interfere with Phylogeny programs.
- The first file in our comparisons will be the *S. cerevisiae* protein sequence. Find the record in the NCBI Protein Database, using the NP_ number as a search term. When the record comes up, click the FASTA link at the upper left side of the record. (Alternatively, you can click to the NP_ record from the Homologene page.)
- Copy the title line (begins with >) and the entire amino acid sequence
- Paste the FASTA record into the Phylogeny data entry box
- Repeat step 3 with each of the sequences that you would like to compare.
- When you are finished, give your project a title, enter your email address (the analysis can take a little time) and click the Submit button. Your results will be posted on a web page.

3. Export the multiple sequence alignment

- Click on the alignment tab to view the multiple sequence alignment.
- Ask for the output in ClustalW format. The output appears on a new web page.
- Right-click on the page and download the Clustal alignment with a new filename that makes sense to you. The page will download as an ASCI text file that you will be able to open with Word or a text editor.
- Open the file in a word processor. Adjust the font size so sequences are properly aligned and fit on the page. Use a non-proportional font such as Courier so all characters line up.

3. Construct a phylogenetic tree.

- Click the tree rendering tab to access your phylogenetic tree.
- Use the editing tools to alter the appearance of your tree as you see fit. Pay particular attention to the legends in the "leaves" of the tree, which may have been altered during the rendering processing. For example, if there are no paralogs for the protein, you may want to replace the NP_ numbers with legends that make the figure more comprehensible to a reader.
- Download the file in a format of your choice.

Discussion questions

Which homolog would you test after the *S. pombe* homolog? Why?

Which gene has the strongest homolog in *E. coli*? *B. subtilis*? *A. thaliana*, etc.?

Look at the methionine superpathway on p. 57. Which parts of the pathway are the most widely distributed? Which reactions are more highly conserved in bacteria? in plants? in animals?

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool *J. Mol. Biol.* **214**: 403-410.

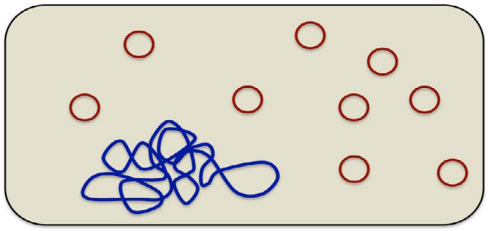
Henikoff, S., and Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* 8**9**: 10915-10919.

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, E.W., 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



Plasmids are cloning vectors that are widely used in molecular biology and they play important roles in the laboratory. Plasmids are small, circular pieces of DNA that replicate independently of the host chromosome. The first plasmids used in the lab were derivatives of naturally-occurring plasmids found in bacteria. Since their discovery, molecular biologists have added many features to plasmids to suit a variety of applications. In this lab, each team will isolate three plasmids from bacterial strains.

Objectives

- Understand the structure and replication of plasmids
- Understand features that have been engineered into plasmids for a variety of experimental purposes
- Learn how the physical properties of plasmids are used in their purification
- Isolate plasmids from *E. coli*
- Use ultraviolet spectroscopy to estimate the concentration of plasmid DNA

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* or *CYS* 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/CYS* 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 usually higher than 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

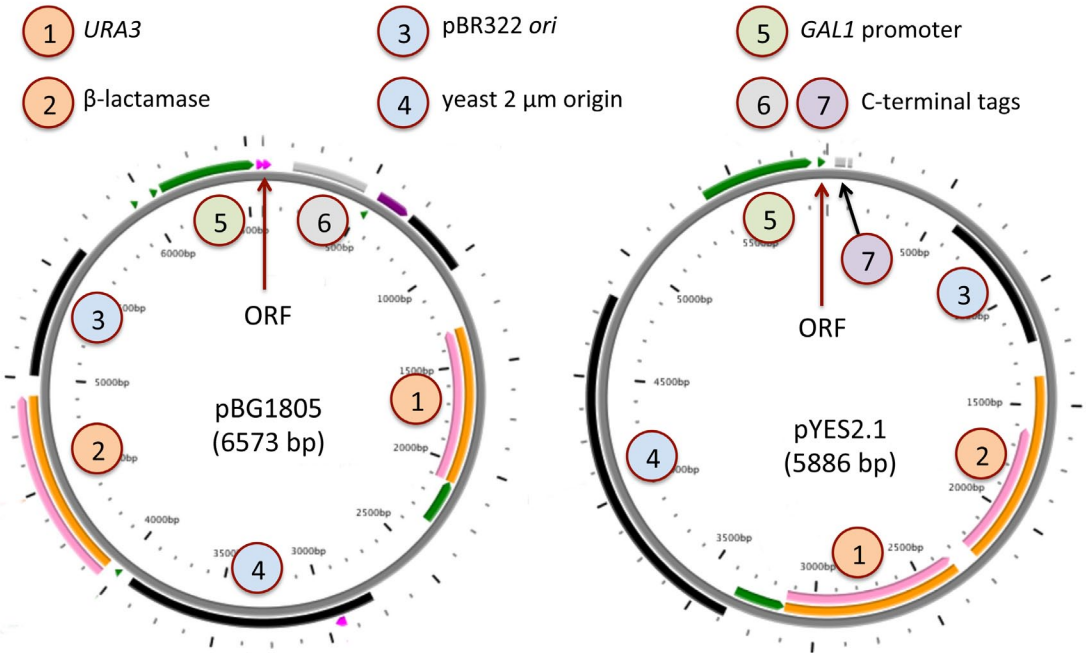
that interferes with cell wall synthesis in bacteria. The plasmids also contain the *S. cerevisiae URA3* gene, which allows *ura3* mutants like BY 4742, the parent strain of our mutants, to grow in the absence of uracil.

Promoters control transcription of plasmid sequences

The plasmids that we will use this semester contain *MET* and *CYS* genes that have been cloned into plasmids containing the promoter sequence for the yeast *GAL1* gene (Johnston, 1987). Transcription from the *GAL1* promoter is normally regulated 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* proteins or their *S. pombe* homologs. The presence of the *GAL1* promoter will allow you to manipulate expression of the Met or Cys 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. From the placement of these elements, however, it is clear that the plasmids were developed independently 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.)

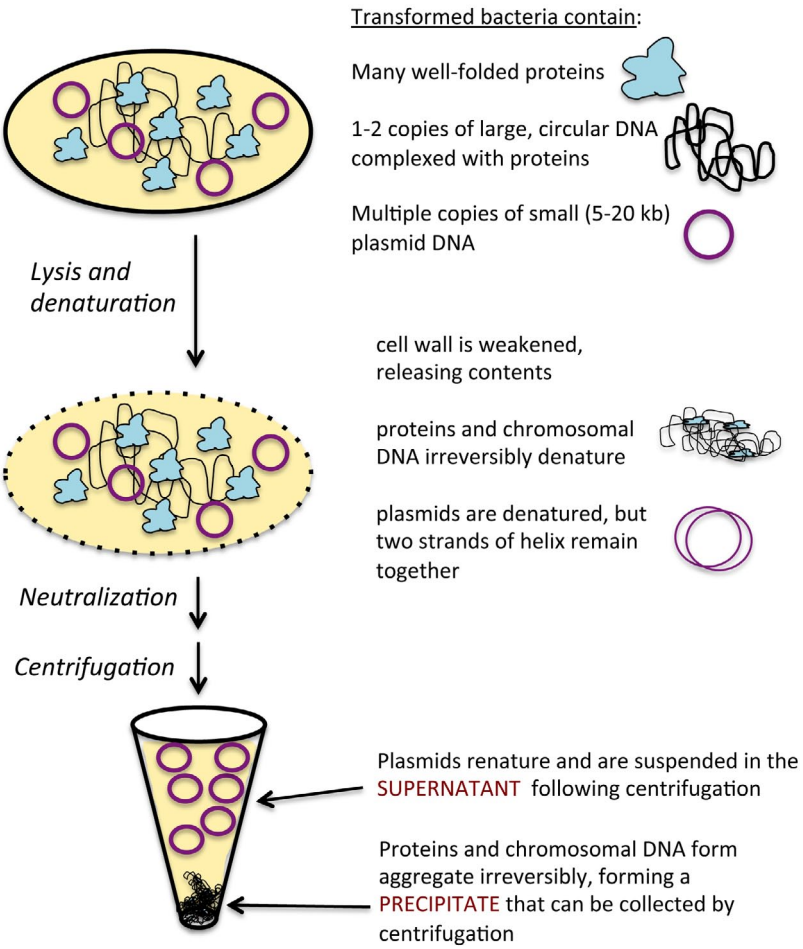
pYES2.1 -LacZ would designate the bacterial LacZ gene cloned into pYES2.1.

NOTE: 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 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 chromosomal and plasmid DNA.
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

Exercise 1 - Plasmid isolation with the Zyppy™ kit

Concentrate the plasmid-bearing bacterial cells

1. Your group will receive 3 bacterial cultures that were grown overnight in Luria Bertani (LB) media containing 100 µg/mL ampicillin. The density of cells in the culture is expected to be 3-4 X 10^9 cells/mL. *What is the purpose of the ampicillin? How does it work?*
2. Transfer 1.5 mL of each 3 mL culture to a microcentrifuge tube. Concentrate the cells by centrifuging them 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 will 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 step 3 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.

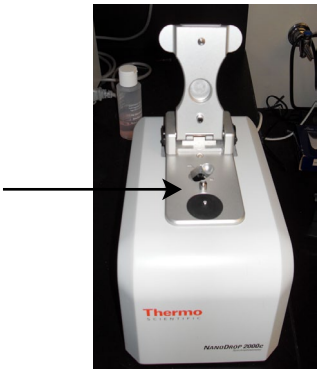
Elute the plasmid DNA

14. Transfer the Zyppy column to a clean (and appropriately labeled) 1.5 mL centrifuge tube, leaving the lid of the tube open.
15. Carefully, add 100 µL of sterile water 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.
16. Allow the buffer to percolate into the column by letting the column stand upright in the microcentrifuge fuge 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.***

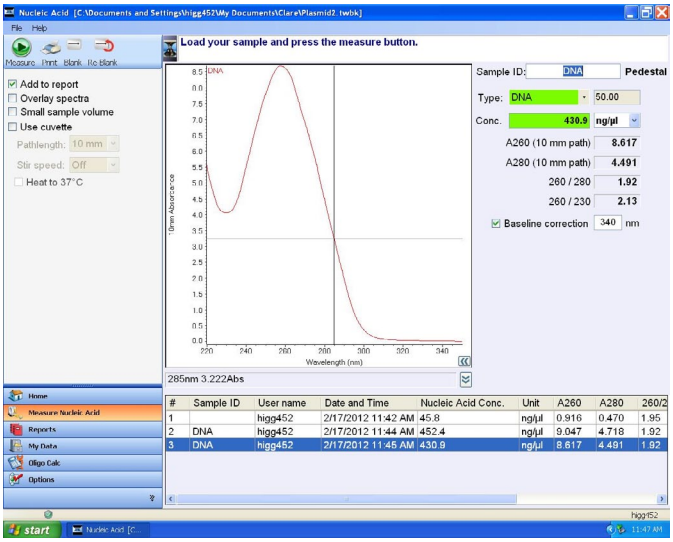
Exercise 2 - Measure the plasmid DNA concentration

The concentration and purity of DNA samples can be estimated from their absorbance of ultraviolet (UV) light. Nucleic acids absorb light strongly in the UV range, with an absorbance maximum of ~260 nm. An A_{260} of 1 corresponds to ~50 µg/mL DNA. Proteins also absorb UV light due to their aromatic side chains, but much less strongly than nucleic acids. Proteins have a different absorbance maximum of ~280 nm. An A_{280} of 1 corresponds to 1 mg/mL protein. Because of the differences in protein and nucleic acid absorbances, the A_{260}/A_{280} ratio of a sample is useful in determining the purity of a DNA sample. Pure DNA has an A_{260}/A_{280} of ~1.8. Protein or chemical contamination will lower this ratio.

In this laboratory, you will use an instrument called the Nanodrop™ to obtain an ultraviolet absorbance spectrum for your plasmid DNA samples. The Nanodrop software will also calculate the concentration of DNA and the A_{260}/A_{280} ratio. The NanoDrop uses fiber optic technology and the surface tension of a sample to provide accurate absorbance measurements over a large dynamic range of DNA concentrations. From a practical point of view, this means that you need to sacrifice only small amounts of your precious DNA sample. The NanoDrop is amazingly simple to use, and your TA will demonstrate the appropriate use of the instrument.



1. Pipette 3 µL of your plasmid sample on the pedestal on top of the pin (shown at the end of the arrow at left).
2. Close the arm and record the absorbance.
3. When the measurement is complete, open the arm and wipe the pedestal surface with a lint-free wipe in preparation for the next sample.
4. Record the absorbance readings in your notebook.



The Nanodrop results screen shows the UV spectrum of the sample and the estimated DNA concentration

References

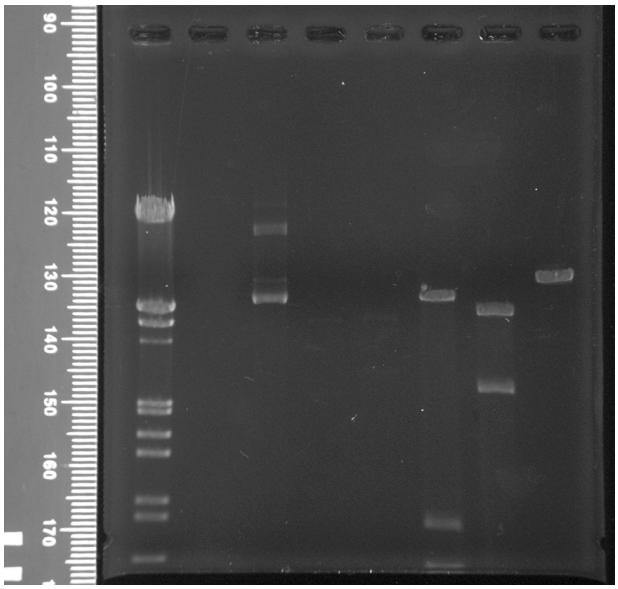
Dong, X, Stothard, P, Forsythe, IJ, & Wishart, DS (2004). PlasMapper: a web server for drawing and auto-annotating plasmid maps. *Nucleic Acids Res.* **32**, W660-W664.

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.

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 carrying are *S. cerevisiae* and *S. pombe* ORFs or bacterial *LacZ*.

Objectives

- Understand the biological origins and functions of REs
- Understand the structures of restriction sites in DNA
- Devise a strategy to distinguish between three yeast overexpression plasmids using RE digests
- Analyze restriction digests on agarose gels

Chapter 11

In the last experiment, you isolated plasmids carrying *S. cerevisiae* or *S. pombe* ORFs from transformed bacteria. Each group was given three different plasmids. One of the three plasmids, carries the *S. cerevisiae MET* or *CYS* 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* or *CYS* 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

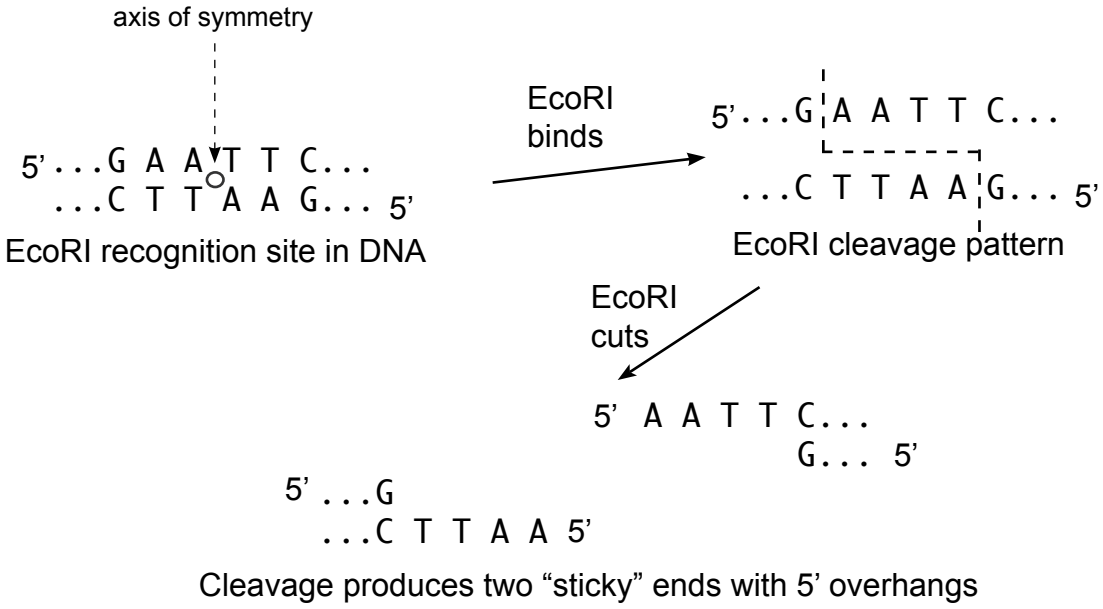
Bacterial restriction/modification systems protect against invaders

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 bacteria protect their own DNA from cleavage with a methyltransferase that modifies the same restriction sites in the bacterial DNA. The combined activities of the endonuclease and methyltransferase are referred to as a restriction/modification system. In Type I enzymes, the endonuclease and methyltransferase activities are part of the same complex. The REs used in this laboratory are Type II enzymes. The Type II endonucleases and methyltransferases do not form a complex and function independently.

To understand how REs work, it may be helpful to look at a single enzyme in greater detail. One of the best-studied restriction enzymes is EcoR1. 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*, or *E. coli*. Its name begins with the genus and species (Eco for *E. coli*), followed next by a strain identifier (R for RY13), and ends with a Roman numeral that distinguishes the different enzymes in the strain. Strain RY13 of *E. coli* contains multiple endonucleases, but only EcoRI and EcoRV, are widely used in molecular biology. Purification of REs from their natural source is a laborious procedure that requires large quantities of bacteria, since the REs are not abundant proteins in bacteria. To facilitate the production of these useful enzymes for routine laboratory use, molecular biologists cloned the coding sequences of the endonucleases into bacterial expression plasmids. Today, most commercially available enzymes are recombinant proteins overexpressed in bacteria. (Along the way, scientists also introduced some mutations that generated enzymes with enhanced stability or higher activity.)

Restriction enzymes cleave specific sites in DNA

Restriction enzymes like EcoRI are frequently called 6-cutters, because they bind to 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. A DNA fragment of that length is very useful in the lab, since it is long enough to contain the coding sequence for some proteins and it is easily resolved on common agarose gels. The sequence that EcoRI recognizes in double stranded DNA is G A A T T C. The sequence is a palindrome with a two-fold axis of symmetry, because reading from 5' to 3' on each strand gives the same sequence. The palindromic nature of the restriction site is more obvious in the figure below. The circle in the center of the restriction site denotes the axis of symmetry. EcoRI binds DNA at the restriction site and catalyzes the hydrolysis of the phosphodiester bond between G and A on either strand.



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.

The structure of EcoRI is consistent with an enzyme that catalyzes identical reactions at two different sites. Crystallization data shows that EcoRI is a homodimer, an enzyme with two identical subunits. The figure on the next page (PDB datafile 1QPS, unpublished) shows two different views of EcoRI bound to DNA. In the left view, we're looking down the axis of the DNA helix. The right view is a side view. Take a close look at both structures, using the DNA helix as the center. In these models, alpha helices and beta sheets are given different colors or shades. The symmetry in the enzyme is quite apparent.

Crystallization structures (1QPS) show that EcoRI is a symmetric homodimer that binds DNA.

Selecting restriction enzymes for experiments

Restriction enzymes are widely used in molecular biology for both the analysis and construction of DNA molecules. The first step in an experiment involving restriction enzymes is to map the position of restriction sites in a target DNA molecules. Many software programs also identify open reading frames as they scan DNA sequence for RE recognition sites. These programs usually generate a graphical output as well as tabular data with details about the lengths and positions of the restriction site 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 enzymes, investigators consider the number and lengths of the fragments that will be generated in a digest, the position of the restriction sites relative to a gene of interest, as well as the kind of ends in the restriction fragment.

Restriction maps are used to identify DNA molecules

The sequence of a DNA molecule determines the distribution of restriction sites. The map of restriction sites in a piece of DNA is known as a restriction map. With careful planning (below), investigators can use restriction maps to design experiments that will distinguish between DNA molecules. 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. In this experiment, you will use restriction maps of your three unknown plasmids to devise an experimental strategy which will allow you to identify the three plasmids.

Restriction endonucleases are used in the construction of DNA molecules

The availability of a large number of REs with distinct specificities has been critical for recombinant DNA technology, which involves the construction of novel DNA molecules with properties suited for a particular purpose. In these experiments, REs serve as the molecule

scissors that cuts DNA into segments that can be joined to additional restriction fragments to produce a desired molecule. REs are used in conjunction with DNA ligases, which serves as the "paste" by catalyzing the formation of phosphodiester bonds between restriction fragments. The kind of end generated by an RE is particularly important when investigators are constructing a recombinant DNA molecule. In the figure on the previous page, you saw how EcoRI produced DNA fragments with 5' overhangs. Other enzymes produce fragments with 3' overhangs, while still others produce blunt ends with no overhangs. Ends with overhangs are often referred to as "sticky" ends, since they will form hydrogen bonds with a complementary sequence on a second DNA fragment. By contrast, fragments with blunt ends (*i.e.* no overhang) can anneal with any other blunt end, regardless of the adjoining sequence.

We will not be generating new recombinant molecules in this class, but it is important to understand their importance to molecular biology. Consider the pBG1805 and pYES2.1 plasmids. From the map on p. 99, you can see that these complex plasmids were constructed by stitching together functional elements from pieces of DNA obtained from a variety of sources.

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 restriction endonuclease functions optimally under a defined set of reaction conditions, including variables such as temperature, pH, and the concentrations of metal ions and salts. The manufacturer's supply restriction enzymes with buffers that generate reaction conditions close to those that are optimal for the enzyme. Some of the buffers contain bovine serum albumin (BSA), an abundant protein from cow serum. BSA is included largely for its stabilizing properties. In general, enzymes are more stable at higher protein concentrations. BSA does not affect the enzymatic activity of the REs, but protects particularly sensitive REs from denaturation. BSA may also saturate non-specific binding sites for REs on the sides of the test tubes used in the digests.

Like all enzymes, REs are subject to spontaneous denaturation and loss of function, so REs need to be handled with care. (By comparison, DNA is an exceptionally stable molecule.) The rate of protein denaturation increases as the temperature increases, so enzymes are kept on ice while reaction mixtures are being prepared. Denaturation is also accelerated at air/water interfaces, so it's important to minimize air bubbles and freeze/thaw cycles for solutions containing restriction enzymes. Manufacturers attempt to minimize denaturation by supplying the solutions in a solution of 50% glycerol, which does not freeze when the enzymes are stored at -20˚C.

Follow some simple rules when you use REs.

- Be sure to use the recommended buffer for a particular RE.
- Keep the reactions on ice until the incubation begins
- Be careful not to introduce bubbles. Mix the reactions by flicking the tubes with your fingers, rather than using the vortex mixer.
- Gather the reaction components at the bottom of the tube, NOT by centrifuging, but with a firm flick of the wrist.

Chapter 11

Exercise 1 - Plan the restriction digest

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

You will first need to construct the sequences of your overexpression plasmids. 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. After the sequence is constructed, you will 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.

1. ***Locate the coding sequence of your gene.***
   Use the NM___ accession numbers of your gene (p. 45 for *S. cerevisiae* genes or p. 94 for *S. pombe* genes) to access the nucleotide record in NCBI's Nucleotide database. The pYES2.1-*LacZ* sequence is posted on Blackboard. Note that this 8964 bp sequence has been modifed from naturally-occurring *LacZ* genes.

2. ***Assemble the complete nucleotide sequence of your plasmid.***

- Open the sequence file for the BG1805 or pYES2.1 vector sequence posted on Blackboard. The plasmid sequences are numbered so that the GAL1 promoter is at the 3'-end of the DNA sequence.
- Paste the coding sequence from the NM_ record (or the *LacZ* sequence) to the end of the plasmid sequence.
- After you paste the coding sequence at the end of the vector sequence, delete the last three nucleotides, which comprise the gene's stop codon. Stop codons are not included in the plasmid sequences, which encode fusion proteins with C-terminal extensions. (Note: the stop codon in the *LacZ* sequence has been removed by the manufacturer.)
- Remember that the first nucleotide of your *MET* or *CYS* gene is one nucleotide higher than the last nucleotide in the vector. For example, the BG1805 sequence contains 6573 nucleotides. The first nucleotide of a *MET* or *CYS* gene inserted into pBG1805 will be nucleotide 6574. The pYES2.1 plasmid contains 5886 nucleotides.

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. 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 RE that cut the plasmid, their recognition sites, the number of recognition sites, and the amount of enzyme activity in each of the four buffers used for digests. ***Your TA will tell you which REs are available in the lab.*** We will be analyzing the restriction fragments on agarose gels, which do a good job of resolving fragments ranging in size from ~200 bp to ~5000bp. Determine which of the REs available in the lab will produce restriction digests with a nice range of fragment sizes.

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 ***ONE*** RE that best allows you to distinguish the three plasmids. The restriction digests will be anallyzed on agaraose gels with **8 wells**.
- One well is reserved for the molecular weight standard.
- Two wells will be used for each plasmid.
  One well will contain plasmid that has NOT been digested with RE.
  A second well will contain plasmid that contains plasmid digested with RE.

Chapter 11

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 activity/ µL. A unit of activity (U) is assessed in a standardized assay. Restriction digests are usually set up to contain 2-5 U per µg plasmid DNA. We will be using 5 µL of plasmid miniprep DNA in each reaction. The REs that you will be using have been diluted to a concentration of 0.5 U/µL.

In your lab notebook, note which RE(s) you have decided to use. Check which 10X salt solution that you should use from the manufacturer's chart of : https://www.neb.com/tools-and-resources/usage-guidelines/nebuffer-performance-chart-with-restriction-enzymes

- Compute the volumes of each component that you will need for restriction digests with a final volume of 10 µL.

  5.0 µL plasmid - to give a final concentration of 0.2 - 0.4 µg plasmid DNA
  1.0 µL 10X buffer
  1.0 µL BSA (from a 1 mg/mL 10x stock) - only IF recommended
  0-1 µL deionized water (to bring total volume to 10 µL)*
  2.0 µL (1.0 U) restriction enzyme

The total reaction volume should be 10 µL.

- Draw a schematic in your lab notebook showing how the digests will be set up.
- Combine the components for each digestion in their own separate tubes, in the order listed above.
- Add the restriction endonuclease(s) last and mix by gently tapping the tube.
- Incubate the samples at 37 °C for 2 hr. Store the samples in the freezer.

Exercise 3 - Analyze the restriction digests on agarose gels

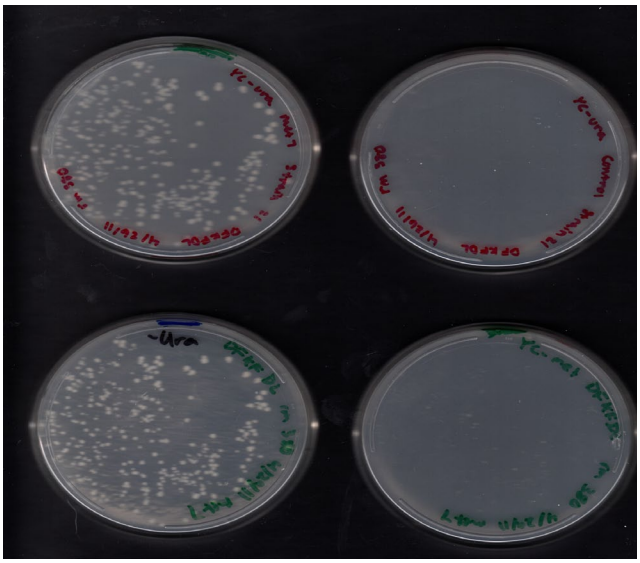
This exercise will be performed in the next laboratory session.

1. Plan your gel. Each group of students will prepare one agarose gel. Each student will run one lane with undigested plasmid and a second lane with plasmid that has been digested with restriction endonuclease(s). An additional lane should be reserved for molecular weight standards. Record in your notebook how you plan to load your gel.

2. Prepare, load and run your gels as described in Chapter 8. Be sure to record the details of every step in your lab notebook.

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* or *cys* 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* or *Cys* genes are functionally equivalent to *S. cerevisiae MET* and *CYS* genes.

Objectives

- Understand the principles of yeast transformation
- Understand how plasmid-encoded genes can complement gene deficiencies
- Transform *met* or *cys* strains with plasmids carrying *S. cerevisiae* and *S. pombe* genes involved in Met or Cys synthesis
- Use replica plating to test transformants for their ability to grow on various selective media

Chapter 12

In this lab, you may receive a preliminary answer to the semester's research question about the functional conservation of Met and Cys 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 and Cys 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* or *CYS* genes in the mutants. In complementation, the introduction of a foreign 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 transformation is still poorly 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.

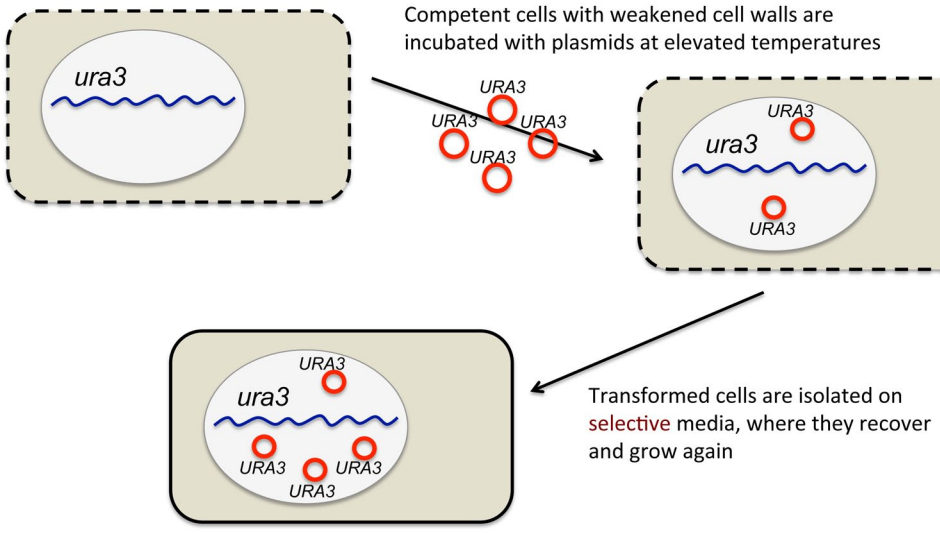
Techniques for yeast transformation are now standard in the laboratory. Depending on the details of the experimental procedure, reactions can yield as many as 10^6 transformants per μg DNA. The structure of the DNA used for transformation greatly affects the transformation efficiency. Transformation 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 incubating cells with a plasmid containing the yeast *URA3* gene at an elevated promoter (top). Transformed cells are selected on media that does not contail uracil (bottom).

Chapter 12

GAL1 promoter - experimental considerations

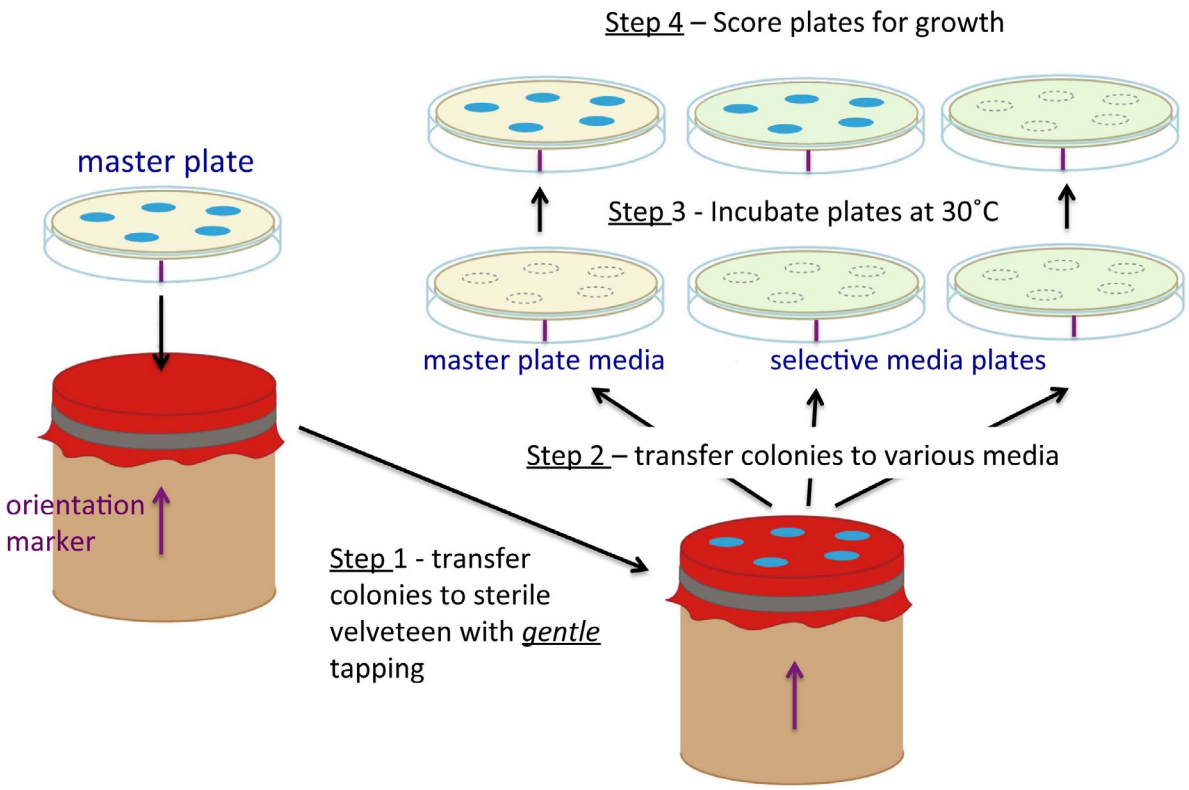
You may be wondering why we are not using *MET* or *CYS* gene complementation to isolate transformants, since this is the goal of our semester's project. There are several reasons for this. First, we need to insure that the overexpression plasmids have successfully transformed the deletion strains. *URA3* gene complementation offers a well-tested and reliable means to assess successful transformation. Successful *URA3* complementation serves as an important positive control for transformation if the plasmid-encoded *MET* and *CYS* genes fail to complement deletion strains.

A second issue relates to regulation of the plasmid ORFs by the *GAL1* promoter (Johnston, 1987) in the plasmids. The *GAL1* promoter is an inducible promoter that is normally repressed when cells are grown in glucose, the preferred carbon source for yeast. When galactose replaces glucose as the carbon source, yeast respond by significantly increasing the transcription of several genes, including *GAL1*, that are involved in the metabolism of galactose. In its normal chromosomal location, the *GAL1* promoter responds to a variety of positive and negative transcription regulators (Chapter 13). The *GAL1* promoter sequence has been widely used in molecular biology because it generally functions well in ectopic locations, such as plasmids.

Once you obtain transformants, you will analyze *MET/CYS* gene complementation using selective media containing either D-glucose and D-galacatose as the carbon source. The transformed cells will contain many more copies of the *GAL1* promoter than non-transformed cells, because both pBG1805 and pYES2.1 are multi-copy plasmids. Because the regulatory balance is altered in transformed cells, galactose and glucose may not function as simple "ON" and "OFF" switches. You should consider the possibility that "leaky" gene transcription could occur in the presence of the normal repressor, D-glucose. If this happens, *MET/CYS* genes would complement *met/cys* mutants grown in D-glucose. It is also possible that transformed cells could produce excessive quantities of Met and Cys 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 or Cys. 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 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/Cys, 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

Chapter 12

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, which would be required for linear pieces of DNA (Gietz and Schiestl, 2007).

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 - for each transformation:

2. Add 15 µL of the denatured salmon sperm DNA (2 mg/mL) to ***labeled*** microcentrifuge tubes.

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 the 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 37˚C with shaking for 30-45 min.

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 10) 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

10. Calculate the fraction of cells that were transformed. For each transformation reaction, divide the number of transformed cells (step 8) by the total number of cells in the transformation mixture (step 9). Use the spot plate data from step 9, correcting for dilution, to calculate the total number of cells in the transformation mixture.

11. Transformation efficiencies are usually expressed by the number of cells transformed per µg DNA. Use your Nanodrop data to calculate the concentration of DNA used for the transformation and your data from step 8 to obtain the number of transformed cells.

Chapter 12

Exercise 2 - 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. Place a piece of sterile velveteen with the nap face up on the replica plating block.
3. 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.
4. 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

References

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

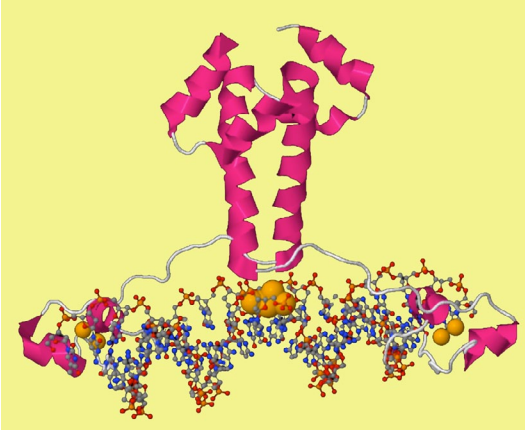
Gelperin, DM, White, MA, Wilkinson *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

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

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 Cys 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

- Understand regulation of the yeast *GAL1* promoter by carbon sources.
- Understand the experimental challenges to preparing cell extracts from yeast.
- Culture yeast with different carbon sources to induce or repress expression from the *GAL1* promoter
- Prepare extracts from transformed yeast grown under repressed and induced conditions

Chapter 13

Over the next few weeks, you'll be analyzing the expression of *S. pombe* and *S. cerevisiae* Met or Cys fusion proteins in your transformed strains as well as untransformed parental *met* or *cys* strains. You have already tested the ability of the overexpression plasmids to complement the *met* or *cys* mutations in your yeast strains. Complementation depends on the presence of functional Met proteins. If you observed a failure to complement the *met* or *cys* deficiencies, this could indicate that proteins were not expressed from the plasmids. Alternatively, the overexpressed proteins may 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 evaluated.

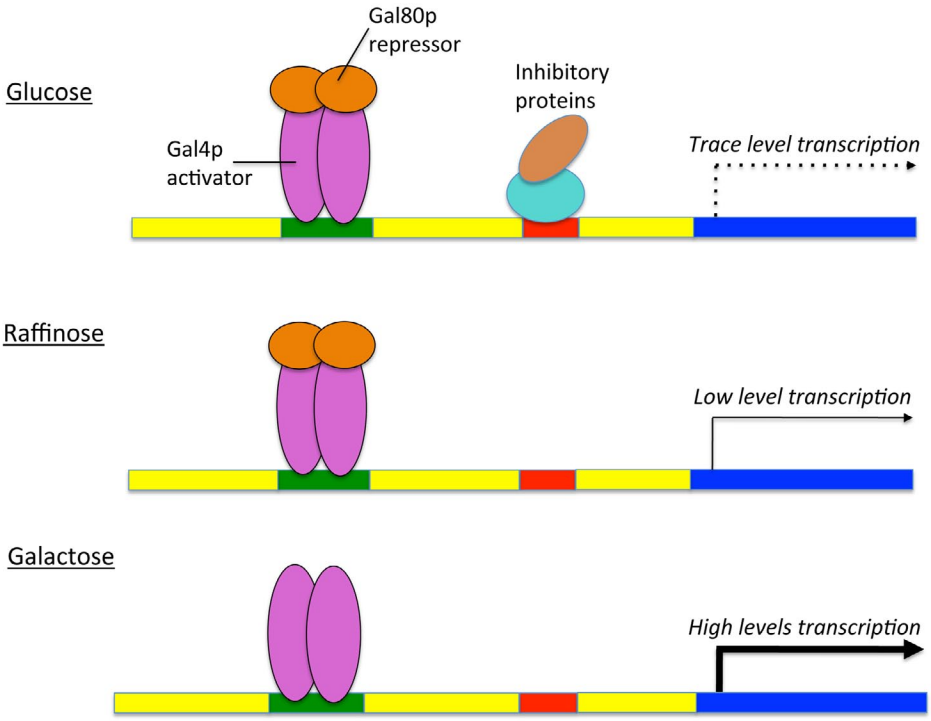
In this experiment, you will prepare extracts from yeast for later experiments (Chapters 14 and 15) in which you will determine if the fusion proteins are being successfully expressed in the transformed yeast strains and how the expression of Met and Cys proteins varies with carbon sources.

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 ultimately convert galactose 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/CYS* genes in our 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



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.

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 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 by the inhibitory Gal80p protein. Gal4p acts as a master regulator of galactose metabolism. In addition to activating *GAL1* transscription, 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.

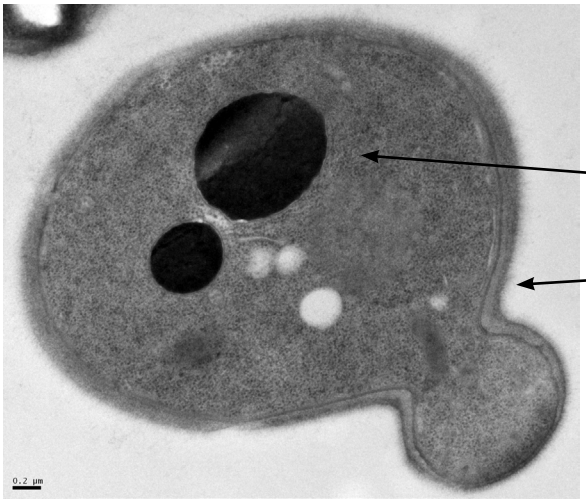
In this lab, you will prepare extracts from transformed *met* or *cys* strains that are growing in either glucose or galactose. In subsequent labs, you will analyze the proteins in these extracts on SDS-PAGE gels (Chapter 14) or western blots (Chapter 15).

Chapter 13

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 many of these experimental barriers.

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 denaturing proteins because of the heat and foaming generated during the processes.



Electron micrograph of an *S. cerevisiae* cell.

Vacuole contains many proteases.

Polysaccharide-based wall surrounds the cell.

Image by Christopher Bruser
Reproduced with permission
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. Detergents can be either nondenaturing or denaturing. Denaturing detergents destroy protein structures by breaking the thousands of

weak bonds that normally hold proteins together. A good example of a denaturing detergent is SDS, or sodium dodecyl sulfate, which we will be using to prepare our cell extracts. By contrast, non-denaturing detergents bind to the surfaces of proteins and consequently preserve enzyme activity.

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) relies on chemical disruption and rapid protein denaturation to denature the yeast proteases. Extracts prepared by this method are suitable for SDS-PAGE and western blot analysis.

Chapter 13

Exercise 1 - Prepare cell cultures for extraction

DAY 1 - First lab session of the week

1. Inoculate 2 mL of YC-URA containing 2% raffinose with a single yeast colony.

2. Repeat step 1 for each of your transformed strains (for a total of 3 cultures).

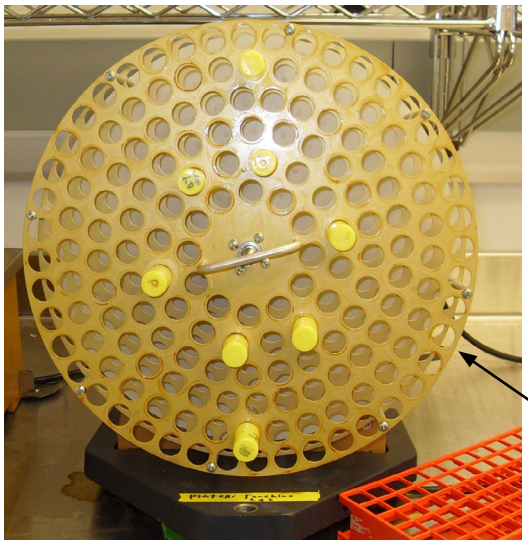
3. Place the cultures on the wheel at 30˚C, ***taking care that tubes are properly balanced.*** Turn the wheel on.

DAY 2 - The day before your second lab session of the week (your off day)

1. Remove your cultures from the wheel.

2. Add 1mL of your Raffinose culture to a tube containing 1 mL YC+Glucose (repression medium). Add the other 1mL of your Raffinose culture to a tube containing 1mL of YC+Galactose (induction medium) ***Be sure that the tubes are labeled with your initials and the sugar source. You should have a total of 6 tubes.***

3. Place the cultures back on the wheel to grow overnight before your class time.



USE THE WHEEL CORRECTLY!

Make sure that wheel is balanced by placing tubes symmetrically with respect to the central axis.

Remember to turn the wheel ON after removing your samples!

On switch is on the right side of the base behind the wheel.

Exercise 2 - Preparing cell extracts

Harvest the cells for extraction

1. Grow the cells under various conditions as instructed.

2. Determine the cell concentration. Vortex the tubbes and transfer 100 µL of each cell culture to 900 µL deionized water and measure the OD_{600} of cultures in the spectrophotometer. Note the values in your lab 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 at top speed- in a microcentrifuge. Decant 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 be present in the cell pellet.

5. 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.

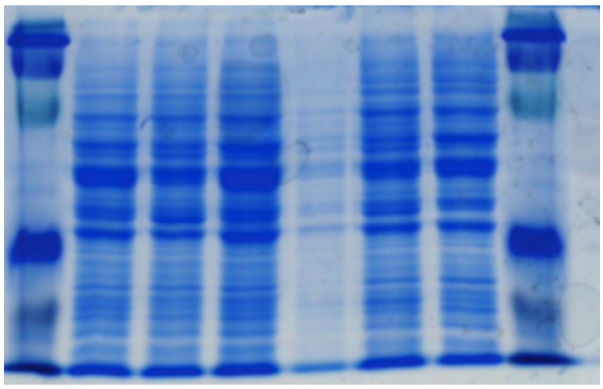
Gelperin, DM, White, MA, Wilkinson *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

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'll use SDS-PAGE to analyze the protein extracts that you prepared from yeast strains overexpressing Met and Cys fusion proteins of LacZ.

Objectives

- Understand the theory governing protein separation by SDS-PAGE
- Learn how to cast and run SDS-PAGE gels
- Understand how stains are used to visualize proteins in gels
- Learn how to 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. You will make two gels in this lab. You will run one gel today and stain it with Brilliant Blue G-250 to visualize the proteins in the extracts. You will store the other gel until the next lab period, where you will use it for a western blot.

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

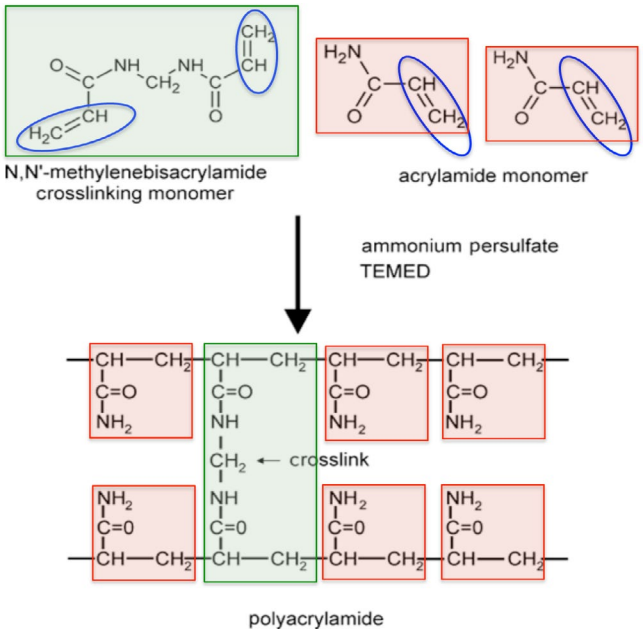
In gel electrophoresis, an electric field is used to move charged molecules through a matrix of a polymerized substance such as agarose or polyacrylamide. The rates at which individual molecules move through the gel depend on the properties of both the separation system and the molecules themselves. 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, i.e. voltage, used for the separation.

You have already used agarose gel electrophoresis to separate DNA molecules. DNA molecules are much larger than protein molecules and they have a uniform charge to mass ratio. Consequently, electrophoretic analysis of DNA molecules is more straightforward than the separation of protein molecules. The agarose gels used to separated DNA molecules have much larger pores than the polyacrylamide gels typically used to separate protein molecules. (However, small DNA molecules, e.g. less than 1000 bp, may be separated in polyacrylamide gels.)

Chemistry of acrylamide polymerization

Proteins are usually separated on polyacrylamide gels 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). Aqueous solutions of APS are unstable, but they can be stored for about a week in the refrigerator or for several months in the freezer without losing potency.



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

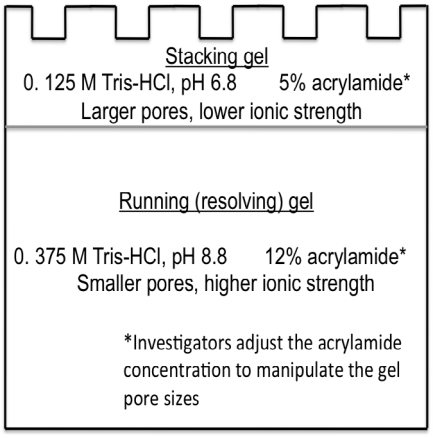
Proteins are denatured prior to electrophoresis

Compared to DNA molecules, proteins are structurally very diverse. Crystallographers have described over 1000 structurally distinct protein folds. Proteins also show tremendous variation in their amino acid compositions and in the distribution of amino acids in the folded structure, features with important implications for electrophoresis. Because the hydrophobic effect is an important contributor to protein folding, hydrophobic amino acids tend to be more frequently found in the interior of proteins than on their surfaces. The surfaces of proteins have a higher frequency of polar and charged amino acids than the interior of the protein, because these side chains are compatible with the aqueous cytoplasm. Even so, protein surfaces are mosaics with respect to the distribution of charged, polar and hydrophobic groups. Consequently, folded proteins lack the uniform charge to mass ratio that characterizes DNA molecules.

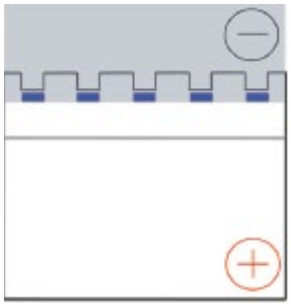
To resolve the proteins in a cell extract, investigators must reduce the structural complexity of proteins in their samples and impart a uniform charge to mass ratio to the proteins. In SDS-PAGE, the solution is to denature the proteins by boiling them with the anionic detergent, sodium dodecyl sulfate (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 binds to the many hydrophobic groups in proteins, reducing the protein to a random coil, coated with negatively charges 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.

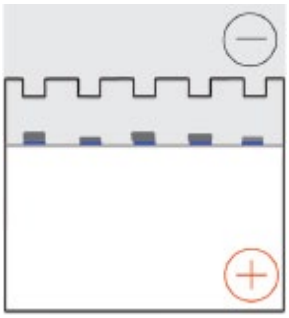
Discontinuities between the stacking and running gels underlie the resolving power of the SDS-PAGE gels



The Laemmli SDS-PAGE system can be considered a 3-component system. The stacking and running (resolving) gels have different pore sizes, ionic strength and pH. 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.

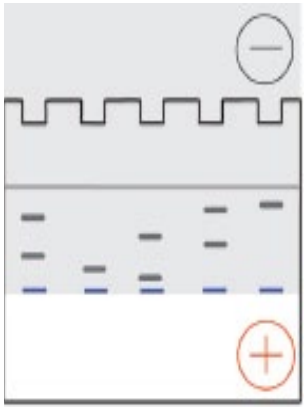


The sample buffer used for SDS-PAGE contains a tracking dye, bromophenol blue (BPB), which will migrate with the leading edge of the proteins being separated on the gel. The sample buffer also contains glycerol, which allows the protein samples to settle into the bottom of the gel wells. The gel is vertically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

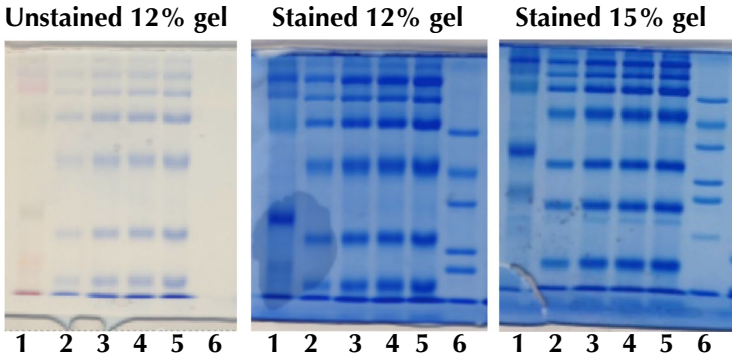


Proteins are visualized with stains.

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 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. The most commonly used dyes are the closely related Brilliant Blue R-250 and G-250 dyes, which bind proteins nonspecifically through a large number of ionic and Van der Waals interactions. Although Brilliant Blue R-250 staining is slightly more sensitive than G-250 for detecting proteins, the G-250 staining procedure is more rapid and does not require destaining with organic solvents. We will use the commercial Simply Blue™ reagent, which contains a colloidal suspension of Brilliant Blue G-250, to stain the gels. 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 a quantitative procedure, because the intensity of a stained band is directly proportional to the amount of protein in a band.

The figure below illustrates these points. 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 wen calculating MWs of unknown 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 loaded in that lane.



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 are 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.

When analyzing experimental data, remember to consider any 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 an 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 adds only 33 amino acids, including a V5 epitope tag and a $(His)_6$ purification tag to the C-termini of overexpressed proteins. Together, these amino acids add 3710 to the MW of the protein.

Casting SDS-PAGE gels

These instructions are designed for constructing two 12% SDS-PAGE gels with the BioRad Mini Protean system. We will need only one gel in this lab, but the extra gel provides insurance against leaks! If neither gel leaks, you will use the second gel for western blots in the next lab.

Assemble the gel casting apparatus

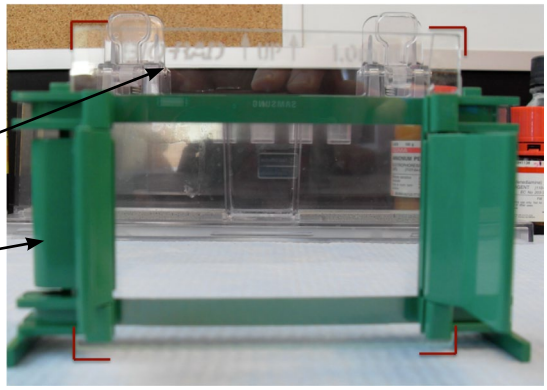
1. Assemble the components that you will need for casting the gel: a tall glass plate with attached 1 mm spacers, a small glass plate, a green casting frame and a casting stand.
2. Place the green casting frame on the bench with the green "feet"resting firmly against the bench and the clamps open (perpendicular 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

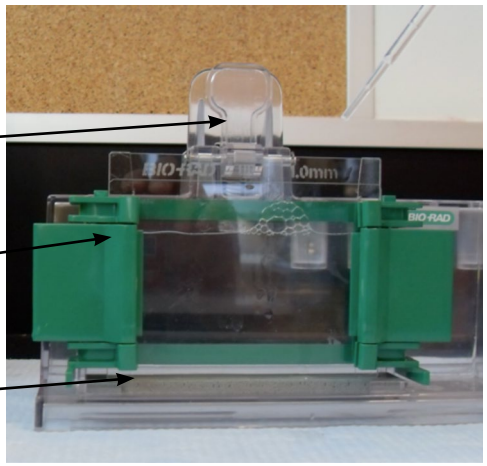
1. Label two 15 mL conical tubes "Resolving gel" and "Stacking gel".
2. Prepare the resolving gel. Mix the acrylamide solution, **pH 8.8** Tris buffer and water, as shown in the chart above. Mix the ingredients gently, trying not to introduce air. Oxygen inhibits polymerization of acrylamide gels.
3. 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 resolving gel. 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.
5. Remove the comb when the gel has polymerized (10-15 minutes). You will be able to see faint schlieren lines (indicators of a change in refractive index) at the top of the spacer gel when polymerization is complete. You can also check the leftover solution in the 15 mL tube for polymerization.

Saving 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 gel in a wet paper towel. Then wrap the gel in plastic wrap to be used in a later lab.

Chapter 14

Running SDS-PAGE gels

Set up the electrophoresis apparatus

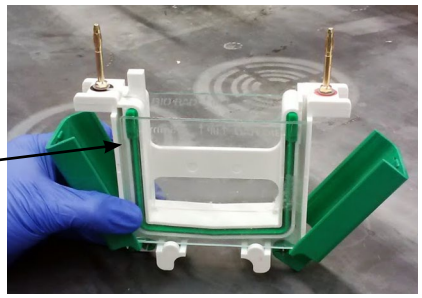
1. After the stacking gel polymerizes, remove the gel cassette sandwich from the casting stand.
2. Carefully remove the comb from the spacer gel.
3. Remove the casting frame from the gel cassette sandwich and place the sandwich against the gasket on one side of the electrode assembly, with the short 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. If you wish, place the sample loading guide in between the two gels in the electrode assembly before loading your samples.
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.
3. NOTE: Be sure to record the order of samples loaded onto the gel.
4. 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).
5. 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. 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 *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop* **19:** 2816-2826.

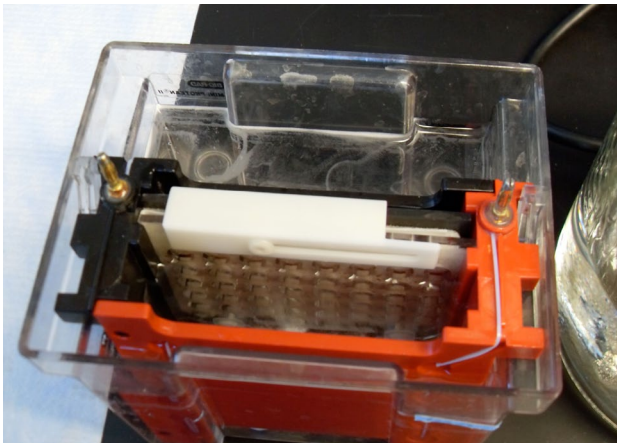
Laemmli, UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680–685.

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'll learn about the different kinds of antibodies used in western blots. You will use western blots to analyze Met and Cys protein expression in your transformed yeast strains.

Objectives

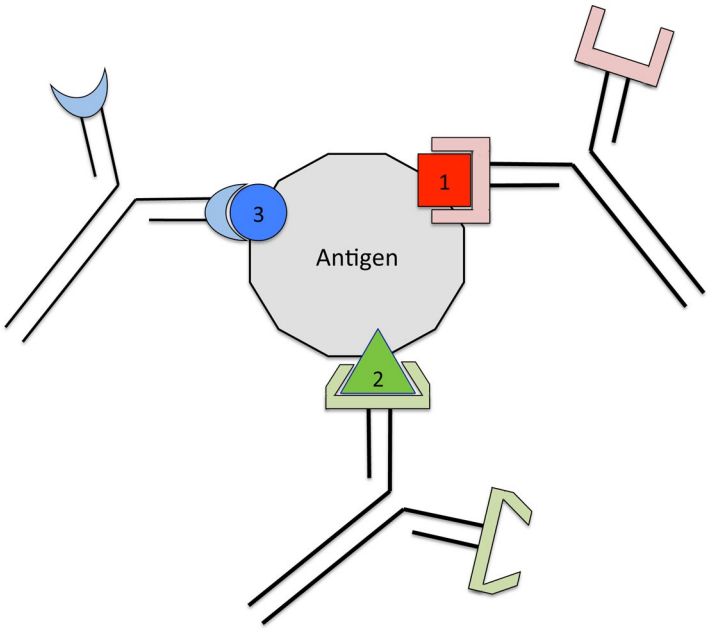
- Understand the structural features of antibody molecules.
- Understand how monoclonal and polyclonal antibodies are produced.
- Understand how epitope tags are used in molecular biology
- Understand how primary and secondary antibodies are used to detect proteins in western blots.
- Use western blots to analyze protein expression in yeast transformed with overexpression plasmids.

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.

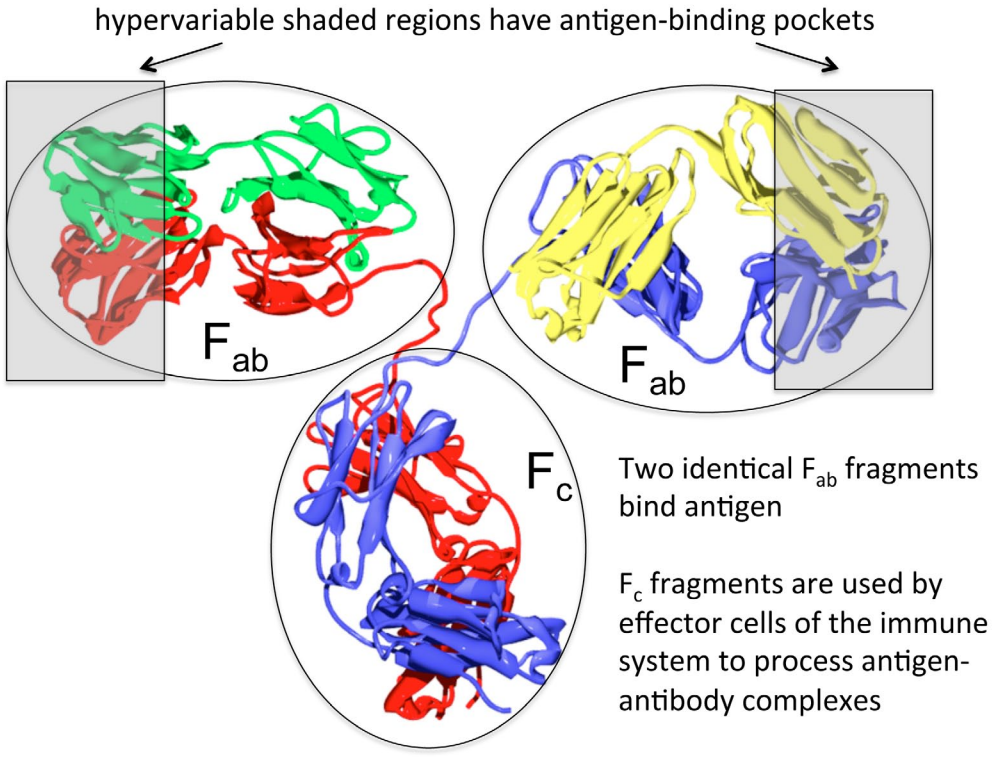


Antibodies binding to an antigen. An antigen with three different epitopes on its surface is bound by three different antibody molecules, each of which binds a single epitope with high affinity.

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.

The antibodies that we will be using in this class are IgGs. X-ray crystallography has given us a detailed view of the IgG structure. IgGs are tetramers consisting of two identical light chains and two identical gamma heavy chains. F_{ab} domains are formed from a light chain and the N-terminal portion of a heavy chain. The N-termini of the heavy and light chains are hypervariable regions that have been generated by DNA rearrangements that occur during lymphocyte maturation (see below). These hypervariable regions form the antigen-binding sites on the antibody. The F_c domain is formed by the C-terminal portions of the two heavy chains. The same F_c domain is found in all immunoglobins of a particular class of antibody.



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

Chapter 15

Antibodies are produced by lymphocytes

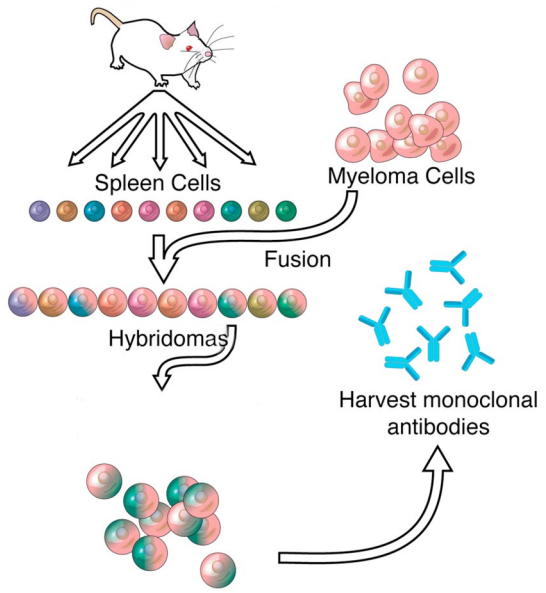
In the initial stages of the immune response, small numbers of immature B lymphocytes are able to bind foreign antigen molecules weakly via the antibodies expressed on their surfaces. Antigen binding stimulates the lymphocytes to proliferate and to differentiate into mature, antibody-secreting lymphocytes. An amazing transformation occurs 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 are known as plasma cells, once they have survived the selection process. ***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

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, serum collected from an immunized animal contains a mixture of antibodies that bind an antigen at distinct sites. These antibody-binding sites on antigens are known as epitopes.

In contrast to polyclonal antisera, which are harvested from the serum of living animals, monoclonal antibodies are produced from cultured hybridoma cells. Hybridoma cells are generated by fusing lymphocytes from an immunized animal, most commonly a mouse, with myeloma cells, cancerous cells that are capable of dividing indefinitely in culture. The figure on the opposite page outlines the procedures used to generate antibody-secreting hybridoma cell lines. Lymphocytes are isolated from the spleens of an immunized mouse and fused with myeloma cells. Standard culture techniques are used to isolate individual hybridoma cell lines from the fusions. Each hybridoma cell line secretes a unique antibody that recognizes a single epitope on an antigen.

Hybridoma technology has revolutionized biomedical research since its description (Kohler & Milstein, 1975), 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.



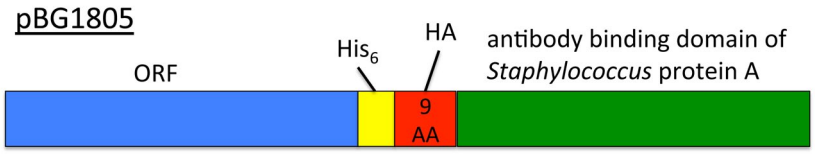
Construction of hybridoma cell lines that secrete monoclonal antibodies.

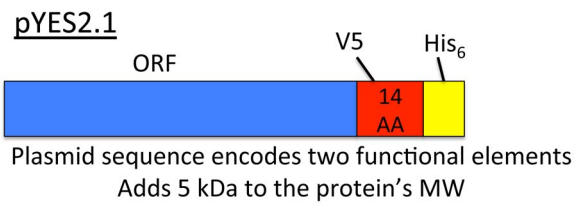
Adapted from http://en.wikipedia.org/wiki/Monoclonal_antibody#mediaviewer/File:Monoclonals.png

Plasmid-encoded proteins have C-terminal tags

In this lab, we will be using antibodies to detect Met and Cys proteins expressed in transformed *S. cerevisiae*. The proteins that will be expressed are not the native proteins, but fusion proteins. The pBG1805 (Gelperin *et al.*, 2005) and pYES2.1 plasmids were designed to be used for multiple purposes. Molecular biologists designed the plasmids by adding DNA sequences that encode functional elements in the same reading frame as the yeast ORFs. Thus, the expressed proteins will have C-terminal tags that can interact with a variety of reagents.

Both plasmids encode short epitope tags that can be detected with antibodies on western blots. These epitopes correspond to highly 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). This HA-tag, can be detected with antibodies on western blots.





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.

Chapter 15

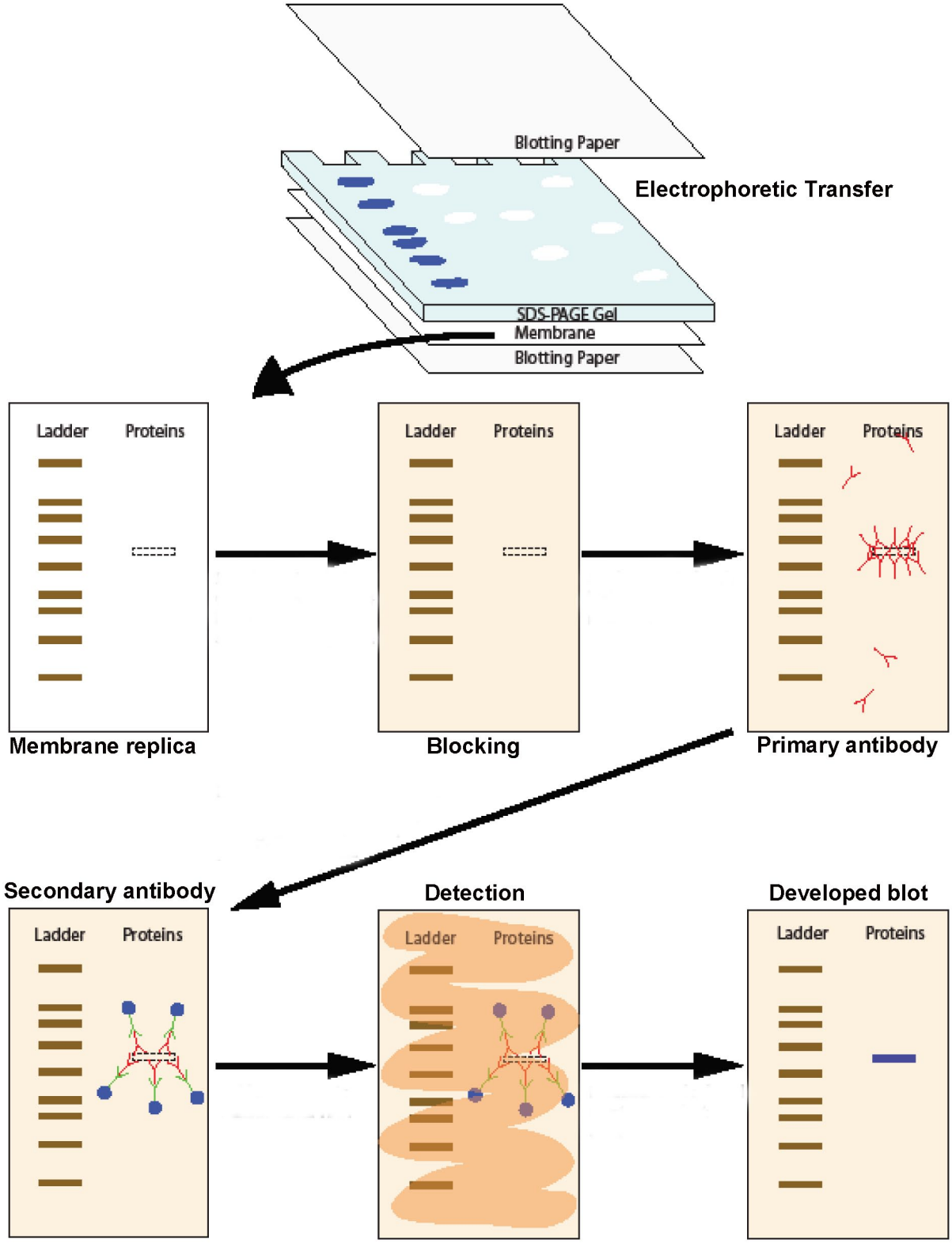
We will not be using an anti-HA antibody to detect proteins expressed from pBG1805-based plasmids. Instead, we will make use of the large C-terminal tag that encodes the ZZ domain of *Staphylococcus aureus* protein A. *S. aureus* uses protein A to evade detection by the immune system, contributing to its pathogenicity. Protein A is a membrane protein with multiple extracellular ZZ domains, each of which is able to bind the F_c domain of an IgG. The secondary antibody that we are using for the blots was produced in rabbits, and its F_c domain binds tightly to protein A.

Both plasmids also encode a third 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. His_6 can be used as an epitope tag, but not very effectively. Anti His_6- antibodies generally have a much lower affinity for their target than anti-HA or anti-V5 antibodies.

Western blot overview

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 specific for the primary antibody
- Detection of the primary antibody-secondary antibody complex


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

Chapter 15

Western blots involve many steps

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 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's 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

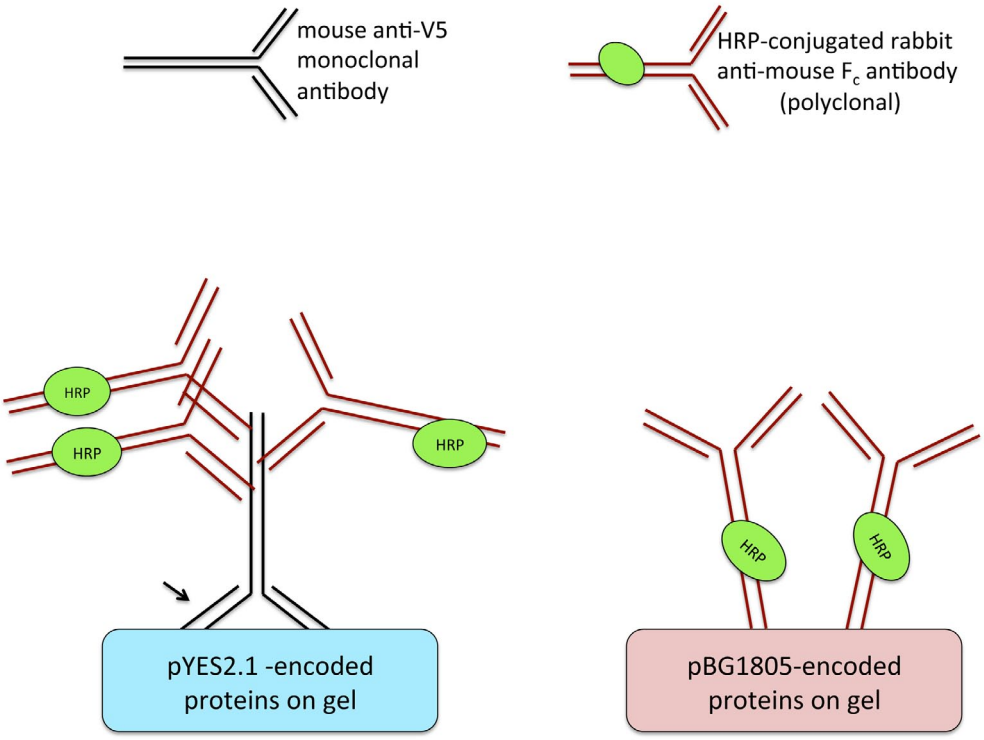
Either polyclonal or monoclonal antibodies can be used as the primary antibody on western blots. Antibodies can be directed toward a naturally-occurring protein or toward an epitope attached to an overexpressed protein (as we are doing). Increasingly, researchers are using epitope-tagged proteins in 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. The primary antibody will not react with the HA-tagged proteins encoded by the pBG1805-based plasmids. Instead, these latter proteins will be detected with the secondary

antibody, which will react with the protein A ZZ domains encoded by the plasmid. (Note: *S. aureus* protein A binds tight to rabbit IgGs, but not to mouse IgGs.)

Secondary antibody binding

Polyclonal antisera are commonly used as the source of secondary antibodies in western blotting. The antisera are prepared by injecting a second species with F_c fragments of IgGs from the species used to generate the primary antibody. Antisera designed to be used on western blots are then covalently linked to a reagent that aids in visualizing bound antibody. In some cases, secondary antibodies are bound to a fluorescent dye, so that blots can be analyzed directly with a fluorescent scanner. In other cases, the antibodies are linked to an enzyme, such as horseradish peroxidase (HRP) or alkaline phosphatase, that can be used for histochemical detection of bound antibody in an additional step. The secondary antibody-binding step plays an important role in amplifying the signal on western blots. Because the antibody is polyclonal, secondary antibodies will bind to multiple sites in the F_c fragment of the primary antibody. If an enzyme is attached to the secondary antibody, the histochemical detection step will provide additional amplification of the signal.

The secondary antibody in our experiment is a rabbit polyclonal antibody prepared against the F_C domains of mouse IgGs. The antibody will bind both the F_C domains of the mouse anti-V5 antibodies bound to the pYES2.1-encoded proteins and the protein A ZZ domains in pBG1805-encoded proteins. The figure below summarizes the binding of the primary and secondary antibodies in our western blots.



Immunochemical detection 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

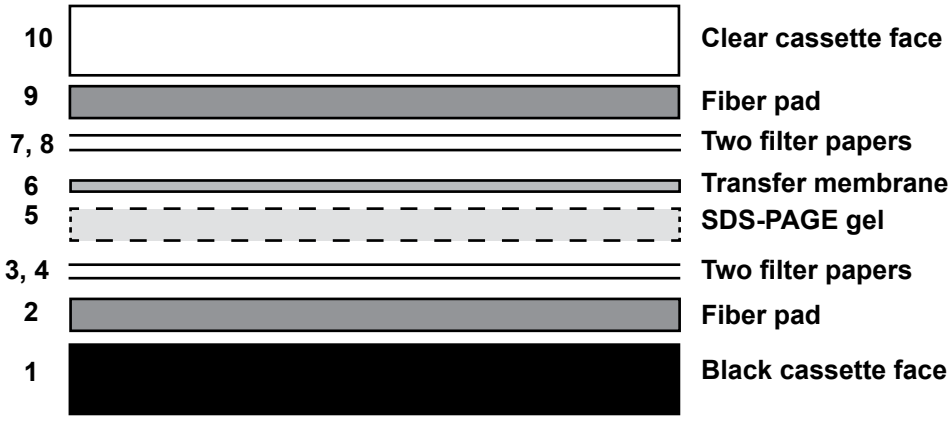
NOTE: DO NOT touch transfer membranes with your fingers. Wear gloves and use filter forceps when you handle transfer membranes.

1. While the gel is running, 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. Fill the buffer tank about half way to the top with transfer buffer. Wet the fiber sponges and filter papers for the transfer cassette by placing them into the transfer buffer in the tank. (Transfer buffer contains Tris and glycine, as well as 10% methanol. Unlike SDS-PAGE running buffer, the transfer buffer does not contain SDS.)

2. 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.)
3. Assemble the transfer cassette as shown be;pw. Be sure that all parts of the transfer "sandwich" remain moist at all times.

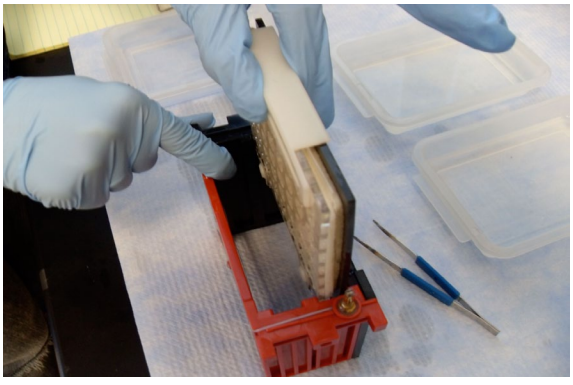


Assembly of the transfer cassette.

- Place a wet fiber pad on top of the black cassette face.
- Add two pieces of filter paper.
- Position the gel 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 blot 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 paper(s) and the fiber pad.
- Fold the clear cassette face 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), the electrophoresis tank should contain an ice pack.
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. Run the transfer at 100V 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 the TBS-T into the waste container.
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 (Higgins 377). 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. Add enough TBS-T to cover the blot. SAVE the plasticwrap! You will need it to cover the container again!
3. Pour enough TBS-T solution 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. Incubate the membrane with 15 ml of secondary antibody solution 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

Stain all proteins on the transfer membrane

In order to verify your proteins have transferred properly to the PVDF membrane, you will stain the proteins on the membrane with a permanent stain, Amido Black (Naphthol Blue Black). Amido Black will non-specifically stain all proteins present on the membrane. Note that once a membrane is stained with Amido Black, it can no longer be probed with an antibody.

As with all stains, wear disposable gloves when working with Amido Black.

1. Wet your membrane with methanol and then with deionized water for 2-3 minutes.
2. Pour off the dH2O and add enough Amido Black stain (0.1% Amido Black in 50% methanol-10% acetic acid) to cover your membrane.
3. Cover your tray with plastic wrap and stain your membrane by shaking for 5 minutes at room temperature.
4. Pour your staining solution back into the stock container (the stain can be used many times), and add enough destaining solution (20% methanol-7.5% acetic acid) to cover your membrane.
5. Allow the membrane to destain for 5 minutes, shaking at room temperature, before pouring the destaining solution into a waste container and adding another aliquot of destaining solution.
6. Destain for an additional 5 minutes or until bands become apparent.
7. Discard the destain solution in the waste container and rinse the membrane with deionized water for 1 minute.
8. Take a picture of the membrane while it is still wet.

References

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 an&d 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.

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.

Glossary

bp
base pair; lengths of DNA molecules are expressed in base pairs

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

locus
position on a chromosome where a gene is located

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 fragment(s)
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