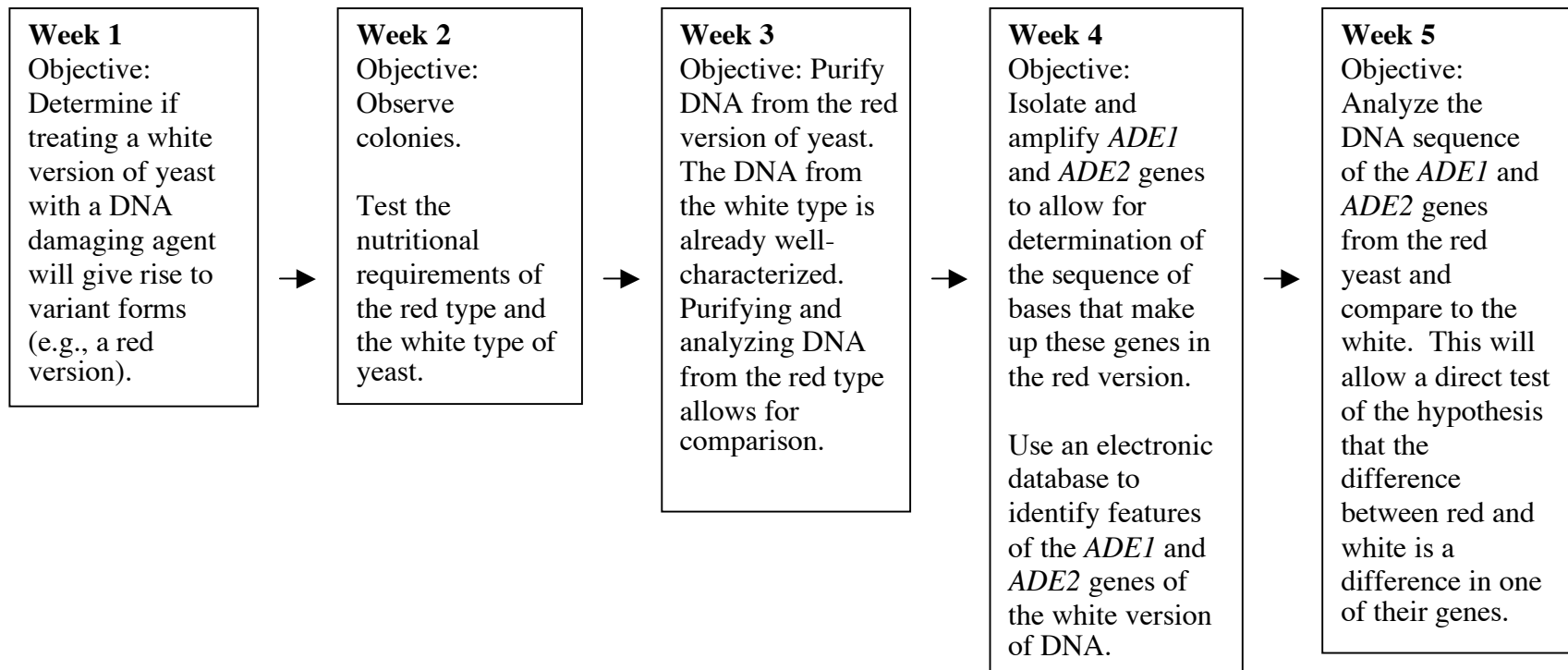


# Table of Contents

Flowchart of the FGPB project.....	2
Objectives of the FGPB project.....	3
Adenine synthesis pathway, simplified (generic names).....	4
Adenine synthesis pathway (specific names).....	5
Purine metabolism pathway.....	6
Can ultraviolet light change the appearance of yeast?.....	7
Characteristics of red yeast: Can they be obtained by treatment of white yeast with UV light? Do they have special nutritional requirements? ....	13
Preparation of DNA from a red yeast strain.....	19
DNA from the red yeast strain: Electrophoresis and PCR.....	23
Analysis of the base sequence in the DNA of the <i>ADE1</i> and <i>ADE2</i> genes..	35

# A flowchart of the From Genes to Proteins to Behavior Project



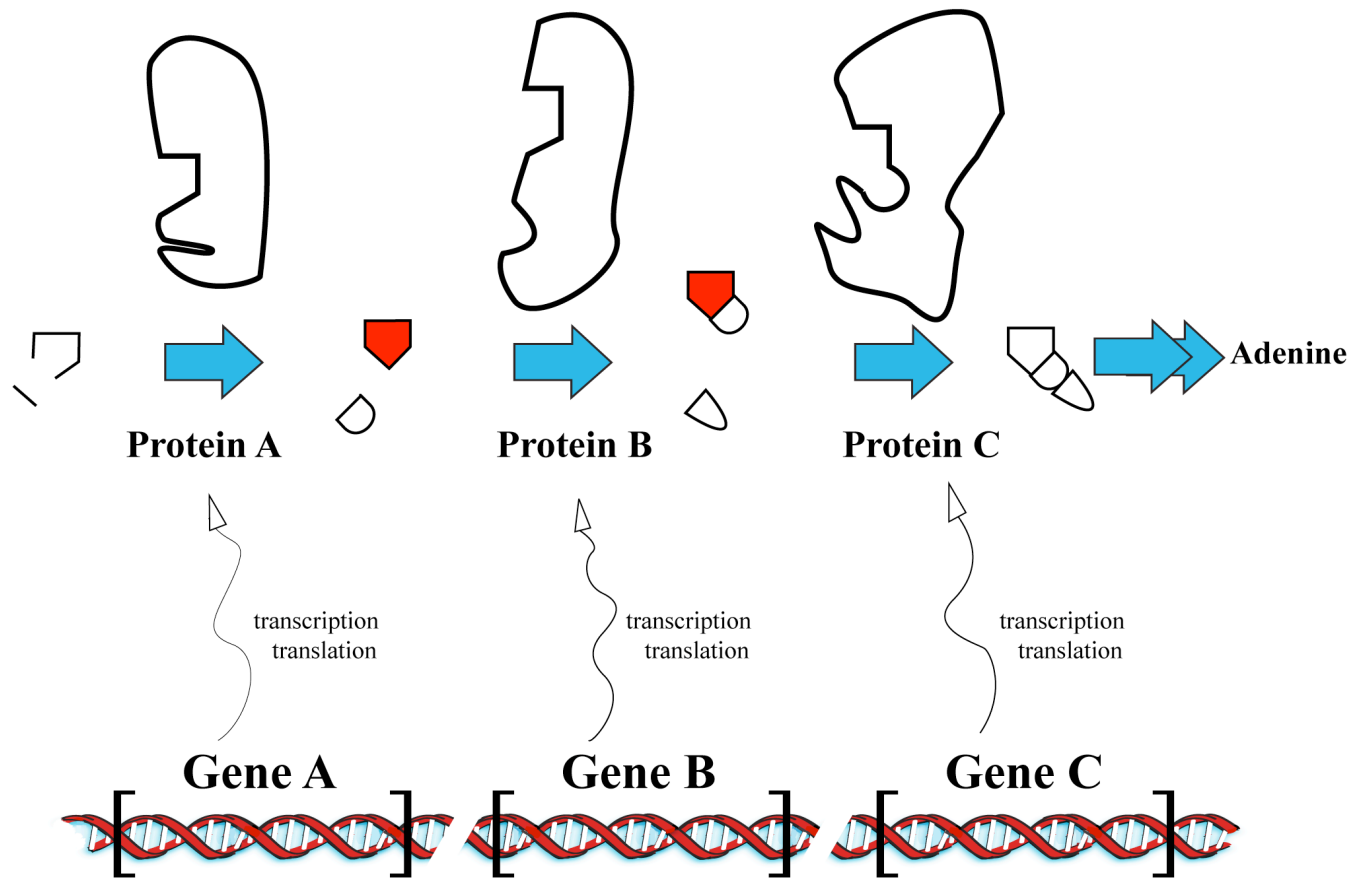
## **Learning Objectives for the From Genes to Proteins to Behavior Project**

The next five weeks during the laboratory period will be a series of related experiences that mimics the process of investigational science and is aimed at accomplishing the following learning objectives.





- (1) To understand the Central Dogma---how information "flows" from DNA to RNA to protein
- (2) To understand how the sequence of nucleotides in a gene provides the information to make a protein
- (3) To understand how the sequence of amino acids of a protein determines the structure of a protein and to recognize the relationship between the structure of a protein and its function
- (4) To understand mutation, its random nature, and the relationship between mutation and variation
- (5) To understand how a series of chemical reactions, each one catalyzed by a specific enzyme, can result in the production of essential "building blocks" of the cell and to understand the consequences of an inability to carry out one of the chemical reactions in the series

During these five weeks some of these learning objectives will also be covered in the lecture period. Other objectives listed above, however, will be introduced during the laboratory period and will then be covered again in the second half of the semester.

## Three steps in the making of adenine inside a yeast cell



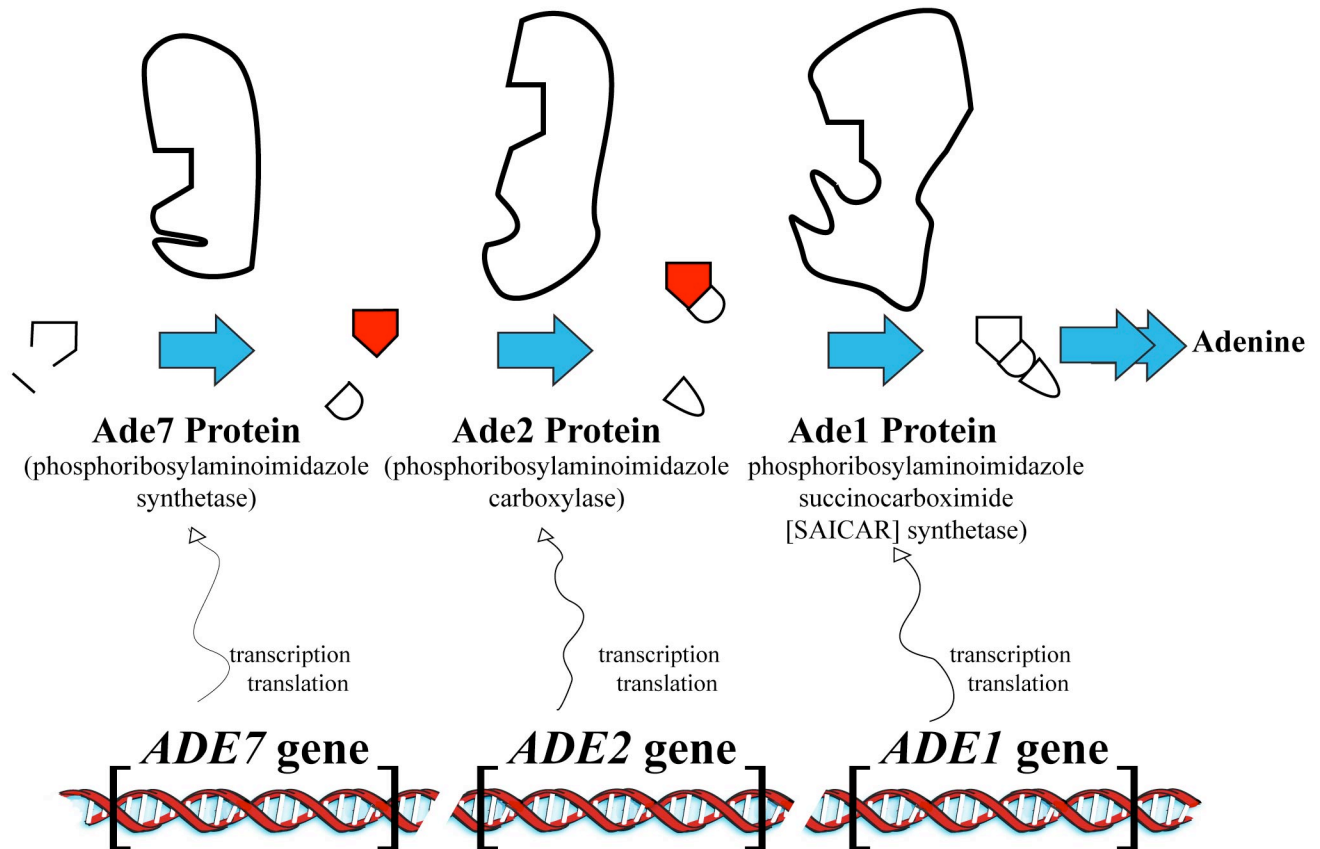
**Legend-** these geometric shapes represent small chemicals that are added to or rearranged in the process of making adenine

-  Precursor (colorless)
-  Intermediate #1 (red)  
(Phosphoribosylaminoimidazole)
-  Intermediate #2 (red)  
(Phosphoribosylaminoimidazole  
carboxylate)
-  Intermediate #3 (colorless)  
(Phosphoribosylaminoimidazole-  
succinocarboxamide)





### Points of Note

- (1) The genes and their corresponding enzyme proteins were arbitrarily assigned a letter to differentiate them from one another. The genes are not adjacent to one another on the DNA.
- (2) The gene encodes the instructions to make the corresponding protein (enzyme).
- (3) The enzyme (shown as a large irregular shape) is required for the chemical reaction depicted.
- (4) A block at any step in the pathway will lead to the accumulation of the molecules just prior to the step that is blocked.
- (5) There are steps prior to the first one shown here and after the last one shown here. For simplicity these steps are not shown.

## Three steps in the making of adenine inside a yeast cell



**Legend-** these geometric shapes represent small chemicals that are added to or rearranged in the process of making adenine

-  Precursor (colorless)
-  Intermediate #1 (red)  
(Phosphoribosylaminoimidazole)
-  Intermediate #2 (red)  
(Phosphoribosylaminoimidazole carboxylate)
-  Intermediate #3 (colorless)  
(Phosphoribosylaminoimidazole-succinocarboxamide)

### Points of Note

- (1) The genes were numbered in the order of their discovery. They are not adjacent to one another.
- (2) The gene encodes the instructions to make the corresponding protein (enzyme).
- (3) The enzyme (shown as a large irregular shape) is required for the chemical reaction depicted.
- (4) A block at any step in the pathway will lead to the accumulation of the molecules just prior to the step that is blocked.
- (5) There are steps prior to the first one shown here and after the last one shown here. For simplicity these steps are not shown.



## Lab, Week 1

### From Gene to Protein to Behavior

## CAN ULTRAVIOLET LIGHT CHANGE THE APPEARANCE OF YEAST?

In the From Gene to Protein to Behavior Project you will be studying baker's yeast strains that differ in appearance. Yeast, schmeast, you may say, why should I care about different kinds of yeast? However, the presence of different types of yeast leads to a fundamental question of biology: What are the underlying causes of biological variation? Why are some yeast red and others white? Why do some of us have black hair and others blonde or red or brown? In the laboratory sessions of weeks 1-5 our overall goal will be to determine the cause of the variation we see in red and white yeast. The insights obtained by this study of yeast can be applied to understanding variation in other organisms, such as humans.

Specifically, in this week's laboratory you will test whether one type of yeast can be changed into another by exposure to UV light. You will examine the effect of different doses of ultraviolet (UV) light on *S. cerevisiae* (yeast) cells and the resulting colonies. One macromolecule that is known to undergo change after UV light exposure is DNA. Can change in the DNA result in a change in the appearance of the resulting yeast colony? That is, starting with a yeast that forms white, smooth, round colonies and exposing it to UV light, will we see any colonies with altered color or morphology?

To most fully measure the effects of UV light you will subject the yeast to increasing doses of UV light. Not only can we answer the question of whether one type of yeast can be generated from another but we can answer more general questions as well. For example: What happens when macromolecules in a cell are subjected to chemical change beyond what normally occurs inside of the cell? Does the cell grow faster than in the absence of the chemical change? Does the cell die as the result of the chemical change? Does the cell acquire a different property as the result of the chemical change? Do all cells react the same way?

#### Some useful vocabulary:

- sterile:** adj. Containing no living organisms.
- medium:** n. What bacteria or yeast grow in or on. Can be solid or liquid.
- culture:** n. Yeast (or bacteria or other microorganism) growing in or on a medium.
- colony:** n. Usually found on a petri plate. A visible "pile" of yeast (or other microorganism) that are all descendants of a single yeast cell.
- strain:** n. A particular variety of a microorganism. Different strains are more similar than are different species. This term is functionally equivalent with the term "breed" (used in animal husbandry) or "variety" (used in agriculture). In this lab you will be working with a white strain of yeast to see if you can alter it to make a different strain with a new appearance. Do not confuse this term with the word "strand" which is a long thin object such as a hair or one half of a DNA helix.
- aliquot:** v. To subdivide a liquid sample into smaller portions. n. A portion of a larger sample.

## Methods

- Pipetting: This lab will require precise measurement of liquids. We will use pipets to measure the liquids. Your instructors will demonstrate use of the pipets. Practice using the pipets before you begin the rest of the lab. We will have a brief exercise to help you practice.
- Sterile technique: Remember, you are trying to see if you can change one type of yeast into another by using UV light. If next week you see any colonies that appear different from the colonies you started with, you want to be assured that these colonies were derived from the yeast in your original culture and are not some random cells from the air that just happened to land in the tubes or on your plates. In other words, it is essential that your yeast cultures be kept free from other microorganisms that are found in the air, on your hands, or elsewhere in environment. You will use **sterile technique** to keep the cultures from becoming contaminated. There are several different components to sterile technique; your instructor will demonstrate those that are necessary for today's lab. You will try some others next week.
- Spread plates: Next week you will want to be able to see individual colonies of the yeast, in order to see if any of them look different from those of the original strain. A spread plate is often used to separate the cells in a liquid culture from each other by spreading them evenly over the surface of a petri plate, so that each cell can give rise to a separate colony. The technique will be demonstrated by the instructor and instructions are also provided below.

Each lab group will prepare four identical spread plates. Obtain four YPD plates. These YPD plates are a nutrient-rich medium for growing yeast.

Obtain a liquid culture of yeast cells marked "HAO". HAO is a strain of white-colored yeast. Check the label carefully – it is critical to use the correct culture! **Just before using the culture, vortex it to mix.**

*The yeast cells are dense and will settle to the bottom of the tube. Thus, to ensure pipetting a homogeneous solution of cells, vortex immediately prior to **each** use.*

Take a 0.1 ml aliquot of the culture using a sterile pipet and sterile technique. Put it in the center of a YPD plate. Make sure your work area is free of papers and other flammable material. Dip a spreader in a small amount of ethanol and pass it **briefly** through the flame of a bunsen burner. A small amount of the flaming ethanol may drip on the lab bench, which is why you want the area clear. Wait until the alcohol burns off the spreader-- the spreader is now sterile. Cool the spreader on the YPD agar surface and then use it to push around the drop of culture on the dish. Try to spread the culture as evenly as possible across the plate, spreading until the plate is dry. Repeat the procedure, including the vortexing of the liquid culture, for the remaining three YPD plates.

- Exposure of the cells to ultraviolet light: You will expose each of the four plates to a different amount of ultraviolet light. UV light can be measured in a unit known as a microJoule. Expose one plate to 3000 microJoules, one to 6000 microJoules, one to 12000 microJoules and one to 24000 microJoules. Label the bottom of each of the four plates with a marker to indicate the amount of UV exposure the plate will receive and the lab group's initials.



*Exposure to the ultraviolet light will be done in an apparatus that looks like a microwave oven. This can be programmed to give a known dose of ultraviolet light.*

The apparatus that delivers the UV light has an LED that can be set. The readout represents the number of microJoules X 100. Therefore if you want to deliver 3000 microJoules the display should read 30. Also make sure the apparatus is on the “Energy” mode and not the “Time” mode. After setting the apparatus, place the appropriate YPD plate in the apparatus and remove the lid.

***UV light will not penetrate the plastic lid!***

Press the “Start” button. It should take between 5 and 30 seconds for the UV light to be delivered.

- **Dilute culture for untreated control plate:** As a control for your experiment, you should make a plate of yeast cells that is not exposed to UV light. The problem with this idea is that it would be useful to be able to count the number of colonies on this plate next week, but the culture we gave you is concentrated enough that there will be too many colonies to count. Therefore, you will dilute the culture before using it in the next step to make your control plate. Each person should make his or her own dilution of the culture. Dilute the cell culture as follows: Mix the culture thoroughly on a vortex mixer just before you begin the dilution process. Using sterile technique, add 1 ml of culture to a tube that has 9 milliliters of sterile water. Vortex to mix the dilution. Be sure to mix thoroughly.

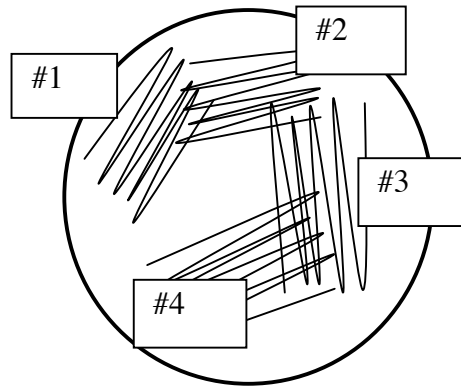
*Don't forget to use sterile technique. Flame the lips of the tubes before and after transfer and don't touch the pipet tip to anything that is not sterile, such as your fingers or the benchtop.*

- **Culture the cells in the absence of UV light:** As a control, you will use your diluted culture to examine the growth of the yeast culture in the absence of the UV light. Each person in the lab group should make a control plate. To do this, obtain one YPD plate for each member of your lab group. Label the bottom of the plate to indicate your lab group and the lack of UV treatment. Mix your diluted culture just before use. Using sterile technique, immediately pipet 0.1 ml of this diluted culture onto your new YPD plate. Use the same technique as described above to spread the culture.
- **Incubate cultures at 30°C:** Place each of your five YPD plates face down in a receptacle at the front of the classroom. It will be convenient if you tape the bundle of plates together. The instructor will place the receptacle at 30°C, an optimal temperature for the growth of this organism.
- **Streak plate:** Next week you will be asked to separate cells from any interesting colonies you obtain from the other colonies surrounding them on the plate. To do this, a technique called a streak plate will be used. This week you'll do a practice streak plate so you can see how this technique is done. Your instructor will demonstrate it for you.

Obtain a liquid yeast culture that contains two types of yeast (it should be labeled “mixed yeast culture”). Flame a loop until red-hot to sterilize it. Open culture using sterile technique (as demonstrated), cap and loop in one hand, tube in other. Briefly flame mouth of tube, and remove a loopful of culture. Replace cap. Streak out on a YPD plate in the four streak pattern (see picture),

flaming loop between streaks 1 and 2, 2 and 3. Streak gently-- try not to break through the surface of the agar with the loop. Incubate the plate at room temperature.

*You should separate the two types of yeast by streaking them around the plate.*

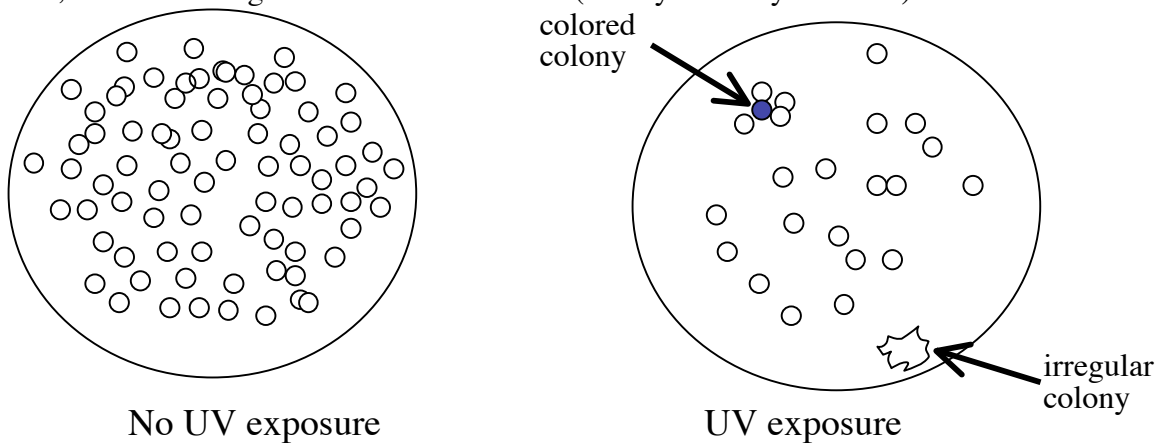


## Lab Assignment

**Due at the start of lab in week 2**

**8 points**

- (2 pts) 1. The following observation is made in comparing a plate that was exposed to UV light to one that was not. Explain how you would experimentally determine if the colored colony resulted, not from UV treatment, but from being in a crowded condition (nearby so many colonies).



- (6 pts) 2. A biologist wanted to understand the quantitative relationship between the amount of UV light (in microJoules) used to expose a culture of yeast cells and the percentage of yeast cells that survived the treatment (where 100% is the number of cells surviving after no UV treatment). She obtained the following result.

<u>Amt of UV light (mJ)</u>	<u>Number of cells surviving</u>	<u>% of cells surviving</u>
0	1,540	100
300	1310	
600		65
900	540	
1200		9
1500	8	

- (a) Fill in the blank spaces in the table.
- (b) Use the computer program Microsoft Excel (or your favorite graphing software) to graph the above results in a scatter graph. The plot should show the percentage of cells surviving as a function of UV dose. Refer to the section of this manual named “Using Microsoft Excel to Create Graphs” for general instructions.

Label the X and Y axes and place a title at the top of the graph. The title should be meaningful but not verbose. For instance, a title of “Yeast experiment” will not have any meaning for anybody who is not in Biology 131 in the spring semester of 2004. On the other hand, “A

study to determine the effects that ultraviolet light has on the ability of a fungus known as *Saccharomyces cerevisiae* to survive on YPD growth medium when incubated at 30°C” is an example of a title that is not succinct enough to be most effective.

(c) What percentage of surviving cells would be expected if 450 microJoules were used? (2 points)

## Lab, Week 2

### From Gene to Protein to Behavior

#### CHARACTERISTICS OF RED YEAST:

CAN THEY BE OBTAINED BY TREATMENT OF WHITE YEAST WITH UV LIGHT?

DO THEY HAVE SPECIAL NUTRITIONAL REQUIREMENTS?

Last week you treated white yeast cells with UV light. It was suggested that UV light can alter DNA. This week you will examine your UV-treated plates to see if any of the yeast cells changed their characteristics as a result of the UV treatment. We are particularly interested in any yeast that changed in color from white to red. *Our working hypothesis will be that these and other changes occurred because UV light altered the DNA of the white yeast.* The rest of the labs in the From Gene to Protein to Behavior project are designed to test this hypothesis.

Our focus this week will be in identifying and purifying colonies that are altered in appearance by the UV treatment. You should also notice more general effects of UV light on the yeast culture – do you notice any overall trends that are related to the dose of UV light received? If so, what might be the explanation for these trends?

Some students in the class will probably obtain red-colored colonies on their UV-treated plates. We are going to focus our future studies on these red yeast. Therefore, our second mission this week will be to begin studying previously identified strains of red yeast. You will use the streaking technique you learned last week investigate the growth characteristics of red and white yeasts. This plate test will address the question of whether red yeast and white yeast have different nutritional requirements. To generalize, we can ask whether there is a single difference between the two types of yeast (the color difference) or whether other differences exist, and if so, whether these differences are correlated. We will provide you with three yeast (*Saccharomyces cerevisiae*) strains (HAO, HB1, HB2) that are growing on rich medium, medium that has all the nutrients that the cells could possibly need. You will test whether these strains can grow on minimal medium, or medium that has limited types of nutrients. In particular, you need to determine if the strains can grow without added adenine, a component of DNA. Some strains can make their own adenine, but others cannot and therefore need to “eat” it from their environment before they can make their DNA.

#### Methods

- Evaluating the effects of UV light on white yeast: Examine your petri plates from last week. Record in your lab notebook the number of colonies on each UV-treated plate and on the untreated control plates. In some cases the number of colonies will be so great that you can only estimate the number. To do this, count the number of colonies in one square centimeter of the plate and multiply by 63.6 (this is the total surface area of the plate in cm<sup>2</sup>). To keep track of which colonies have been counted, it is helpful to put a small dot of sharpie marker on the *bottom of the plate* over each colony as you count it.

*Later in the semester, you will need these numbers, as well as notes about the other observations you will make regarding this project, for your lab report. Make sure you write your notes in a safe place and that EACH person in the lab group has ALL of the information every week.*

You also will need to make a correction when comparing the number of colonies on your UV-treated plates to that of the untreated plates. Remember, when you made your control plates you diluted your culture before plating it. You need to determine what was the dilution factor in this case.

*Here's an example of how dilution factor is calculated: if you had diluted 0.5 ml into 9.5 ml of water, then you would have diluted your 0.5 ml culture to make 10 ml of dilute culture. Your dilution factor would be  $10 \div 0.5 = 20$  fold dilution. Calculate the dilution factor for the dilution you used on your control plates.*

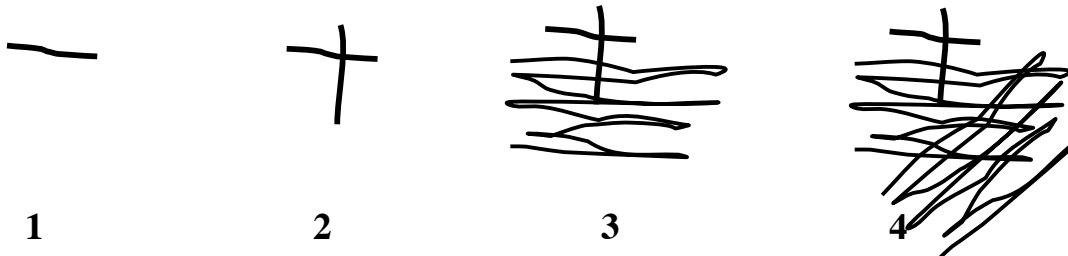
Now you need to use this dilution factor in comparing the number of colonies on the UV-treated plates to your control plates. How do you think you can use the dilution factor in making this comparison?

Also record any differences you see between colonies on plates that had been subjected to UV light and colonies on plates that had not. Are there any examples of colonies that look unlike all other colonies (for example, red rather than the typical white color, have strange colony morphology, etc.)?

- Evaluation of practice streak plate: Last week you used a sterile technique called a streak plate to try to separate a mixed culture of yeast. Look at your streak plate from last week. Did you succeed in separating the two types of yeast into single colonies of each?
- Streak plates to test nutritional requirements of red vs. white yeast: In addition to using the streak plate technique to purify colonies, it can also be used to place yeast on different growth media in testing their nutritional requirements. Obtain a minimal medium plate, a “minimal plus adenine” plate and a “minimal plus amino acids” plate. Using a marker on the **bottom** of the plates, draw lines dividing each plate into thirds (like a Mercedes hood ornament!) and label the three areas HAO, HB1, HB2. Also write on each plate what type it is (minimal, min plus adenine, min plus amino acids).

You are now going to streak three strains onto each of the three plates. In order to make the streaks fit into their sections of the plate, you will use a slightly different streaking pattern than you did last week. Obtain plates (previously prepared by the instructors) with the yeast strains HAO, HB1, and HB2. Flame your loop until red-hot to sterilize it. Cool loop by sticking it into the agar of a plate (not into a colony). Touch the cooled loop to a single colony of HAO. Don't scrape up a large amount of the colony – a barely visible or invisible amount is what you want. Remember cells are too small to see, so there are thousands of cells clinging to your loop, even though you can't see them. If you pick up too many cells it will be difficult to interpret the experiment later. Touch the loop to the minimal medium plate and make a small horizontal streak (see diagram below, 1) near the edge, in the third of the plate that you marked “HAO.” Flame and cool the loop and make

another small streak, this time vertically, that crosses the first streak (2). Flame and cool the loop again, then streak back and forth across the last streak to spread cells over about half of the area of the HAO section (3). Then, streak slantwise across those streaks to spread the cells out further still (4). Try to use most of the room available to you in that third of the plate.



Repeat this process for the HB1 and HB2 strains, making sure to sterilize the loop before touching each strain. Now, using the same procedure, have your lab partner streak all three strains on to the “minimal plus adenine” plate. Now fight over which one of you will repeat the process for the “minimal plus amino acids” plate.

- Streak plates to isolate red colonies from UV-treated plates: If you obtained any red colonies on your UV-treated plates, obtain one YPD plate to use in purifying the colony. Flame a loop, cool it in agar, and touch it *lightly* to your red colony. Use the streak plate sterile technique and the new YPD plate to purify the colony. Use the whole area of the plate and four streaks in a big square, as you did last week, to purify one red colony. If you succeed, you can use your pure red colonies next week to test your red strain’s nutritional requirements. If you found more than one red colony, please pass the extras on to lab groups that did not obtain a red colony.

*Don’t worry if you did not get to do this step; we will give you a red strain to work with in future weeks.*

- Aliquoting and inoculating liquid medium: Next week we will give you liquid yeast cultures that you will use in isolating DNA. This week you will practice preparing such a culture, just so you can see how it is done and to practice your sterile technique. Each person should do this practice exercise independently.

Label two sterile test tubes: “control” and “yeast culture”. Put sterile pipet in pipet aid. Pipet 5 ml from a tube of liquid YPD to each of the test tubes. Manipulate tube caps with pinkie, flame lips of tubes before and after transfer. Put used pipet in pipet canister.

*The flame not only kills bacteria and fungus on the lip of the tube, but also causes air to move out of the tube, so that microorganisms floating in the air do not enter the tube. Handling the lid with your pinkie frees up the rest of your hand for other jobs.*

Obtain a petri plate containing yeast colonies (use a plate prepared earlier by the instructors). Flame your loop until red-hot to sterilize it. Cool loop by sticking it into the agar of the plate (not into a colony). Touch the cooled loop to a yeast colony. Dip the loop containing the yeast into the tube of nutrient broth you prepared (labeled “yeast culture”), flaming the lip of the tube before and

after inoculation. Leave the control tube alone. Put tubes in rack in front of room for 30° incubation. Next period examine these tubes for growth.



**Lab Assignment**  
**Due at the start of lab in week 3**  
**15 points**

Using a word processor, answer the following questions.

- (2 pts) 1. Which two sterile techniques allow you to separate cells in a culture into individual colonies? Explain why these techniques can accomplish this separation.
- (1 pt) 2. In the inoculation experiment, do you expect growth in the control tube? What would it mean if there were growth?
- 3 Consider the following results:

Strain	Minimal	Minimal plus adenine
1	no growth	growth
2	growth	growth
3	no growth	no growth

- (1 pt) (a) Can strain 1 make its own adenine?
- (1 pt) (b) Can strain 2 make its own adenine?
- (2 pts) (c) Explain how you arrived at your answers for (a) and (b).
- (2 pts) (d) What situations might explain the results obtained with strain 3? The best answer will provide two different hypotheses. Consider possible technical problems and possible nutritional requirements of the strain in formulating your hypotheses. Propose a nutritional test experiment that could differentiate between the two possibilities
- (2 pts) 4. In your lab report, you will be asked to compare the number of colonies on your UV-treated plates to your control plates. How do you think you can use the dilution factor in making this comparison?
- (2 pts) 5. In the following weeks you may be asked to interconvert milliliters ( $\text{ml} = 10^{-3}$  liters) and microliters ( $\mu\text{l} = 10^{-6}$  liters). How many microliters are in 0.3 ml? How many milliliters are in 250  $\mu\text{l}$ ? (please handwrite this answer to show your work)
- (1 pt) 6. Look at the flowchart entitled “Simplified adenine synthesis pathway” in the lab manual’s table of contents. The flowchart shows a small portion of the steps needed to make adenine. Three genes encode three proteins (enzymes) – for example, gene A contains a code telling the cell how to build protein A. Each protein does a job. For example, protein A converts a colorless chemical, represented by the open shape, into a red chemical, represented by the filled-in bullet shape. The proteins must do their jobs in the order indicated – protein B cannot do its work until protein A has done its job. A student from last year’s class claims that if a cell contains an altered version of gene A that encodes a non-functional version of protein A, then the cell will remain colorless (white). However, if a cell

contains an altered version of gene B that encodes a non-functional version of protein B, then the cell would turn red. Does this seem possible? Explain your answer.

- (1 pt) 7. Would an alteration in gene C leading to a non-functional protein C result in red or colorless/white yeast? Briefly explain.

## **Lab, Week 3**

### **From Gene to Protein to Behavior**

#### **PREPARATION OF DNA FROM A RED YEAST STRAIN**

In the first week of lab you subjected a white yeast strain to ultraviolet light. After the UV treatment some of you observed a red colony amidst the thousands of white colonies.

In other studies you have seen that red yeast and white yeast strains isolated previously have other differences in addition to color. The two types of strains also showed a difference in their growth properties. The red strains were unable to grow unless adenine was provided to the agar nutrients while the white strain was able to grow in the absence of adenine. Today you will check any red strains you or your classmates isolated to see if they too require adenine for growth.

Because UV treatment is known to alter DNA, both the difference in colony color and nutritional requirements could be due to an alteration in the DNA. A particular DNA change could convert a white, adenine-synthesizing strain into a red, adenine requiring strain. This is our working hypothesis. If our hypothesis is correct then we should be able to identify a difference in the DNA when we compare the DNA of the red and white types. To study the DNA we will separate it from all of the other cellular constituents. This week you will prepare a purified sample of DNA from cells that formed red colonies after being subjected to UV light. DNA from cells that form white colonies has already been analyzed in an international multi-million dollar collaboration similar to the much-touted human genome project. We can therefore compare the results from the DNA analysis from the cells that formed red colonies after UV exposure to the DNA from the cells that form white colonies.

#### **Methods**

##### **DNA isolation**

The following protocol has many steps. It is like a complex recipe. Make sure you complete each step before moving onto the next. The order is very important. Make sure you are comfortable with the use of pipetmen before you begin.

- 1) We will give you a liquid culture of a red yeast strain obtained from one of the lab groups. Harvest the red yeast cells by centrifuging for 5 minutes at 2000 RPM (revolutions per minute). An appropriate centrifuge will be placed on one side of the laboratory.
- 2) Gently pour the liquid out of the tube and, using a P1000 pipetman, resuspend the pellet of cells in 600 microliters of a buffer labeled "Sorbitol buffer". Transfer the liquid to an eppendorf tube (small plastic tube). Then, using a P20, add 10 microliters of a solution marked "lyticase". Invert the tube several times to mix the cells with the lyticase and incubate the tube at 30°C for 30 minutes.

*Yeast cells have a very sturdy cell wall that will prevent the liberation of the DNA. In order to disrupt the cell wall you are adding an enzyme, lyticase, which will catalyze the destruction of the cell wall.*

- 3) Pellet the cells by centrifuging for 3 minutes at 2300 RPM in a microcentrifuge. Use a P1000 to pull off the bulk of the supernatant (liquid layer). Then use a P200 to remove the small amount of remaining liquid.

*After the initial centrifugation the pellet may have a “tail” up the side of the tube. If the tail flops over into the liquid while you are trying to remove the liquid, respin the tube in the microcentrifuge for 15 seconds. Remove the remaining liquid with a P200.*

- 4) Once you have removed all of the liquid, resuspend the cells in 180 microliters of buffer labeled “ATL”. With a P200 set at 120 microliters, gently pipet the mixture up and down to homogenize the solution.
- 5) Add 20 microliters of the enzyme Proteinase K, mix by vortexing for approximately 5 seconds, and incubate at 55°C for 30 minutes.

*ATL buffer provides the optimum conditions for the enzyme Proteinase K to work. The buffer also contains detergent to dissolve the cell membrane. Proteinase K is an enzyme that catalyzes the breaking of peptide bonds that hold together proteins. Some proteins help hold the cell together and others are wrapped closely around DNA. The proteinase K can remove these proteins.*

- 6) After the 30 minute incubation in the previous step, vortex the sample for 15 seconds. Then add 200 microliters of solution “AL”, vortex briefly but thoroughly (15 seconds) to mix the sample. Then incubate the sample at 70°C for 10 minutes.

*A white precipitate may form upon addition of AL but this is reversed after the 70°C incubation.*

- 7) After the 10 minute incubation add 200 microliters of ethanol. Vortex to mix the solution.

*At your lab bench you should find two small tubes, one nested inside the other. The tube on top is not just a tube (in fact it is called a “column”). It has a white membrane that will absorb DNA under some conditions but release it under others. You will apply the solution of DNA and other macromolecules onto this column under a condition that will trap the DNA but other molecules will wash through. Then you will wash the column under a variety of conditions to ensure that all macromolecules other than DNA have been washed through the membrane. Finally, you will wash the membrane under a condition that will cause the DNA to be released.*

- 8) Label the cap of the column and the side of the collection tube. Pipet the solution onto the white membrane making sure not to stab the membrane with the pipet tip.
- 9) Wash the contents through the column by applying centrifugal force---that is, place the column nested in its collection tube in a microcentrifuge and centrifuge at maximum speed for one minute. Discard the fluid in the collection tube.

*At this point the DNA is stuck on the white membrane.*

- 10) Place the column in the emptied collection tube and add 500 microliters of a buffer labeled “AW1” and, as in the step before, centrifuge. Again, discard the fluid in the collection tube.
- 11) Place the column in the emptied collection tube and add 500 microliters of a buffer labeled “AW2” and, as in the step before, centrifuge. However, this time centrifuge the sample for 3 minutes at full speed. This will eliminate almost all liquid from the column in preparation for eluting the DNA. Throw out the contents of the collection tube. SAVE the column.
- 12) Place the column in the emptied collection tube and use the microcentrifuge to spin the column again, this time for one minute.
 

*This second centrifugation will ensure that all liquid has been removed from the surface of the membrane. Although dry, the membrane has DNA stuck to its surface.*
- 13) Get a fresh plastic collection tube for this step. Label the tube with your lab group’s initials. Place the column inside the collection tube and then add 200 microliters of “AE” directly onto the membrane of the column without stabbing the membrane. Let the “AE” incubate on the membrane for one minute at room temperature. Then place the column and collection tube into the microcentrifuge. Centrifuge the sample for one minute at full speed. Save the liquid sample; this has your purified DNA! Transfer the liquid sample to a small 1.5 ml plastic tube with a lid (ependorf tube) for safekeeping. Save the column and the collection tube for the next step.
- 14) In order to maximize the amount of DNA recovered, put the column back in the now-empty collection tube. Add another 200 microliters of “AE” to the column, incubate for one minute, then centrifuge as in the step above. Add the solution in the collection tube to the contents of the ependorf tube from the step above. Label the tube with your lab group’s initials and “red yeast DNA” and place it in a rack to be stored in the freezer.

Voilà. This ependorf tube has the DNA from the cells that you harvested 3 hours ago. The DNA is now ready for analysis. This will take place next week.

### **Analysis of nutritional requirements of red yeast isolated by you or your classmates**

If all went well, you or your classmates identified some red colonies after UV treatment of the white yeast strain. These colonies should have been purified last week by using a streak plate to generate isolated pure red colonies.

Find a pure red colony that is well isolated from its neighbors on your streak plate, or that of a classmate. Obtain minimal and minimal plus adenine plates and mark the bottom in thirds, as you did last week. Using the same technique as last week, assess whether this new red strain requires adenine to grow by streaking it onto minimal and minimal plus adenine plates. Also streak white HAO and a previously characterized red strain such as HB1 or HB2 as controls.

### **Assessment of inoculation experiment**

Look at the tubes of liquid media you prepared last week. Did you succeed in growing a yeast culture similar to the one you used today? Is the uninoculated tube free of contamination by bacteria or fungi?

You do not need to report these results to us, but they should give you a check on how well your sterile technique is working, so you might want to consult us about any problems.

## Lab, Week 4

### From Genes to Proteins to Behavior

# DNA FROM THE RED YEAST STRAIN: VISUALIZING BY AGAROSE GEL ELECTROPHORESIS AND AMPLIFYING A SEGMENT BY THE POLYMERASE CHAIN REACTION & CHARACTERISTICS OF THE *ADE1* AND *ADE2* GENES FROM THE WHITE YEAST

For the past several weeks you have been working with two strains of yeast, one whose cells form red colonies and one whose cells form white colonies. In the first week of lab you subjected the yeast that gives rise to white colonies to UV light. On rare occasions this treatment resulted in colonies that were red. The lab work at this point forward is aimed at defining the underlying difference between the red and the white types of yeast that explains the difference in color. And does this difference explain the difference in nutritional requirements?

Because UV light is known to mutate (change) DNA, we are focusing our attention on comparing the DNA from the red and white yeast. In order to analyze the DNA we needed to separate it from all of the other cellular constituents (parts). This is what you did last week in the laboratory.

There are three objectives in today's laboratory period. The first objective is to test whether you recovered DNA from last week's laboratory period. The procedure that you used involved many steps and it is not difficult to imagine that one of the steps went wrong. You will use a technique called agarose gel electrophoresis in order to determine whether there is DNA in the final solution you recovered last week. The second objective is to make many copies of two segments of DNA (genes; *ADE1* and *ADE2*). The technique known as polymerase chain reaction (PCR) will be used. This technique will allow us to generate enough copies of the DNA to allow for the determination of the sequence of bases of the genes *ADE1* and *ADE2*. (Note that to successfully complete the second objective you have to successfully complete the DNA preparation). The analysis of the base sequence of the two genes from the red yeast will be the goal of next week's laboratory period. The third objective of this week's lab is to learn about the characteristics of the *ADE1* and *ADE2* genes from the white yeast; this information is accessible in a public database.

## I. AGAROSE GEL ELECTROPHORESIS

An introduction to gel electrophoresis can be found in your textbook (Purves, 6th edition, p.312-313). Below is a brief description.

DNA fragments differing in length can be detected and separated by gel electrophoresis. A mixture of negatively charged DNA fragments is placed in a porous gel and placed in the presence of an electric field. The negative charge on each DNA fragment is proportional to its length because each phosphate group in the chain is negatively charged. As the fragments move through the electric field, attracted by the positively charged electrode, the gel resists their movement but with different intensity; the longer fragments are retarded more strongly than the smaller fragments. Therefore the smaller fragments move the fastest and travel the

farthest across the gel. When the electric field is removed, each DNA fragment will have stopped at a different point on the gel. Each visible "pile" of DNA is called a band, and the path that the DNA travels in is called a lane. If desired, the separated fragments can be removed from the gel in their pure form.

Different samples may be "run" side by side in different "lanes" in the electric field. Conventionally, and in today's laboratory experiment, DNA fragments of known length are run next to experimental mixtures to provide a size reference.

The DNA on the gel can be visualized by staining the gel with a fluorescent dye and subsequently placing the gel under ultraviolet light. This dye is known as ethidium bromide and has been incorporated into the agarose gel. **CAUTION!! Ethidium bromide is a carcinogen, so wear gloves whenever you handle or load gels.** We will be working with low levels of ethidium bromide in this experiment. **In addition, electrophoresis involves electrical current high enough to cause severe injury. NEVER touch an electrophoresis apparatus unless the power supply is turned off and the leads are disconnected.** As an extra precaution, connect gel to the power supply only when it is turned off and using only one hand.

Recall that last week you purified DNA from yeast cells. We have no verification however, that you were successful. Everyone in the lab has an eppendorf tube with a clear liquid solution. Is there any DNA in that solution? Today we will use agarose gel electrophoresis to determine if your tube does indeed have DNA and assess whether the DNA has a high molecular weight---that is, whether DNA is a very long polymer of nucleic acids.

### Methods

- Obtain your yeast DNA sample from the instructor.
- Using a pipetman (P20) withdraw 15 microliters of your DNA sample and pipet it into an eppendorf tube.
- Again using the P20, add 3 microliters of gel loading dye to your sample. Close the lid and mix the dye and your DNA sample together.

*The loading dye serves two purposes. First, it gives the sample a density greater than water. This means when you add your sample to an aqueous medium it will sink. Second, the migration of the dye solution will give an indication of how long the sample should remain in the electric field.*

15) The agarose gel has been prepared in advance. Take your sample mixed with dye to the electrophoresis chamber. The gel will be loaded as a group effort. Take careful notes so that you know which lane has your sample and which samples were loaded in the other lanes.

After all of the samples are loaded on the gel, a current will be passed between the two electrodes. This will start the process of electrophoresis. After about 45 minutes the samples should be ready to visualize.



## II. SINGLE COLONY ISOLATION OF A RED STRAIN OF YEAST

Last week you streaked a red colony that was found on the plates that had been exposed to UV light. Find the plates that you streaked last week. Examine them carefully to determine if any of the colonies on the plate are red. If you see red and white colonies on the plate, propose a reason that not all of the colonies are red.

Find a well-isolated red colony on this plate and re-streak it on a new YPD plate. Label the plate using the following format: initial of first names of lab partners followed by a slash, followed by the day the lab section meets, followed by a slash, followed by the date. For example, Hilary and Chip would have a plate labeled HC/Tu/27 if they were in the Tuesday lab section and the date was January 27.

## III. THE POLYMERASE CHAIN REACTION (PCR): *IN VITRO* AMPLIFICATION OF DNA

After subjecting *S. cerevisiae* cells to ultraviolet light, several cases of cells that had acquired a new property (red color) were observed. Because one effect of UV light is to chemically alter the DNA it is suggested that the DNA in the red strain of yeast will be different from the DNA in the normal, cream-colored strain of yeast. In order to test this hypothesis we will examine the sequence of bases in a particular region of the DNA from the red *S. cerevisiae* cells. This can be done using the DNA you prepared last week.

Determining the sequence of bases is technically challenging. The amount of DNA you harvested last week, while detectable, is too small to allow for such analysis. Further, we will only examine a small region of DNA, accounting for approximately 0.0001% of the DNA in the yeast cell. Thus, our goal for today is to synthesize large quantities of the small region of DNA of interest. This will be done using the technique known as the polymerase chain reaction (PCR). Given a large quantity of DNA, the base sequence can then be determined.

PCR is an *in vitro* (in test tube) method to replicate DNA. It bears some basic resemblance to the way DNA is replicated inside the cell but also differs in some fundamental ways. The power of PCR technology lies in the ability to start with tiny quantities of DNA material that are essentially undetectable (and in some cases in a very crude state) and selectively amplify a region of interest into such amounts that the amplified DNA is simple to detect. PCR technology is a staple in most genetics and biology laboratories. In this laboratory period, you will become acquainted with the polymerase chain reaction technology and use this technology to amplify a small region of DNA for subsequent base sequence analysis.

The strains of yeast with red color have been associated with a defect in the ability to make adenine. Adenine is chemically related to several molecules that have a red color, although adenine itself does not. Normally these molecules that have a red color are converted to adenine by two different enzymes. Two genes, *ADE1* and *ADE2*, encode these enzymes. A defect in either one of these enzymes can lead to the accumulation of one of these molecules that have red color, thus the red color yeast. In this lab you will amplify DNA from the *ADE1* gene and the *ADE2* gene. In the following week we will examine the base sequence of parts of these two genes.

A description of PCR is provided by Purves (Life: The Science of Biology, p 214-216 in the 6<sup>th</sup> edition). You can also refer to a figure (Figure 11.21 on p.216 in the 6<sup>th</sup> edition) that demonstrates, in a color schematic, the process of PCR. In the lab you will work on a critical thinking exercise that should demonstrate how PCR works. Furthermore, we will show an animation on the Cold Spring Harbor laboratory website (<http://www.dnalc.org/ddnalc/resources/pcr.html>) that also demonstrates the polymerase chain reaction.

You should recognize that to perform a PCR reaction the following components are needed:

- 1. Source of DNA to be amplified.** This starting material may be purified DNA or can be small amounts of tissue or cells. In forensic medicine, one is often forced to use small bits of tissue left at the crime scene. In today's lab you will use your DNA preparation from last week.
- 2. A pair of short, single stranded DNA molecules called oligonucleotides** that will be used to prime the replication reaction. Because of their function these oligonucleotides are also called *primers*. In this particular case, you will be given primers of the appropriate base sequence to amplify the *ADE1* gene and the *ADE2* gene.
- 3. Thermostable DNA Polymerase enzyme.** DNA polymerases are enzymes that have the ability to incorporate a nucleotide into a growing chain of DNA. For PCR, it is essential that the enzyme be heat stable (thermostable). Heat stable enzymes are generally found in bacterial sources that grow in very harsh environments. The particular enzyme used most often is called Taq polymerase. *Thermus aquaticus* grows in and around hot geysers.
- 4. deoxynucleotide triphosphates (dNTP's).** These are the building blocks of DNA and are the molecules that the polymerase enzyme incorporates into the growing DNA chain.
- 5. Buffer.** This will make sure that the pH and ionic requirements for the DNA polymerase are met.

Now that you have learned about how PCR works, it is time to try to set up an actual reaction:

### Methods

Each lab pair will do two PCR reactions.

- Obtain two small eppendorf tubes containing a “Ready-to-go PCR bead”.

*Today’s techno-wizardry allows for the packaging of three of the key components in one dried pellet. This bead has the dNTP’s, the DNA polymerase, and the buffer. Traditionally there would be three separate solutions to pipet (dNTPs, DNA polymerase, and the buffer).*

- Write your lab initials and a “1” on one of the two eppendorf tubes with a bead. Add to that eppendorf tube 12.5 microliters of a solution labeled “primer mix 1”.

*This solution has an equal amount of two primers that will amplify a gene known as ADE1. You should recognize that two primers are required for PCR amplification of DNA.*

- Next add 12.5 microliters of your DNA solution from the previous week’s lab period to the eppendorf tube that already has the bead and primer mix 1. At this point all of the key ingredients have been added that will allow for amplification of the ADE1 gene.
- Write your lab initials and a “2” on the other eppendorf tube that contains a bead. Add to that eppendorf tube 12.5 microliters of a solution labeled “primer mix 2”.
- Next add 12.5 microliters of your DNA solution from the previous week’s lab period to the eppendorf tube that already has the bead and primer mix 2. At this point all of the key ingredients have been added that will allow for amplification of the ADE2 gene.
- If needed (ask your instructor), add one drop of mineral oil to each of the two samples.

*The purpose of the mineral oil is to prevent the evaporation/condensation of the aqueous sample onto the inside of the lid of the tube. Mineral oil floats on the surface of the sample and does not evaporate even at high temperatures. Some types of thermocyclers do not require the use of oil.*

- Place your samples in the thermocycler (the machine that carries out the temperature regimen for PCR). The machine is programmed to run the following temperatures and incubation times:  
4 minutes, 95°C followed by 35 cycles of: 90 seconds, 92°C; 60 seconds, 45°C, 120 seconds, 72°C.  
Be sure you understand the reason for each of these three temperatures (high, low, medium).

## IV. CHARACTERISTICS OF THE *ADE1* AND *ADE2* GENES FROM THE WHITE YEAST

### Mining a database for information about the *ADE1* and *ADE2* genes and their corresponding proteins

A gene is a segment of DNA that codes for a protein. A unicellular organism such as *S. cerevisiae* has several thousand genes. In the case of this yeast, the sequence of bases that make up all of the genes is known. This is a Herculean task and has therefore only been done for the normal white type and not for the red type of yeast. Today's objective is to locate in a database the two genes of the white form of yeast that we are interested in *ADE1* and *ADE2*. Since the sequence of the bases is already known, we should be able to find this information on the internet. We will also find the sequence of amino acids that make up the protein that is encoded by each of the two genes. Remember that the long-term objective of the project is to determine if there has been a change in the sequence of bases in either of two genes, *ADE1* and *ADE2*, in the yeast that form the red colonies compared to those that form the white colonies. This objective will be accomplished by next week.

For today, each lab pair will divide the tasks. One person will work with the *ADE1* gene while the other will do the same for *ADE2*. For those that are working on the *ADE2* gene, skip down to the heading: "**Finding the *ADE2* gene in the database**".

### Finding the *ADE1* gene in the database.

1. Open the internet browser and type in the address

<http://www.yeastgenome.org>

This is the portal to the *Saccharomyces* Genome Database (SGD), the database that will have the sequence of bases for all six thousand genes. But note that this database is only for the white type of yeast.

2. In the box that is labeled "QuickSearch" type "ADE1" and click on the "Submit" button.

3. The screen that is showing is fairly complicated. The title for the page is ADE1/YAR015W. The *ADE1* part should be comforting and there is no need to get anxious about the YAR015W. Each gene in the database has a formal identifier like *ADE1* and another identifier that indicates (in a non-obvious way) the location ("address") of the gene on a particular chromosome. On the right half of the screen there is a column of buttons. Move the cursor to the button that has a heading "Retrieve sequences" and is labeled "Genomic DNA" and click and hold. You will see eight options to select from on the drop-down menu. Choose "coding sequence" and then click on the button to the right "Retrieve". Like all genes, the *ADE1* gene codes for a protein. By selecting this option the display now shows the DNA sequence found in the white type of yeast that codes for this protein.

The current screen shows the bases that make up the *ADE1* gene. There are other ways of arranging this same data, one of which turns out to be much easier to determine features such as length (in number of bases). This format is known as GCG. To see the GCG format click on "GCG" next to the phrase "Other formats available". Notice that this format has numbers that signify the number of bases from the first base at the start of each row. The bases are presented in groups of 10 with a space between each group provided for your viewing pleasure. Just prior to the presentation of the



Click on the “back” button twice on the internet browser to return to the original screen. Under the heading “Retrieve sequences” select the option “ORF Translation” and then click on “Retrieve”. (ORF is an abbreviation for open reading frame. An ORF is a sequence of DNA that can code for a protein.) Note that the screen that is generated has characters of almost all letters in the alphabet. This string of characters on the display represents single letter abbreviations of amino acids. Check to see if the first ten amino acids that you predicted in step 3 are consistent with what is being displayed here. In order to see the same data displayed in a different manner, click on “GCG” next to the phrase “Other formats available.”

*How many amino acids are connected to make the protein encoded by the ADE1 gene?*

*How many nucleotides were required to code for this many amino acids?*

5. Once more click on the “Back” button twice on the internet browser to return to the *ADE1* main page. On the left side of the page examine the information under “Classical Genetics” (you do not have to follow the link; just look at the information here). Is the information provided consistent with the hypothesis that a change (mutation) in the *ADE1* sequence could lead to the red color property? Also note the use of a term “auxotrophy”. The term “auxotroph” means that something must be added to the diet in order for the organism to grow. Is this information also consistent with the hypothesis that we are pursuing?

Next look at the entry under "Biological Process" on the left side of the page. Here the entry states "purine nucleotide biosynthetic process". Biosynthesis is the process of construction of biological molecules, in this particular case purine nucleotide bases. Do you have any experimental evidence (if not, how could you attempt to gather such evidence) that cells that give rise to red colonies are also affected in purine base biosynthesis?

6. Near the top of the page under “Description” we can see that the protein encoded by the *ADE1* gene is called N-succinyl-5-aminoimidazole-4-carboxamide ribotide (SAICAR) synthetase. This is certainly a mouthful but note the suffix "-ase" that always signifies that the term is describing an enzyme. A bit further down the page, under the heading “Molecular Function” you will see an alternative name for the same enzyme, phosphoribosylaminoimidazolesuccinocarboxamide synthase. You will see this name again next week when we find a human protein that has the same name.

### Finding the *ADE2* gene in the database.

1. Open the internet browser and type in the address

<http://www.yeastgenome.org>

This is the portal to the *Saccharomyces* Genome Database (SGD), the database that will have the sequence of bases for all six thousand genes. But note that this database is only for the white type of yeast.

2. In the box that is labeled “QuickSearch” type “ADE2” and click on the “Submit” button.
3. The screen that is showing is fairly complicated. The title for the page is ADE2/YOR128C. The *ADE2* part should be comforting and there is no need to get anxious about the YOR128C. Each gene in the database has a formal identifier like *ADE2* and another identifier that indicates (in a non-obvious way) the location (“address”) of the gene on a particular chromosome. On the right half of the screen there is a column of buttons. Move the cursor to the button that has a heading “Retrieve sequence” and is labeled “Genomic DNA” and click and hold. You will see eight options to select from on the drop-down menu. Choose “coding sequence” and then click on the button to the right “Retrieve”. Like all genes, the *ADE2* gene codes for a protein. By selecting this option the display now shows the DNA sequence necessary to code for this protein.

The current screen shows the bases that make up the *ADE2* gene. There are other ways of arranging this same data, one of which turns out to be much easier to determine features such as length (in number of bases). This format is known as GCG. To see the GCG format click on “GCG” next to the phrase “Other formats available”. Notice that this format has numbers that signify the number of bases from the first base at the start of each row. The bases are presented in groups of 10 with a space between each group provided for your viewing pleasure. Just prior to the presentation of the base sequence is a reference to the original scientific manuscript that first identified and reported the base sequence of *ADE2*.

The bases that are shown represent the sequence of ONE strand of the DNA. For the questions below you do NOT have to figure out the sequence of the RNA, as you had to do in a homework assignment. Recall, one strand of the DNA in a gene is used as a **template** to form an mRNA, and then this mRNA is used in translation as instructions for building a protein. The DNA sequence that you are viewing is the **non-template strand** of the *ADE1* gene. Like the mRNA, the non-template strand is complementary to the template strand of the DNA. Thus, the mRNA sequence and the sequence of the non-template strand of the DNA are **identical**, except that the mRNA contains U’s where the DNA contains T’s:





5. Once more click on the “Back” button twice on the internet browser to return to the *ADE2* main page. On the left side of the page examine the information under “Description”. Based on the information provided, is it possible that a change (mutation) in the *ADE2* sequence could lead to the red color property? Also, examine the heading “Mutant Phenotype, Classical Genetics” and note that a comment similar to the one under “Description” is here. In addition, note the additional comment in this section “auxotroph”. The term “auxotroph” means that something must be added in order for the organism to grow. Is this information consistent with the hypothesis we are testing and any observations that you have made?
6. You will now collect a bit more information about the *ADE2* gene and the protein encoded by the *ADE2* gene. The top of the page gives a description of some of the properties of the gene and the protein. Under the heading “Description” is the name of the enzyme encoded by *ADE2*. It is certainly a mouthful--- phosphoribosylaminoimidazole carboxylase---but note the suffix “-ase” that always signifies that the term is describing an enzyme. We will come back to this name in the computer lab next week when we find a human protein that has the same name.

Next look at the entry under “Biological Process” on the left side of the page. Here the entry states “purine base metabolic process”. Metabolism is the process of construction or destruction of biological molecules, in this particular case purine bases. Do you have any experimental evidence (if not, how could you attempt to gather such evidence) that cells that give rise to red colonies are also affected in purine base metabolism?

## Lab Assignment

**Due at the start of lab in week 5**

**10 points**

Next week we will be analyzing the PCR reactions using agarose gel electrophoresis. In the meantime, you should answer these questions about this lab.

1. Consider a typical enzyme found in a human cell.

(a) Graph the activity of this enzyme (the ability of the enzyme to do work) as a function of temperature (temperature should be on the X-axis). The X-axis should go from 10°C to 100°C. (1 point)

(b) Now, on the same graph plot a similar relationship for Taq DNA polymerase, showing the way in which it differs from a typical human enzyme. (1 point)

2. Assume one lab group fell behind and was not able to put their samples into the thermal cycler until after the 94°C step was over. Thus, their sample was added to the machine when it was at 45°C. After the last step of the first cycle (that is, after the 72°C step) would there have been any DNA synthesis in this group's sample? Explain your answer. (3 points)

3. List the distinctions between DNA replication by the polymerase chain reaction and DNA replication *in vivo* (in the cell). (5 points)

## Lab, Week 5

### From Genes to Proteins to Behavior

#### ANALYSIS OF THE BASE SEQUENCE IN THE DNA OF THE *ADE1* AND *ADE2* GENES

In the first laboratory period you examined the effect of UV light on *S. cerevisiae*. Many of you were able to note that this treatment resulted in a number of different types of variants, including colonies that appeared red. That the red colony types arose from the white suggests that the two are highly related (highly similar). Are there any other differences between red and white strains in addition to the difference in color? You tested the growth properties of the red and white yeast under two conditions, with adenine provided in the agar and without adenine provided in the agar. The two types of yeast showed a difference with respect to their growth. Thus the two types of yeast, although presumably highly similar, seem to differ with respect to two characters, colony color and nutritional requirements.

So in those colonies that were red that came from a white progenitor, what changed? What is the underlying difference that leads to the difference in colony color?

In the third week of laboratory you prepared DNA from the red type of yeast. We are testing the hypothesis that the difference between the red and the white yeast is a difference in one of two genes. The names of the two genes are *ADE1* and *ADE2*. Genes are segments of DNA and therefore in order to test this hypothesis we need to obtain and characterize the DNA from the red and the white yeast. We will examine *ADE1* and *ADE2* genes (there are a total of 6000 genes in the yeast) because in previous cases red colonies were associated with changes in one of these two genes. You noted this in the computer work you did last week in lab. By the end of the third laboratory period you obtained DNA from the red type of yeast.

(There are no set rules for naming of genes but it is typical that the name of a gene reflects the function of the protein encoded by the gene. Recall that in the first week you showed that the red version of the yeast required adenine in order to grow while the white version of yeast did not require adenine to grow. Previous researchers have shown that in versions of yeast where cells are red and adenine is required to grow there is a change in one of two genes. Because changes in each of these genes results in a requirement for adenine the genes were named *ADE1* and *ADE2*.)

In the fourth week of laboratory you performed the polymerase chain reaction (PCR) on the DNA from the red yeast strain. This DNA contains the entire repertoire of 6000 yeast genes. Remember, the hypothesis that we are testing is that the difference between the standard (wild-type) white yeast strain and the red yeast strain is due to differences in the base sequence of either the *ADE1* or *ADE2* genes. The PCR was done in order to restrict our analysis to only two of the 6000 genes and to obtain sufficient quantities of the DNA of the two genes (*ADE1* and *ADE2*) so that the base sequence of each of these two genes could be determined. Also remember that we already know the sequence of the two genes in the standard white type of yeast. This is what you found on the computer last week. We have taken those DNA molecules generated by PCR and sent them to a DNA sequencing facility. Such facilities have the equipment to determine the base sequence in the DNA. A brief animation of how this sequencing technology works will be shown.

The data have come back from the sequencing facility and we will spend today's laboratory period analyzing and interpreting data through the use of computer software. The data analyzed will be from only one of the red mutant yeast strains. These data will then be compared to data you obtained last week on the standard (wild-type) white strain. Will there be any differences between the data we have collected from the red strain when compared to the sequence data of the white strain? If so, will differences be detected in both the *ADE1* and *ADE2* genes?

Before accomplishing these goals we will first check to see if the polymerase chain reaction that you did last week was successful. The amplification of the DNA from the *ADE1* and *ADE2* genes of the red yeast is required in order to have enough material to determine the base sequence of these two genes. In order to determine if the PCR was successful, we will use the same technique that we used last week, agarose gel electrophoresis.

### Agarose gel electrophoresis of the PCR reactions

- Retrieve your samples from the PCR experiment that you did last week. You will analyze these three samples at the same time.
- Cut a square of parafilm approximately 3 inches x 3 inches. Pipet 3 microliters of DNA loading dye on two well-separated spots on the waxy surface of the parafilm. Set a P20 pipetman to 15  $\mu$ l. Place the pipet tip near the bottom of the tube that holds the *ADE1* PCR sample, making sure the pipet tip is below the layer of oil, if oil is present. Withdraw 15 microliters of the sample from the tube. Then pipet the contents of the tip into one of the 3 microliter drops of loading dye. Mix the sample with the loading dye. To do this, pipet the sample up into the pipet and then back onto the parafilm. Repeat this several times. By now, the blue color of the loading dye should be well mixed with your sample. Eject the pipet tip and place a new tip on the pipetman.

The loading dye serves two purposes. First, it gives the sample a density greater than water. This means when you add your sample to an aqueous medium it will sink. Second, the migration of the dye solution will give an indication of how long the sample should remain in the electric field.

- Again using a pipetman (P20) withdraw 15 microliters of the *ADE2* PCR sample and pipet it into the other 3 microliter drop of loading dye. Mix the sample with the loading dye by pipetting up and down several times as described above.
- The agarose gel will be prepared by your instructor. When the gel is ready, **carefully** take your piece of parafilm that has your two samples mixed with dye to the electrophoresis chamber. The gel will be loaded as a group effort. Take notes so that you know which lane has your sample and which samples were loaded in the other lanes.
- After all of the samples are loaded on the gel, a current will be passed between the two electrodes. This will start the process of electrophoresis. After about 75 minutes the samples should be ready to visualize.

Any sample that has detectable amounts of DNA will be made visible when the agarose gel is viewed under UV light. If your PCR reaction(s) did not work then you will not see any bands that fluoresce

under UV light. If your PCR reaction did work there should be a single discrete band (“pile” of DNA). You should be able to compare the size of this band to the size of the bands in the molecular weight standards to get an **approximate** size of the DNA made in the PCR reaction. Note that this is an **approximation** of the size, because DNA sometimes runs a bit differently than the size standards when there is a lot of DNA and the bands are thick, as yours are likely to be. For example, you might be off in your estimate by a 100 bases or so.

Estimate the size of your DNA fragments generated by PCR. Remember you can only estimate the size--you should be able to state with confidence that the size is within a certain range. How do these sizes compare to the sizes of the ADE1 and ADE2 genes from white yeast that you calculated in the computer lab last week (this calculation is exact---to the base-pair)?

Compare the distance traveled by the DNA from each of the two PCR samples. Additionally, compare these distances with the distance traveled by the DNA you recovered from the red yeast in the third week of lab. Based on the gel, are the ADE genes smaller or larger than the total yeast DNA pieces you isolated in the third week of lab? Does this make sense?

## DNA Sequence Analysis

The technology of DNA sequencing has revolutionized biology. This technique is described in your textbook; you will also work through some critical thinking questions to better understand the technique. In addition, a short animation describing how DNA is sequenced will be shown. This will give you an indication of how the following sequence data (see Step 1) was obtained. The computer files that you will need are on the CD labeled "principles- ade exercise." The computer programs you will need are on the desktop or in the applications folder on the hard drive or on the CD. These programs are Editview and DNA Strider. Note that each question that is both underlined and in italics will be questions for the lab assignment.

During today's laboratory period you will use the computer to analyze the *ADE1* and *ADE2* genes from the red yeast and compare the data with the *ADE1* and *ADE2* genes of the standard white yeast. One person in each lab group will study the *ADE1* gene while the other person will study the *ADE2* gene. You will then share your results with your lab partner. If you will be studying *ADE1*, then start reading below. If you will be studying *ADE2*, skip about eight pages and start on the page the top of which reads "Below are the objectives and procedures for studying the *ADE2* gene."

### **Below are the objectives and procedures for studying the *ADE1* gene.**

- **Examining the primary data of the *ADE1* gene**
- **Converting the *ADE1* sequence to that on the complementary DNA strand.**
- **Comparing the nucleotide sequence of the *ADE1* gene from the red yeast to the *ADE1* gene from the white yeast**
- **Determining the amino acid sequence of the protein encoded by the *ADE1* gene**
- **Comparing this amino acid sequence to the sequence from the standard (wild-type) white strain**
- **Examining the database for a human gene that encodes a (weakly) similar protein that carries out the same function**

#### **1) Examining the primary data of the *ADE1* gene**

Open up the file labeled "ADE1r.seq" using the program "Editview". To do this double-click on the "Editview" icon, and under the pulldown heading of "File" select "Open". Then select the file name "ADE1r.seq". This computer file contains the sequence data we obtained from the PCR product generated with the *ADE1* primer set in the previous lab session. This is the sequence of bases corresponding to the *ADE1* gene from the red yeast strain. At this point you should see peaks of four different colors. There is also an X-axis that consists of a time unit of measurement, a nucleotide number, and a nucleotide (designated A, C, G, or T). There are a few things to note:

i) at many points there is only one peak and that the color of the peak determines the nucleotide indicated on the X-axis.

ii) at some points, there is more than one peak (see bases near the beginning, for example).

This is due to technical problems and makes it difficult or impossible to predict the base that is at that position on the DNA strand. This is why there is occasionally a "N" instead of an A, C, G, or T. "N" designates that the identity of the base at this position was not clearly determined.

iii) it is technically more difficult to determine the base sequence near the beginning. It is also technically difficult to determine the base sequence much more than about 500 nucleotides from the start. Therefore we will limit our analysis to the base sequence between positions 115 and 540.

In the bottom left corner of the window there are four icons. One of them reads "CATG". Click once on this icon. This changes the information that you were looking at into text only, rather than the multicolored peaks with the various other annotations. Notice there are several N's near the end of the text string. Highlight the characters between position 115 and 540 (from ACACCG to TTTTACG) and then select "copy" from the "Edit" pull-down menu. You will now move this sequence of bases into a program that allows a search of the *Saccharomyces* database for any similar sequences.

**2) Converting the *ADE1* sequence to that on the complementary DNA strand.** Remember that DNA has two strands of nucleotides that run in opposite directions and are complementary. The sequence you have shows only one strand of the DNA, but because of the strict base pairing rules you can derive what the other strand must look like. For example, if the sequence of one strand is 5' CCACGT 3' then the sequence of the complementary strand must be 5' ACGTGG 3'. As luck would have it, the sequence you have received is actually the sequence of the gene on the template strand, read from the end of the gene to the beginning. In other words, the sequence is in reverse and complementary to the mRNA sequence that would be used to make a protein. In order to make the analysis of the *ADE1* gene easier, we need to turn the sequence around so that it reads like the other DNA strand (called the "non-template strand"). To do this we will use a program called DNA Strider. We could do this manually, but it would be very tedious to write out the complement of hundreds of bases.

Open DNA Strider by double-clicking on the DNA Strider icon in the "Applications" folder or on the CD. Note that it appears that nothing has happened because there is no open window but the program is open. There is now a menu bar that should not look familiar. Under the pull-down "File" select "New" and under "New" select "DNA". A window will open into which you will paste the DNA sequence of the *ADE1* gene from the red yeast. Paste what you had copied in step 1 into this new window by selecting "Paste" from the "Edit" menu or by holding down the apple and the "V" keys. You have now transferred the base sequence from the program "Editview" into the program "DNA Strider."

To convert the base sequence into the complementary base sequence, highlight the base sequence that is in the DNA Strider window. Then under the "Conv" (Convert) pulldown, select "antiparallel". Note that the new sequence that is generated is simply the sequence of the second strand of the original DNA displayed 5' -> 3'. Now select this new sequence and select "copy" from the Edit menu.

**3) Comparing the sequence of the *ADE1* gene from the red yeast to the sequence of the *ADE1* gene from the white yeast.** In the following steps you will search SGD, the database that has the DNA sequence of all genes from the standard (wild-type) white yeast that you used last week, to see if any have the same, or similar, sequence of bases as the sequence you have determined for the *ADE1* gene from the red yeast.

**a) Accessing the database via the internet**

Open Netscape and type in the following address

<http://www.yeastgenome.org>

At this website, the entire base sequence of the standard (wild-type) white strain of *Saccharomyces cerevisiae* is stored. This represents about 10 million base pairs. Nested somewhere in this 10 million base pairs is the sequence for the *ADE1* gene (only about 1000 base pairs long). At the top of this web page within the beige bar select “**BLAST**”. This program will search the entire sequence of bases of the *Saccharomyces cerevisiae* in attempt to find some sequence that is either identical or similar to the “query” sequence (BLAST is an acronym for **B**asic **L**ocal **A**lignment **S**equencing **T**ool). The query sequence is that which you type or paste into the window, as in the following step. In this case, the query sequence is the *ADE1* base sequence from the red yeast strain.

**b) Searching the DNA database:** Now place the cursor into the window above which is written “Type or Paste a Query Sequence” and paste the sequence that you copied in step 1. Use either the Paste command from the Edit menu or hold down the apple and “V” keys to paste. In the window “Choose one or more Sequence datasets” select “Open Reading Frames (DNA or Protein)”. Next select the button “Run WU-BLAST”. The query sequence (the *ADE1* sequence from the red yeast strain) is now being compared to all 10 million base pairs in the database of the standard (wild-type) white yeast strain to find the sequence or sequences that are most similar.

**c) Matches to the query DNA sequence:** After a short time you will see the results of this search although it will be difficult to interpret. Near the top of the page you will see what looks like a number line. This line represents the sequence you put in, with the numbers indicating the nucleotide number. Under this numbered line you will see one or more colored thick lines (actually they are arrows). Each thick line represents a match to the query sequence. In this case there is one match (there is one line), represented by a thick green arrow. In order to see how closely the two base sequences match scroll down; you should see lines of base sequence (strings of A, C, G, and T) and hash-marks between them. Note the similarity in the DNA sequence between the query (the *ADE1* gene of the red yeast) and the DNA sequence found in the database (labeled "Sbjct"; this is the *ADE1* gene of the white [wild-type] yeast). Note that one of the identifiers of the subject sequence is *ADE1* and that the description includes the name of an enzyme that we have seen before and the comment that the gene is required for the synthesis of adenine. The query sequence starts with base number (coordinate) 1 and goes to coordinate 426; the "Sbjct" sequence starts at coordinate 399 and goes to coordinate 824. Note that each line of sequence alternates between the query and the sbjct sequence. A hash-mark between the query and sbjct sequences indicates that the base sequence is identical between the query and the sbjct at that coordinate. The top portion of the screen will look like this:



>YAR015W ADE1 SGDID:S00000070 Chr I from 169370-170290.  
 Verified ORF, "N-succinyl-5-aminimidazole-4-carboxamide ribotide (SAICAR) synthetase, required for 'de novo' purine nucleotide biosynthesis; red pigment accumulates in mutant cells deprived of adenine"  
 [ Retrieve Sequence / ORF Map / Genome Browser / SGD Locus page ]

Length = 921

Score = 324.3 bits (2121), Expect = 6.3e-92, P = 6.3e-92  
 Identities = 425/426 (99%), Frame = +1 / +1

```
Query: 1  CGTAAAAACAGGTACTGTGCATGGTTTGAACAACCTCAAGGACTTAAAGAATCTCAAGA 60
          |||
Sbjct: 399 CGTAAAAACAGGTACTGTGCATGGTTTGAACAACCTCAAGGACTTAAAGAATCTCAAGA 458

Query: 61  GTTCCCAAGAACCAATCTTACCCCATCGACCAAGGCTGAACAAGGTGAACATGACGAAAA 120
          |||
Sbjct: 459 GTTCCCAAGAACCAATCTTACCCCATCGACCAAGGCTGAACAAGGTGAACATGACGAAAA 518
```

**d) But is it an exact match?** Recall that our hypothesis is that the difference between the white and red yeast strain is attributable to difference in the base sequence of the *ADE1* or *ADE2* genes. Such a difference could be fairly small so check carefully to see if the query sequence and the Sbjct sequence are exact matches.

i. Do you see any differences? Any difference would mean that the base sequence of the *ADE1* gene of the red yeast is different from the base sequence of the *ADE1* gene of the white yeast. If a difference was observed, what treatment of the yeast in week 1 led to this difference in the base sequence? As a reminder, the base sequence in the database (the "Sbjct") came from the standard (wild-type) white yeast.

**Summary of Steps 1-3.** At this point you have determined the sequence of bases that make up the *ADE1* gene from the red yeast, searched a database and compared the sequence of bases of the *ADE1* gene from the white yeast to the *ADE1* gene of the red yeast. In the next step you will use the base sequence of the *ADE1* gene from red yeast to determine the amino acid sequence of the protein encoded by the *ADE1* gene from the red yeast. You will then compare the amino acid sequences from the red and white yeasts.

**4) Determining the amino acid sequence of the Ade1 protein from the red yeast.** In this section you will use the program DNA Strider to determine the amino acid sequence of the protein encoded by the *ADE1* gene.

**a) What's in a DNA sequence?** Recall our definition of a gene: a gene is a segment of DNA that codes for a protein. As was the case last week, the computer shows you the sequence of the non-template strand of DNA, the strand that looks like the mRNA, except for the presence of T's instead of U's. This strand can be directly analyzed using a genetic code table to decipher what amino acids are encoded. The DNA code is written in groups of three bases. Note that the "reading frame" of the sentence is quite important. For instance, if I wrote a sentence of three letter words in the midst of a random assortment of letters it would be nonsense unless you started in the correct reading frame. The three reading frames of the sentence TETHECATATETHEBIGRATTT would be:

Reading frame (1) TET HEC ATA TET HEB IGR ATT T  
 (groups of three start at the first letter)

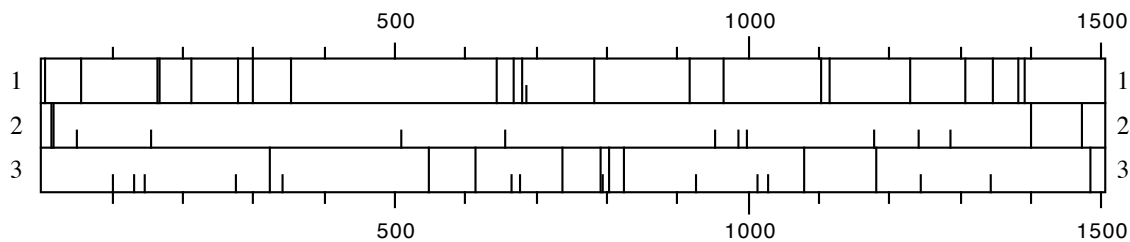
Reading frame (2) T ETH ECA TAT ETH EBI GRA TTT  
 (groups of three start at the second letter)

Reading frame (3) TE THE CAT ATE THE BIG RAT TT

(groups of three start at the third letter)

There is only one of three reading frames in which you will be able to decipher the meaning of this sentence. So in this example, unless you start reading in the “3<sup>rd</sup>” reading frame, this string of letters would have no meaning. Likewise, to decipher the amino acid sequence of a gene the information must be read in the correct “reading frame”. The computer program DNA Strider will examine each of the three possible reading frames to determine which one has a “sentence”. In this case a sentence means a long string of amino acids that are encoded without any “stop” codes. Recall from the homework assignment that you did last week that some 3-letter combinations represent a “stop”---that is, rather than an amino acid this stop signal results in the termination of translation. The codes for stop are TGA, TGG, and TAG.

**b) What and where is an ORF?** An “ORF” is an abbreviation for an open reading frame and is metaphorically the same as finding a sentence in the midst of a number of letters strung together. The computer program "DNA Strider" has tools designed for this purpose. Return to DNA Strider. Under the pull down “AA” (which stands for amino acids), select “ORF” and “3-phase”. This operation will “decode” (translate) your DNA sequence into an amino acid sequence by analyzing the nucleotide sequence in groups of three, noting any cases of stop codes. The map that you are currently looking at is plotting this translation in each of the three possible reading frames the presence of “periods” (stop codes; TGA, TGG, and TAG). Each of the three reading frames is indicated by a number next to a rectangle. The rectangle represents the DNA sequence. Wherever stop codes occur in the sequence they are plotted as a solid vertical line from the top to the bottom of the reading frame’s rectangle. An example of the output on a randomly chosen gene (i.e., not *ADE1*) is shown below. Your output should look similar but not identical to this:



An open reading frame would be a large area in which there are no long vertical lines. In the example provided above, the ORF is in the second reading frame. Do you see an ORF in the *ADE1* sequence? If you do, this means that this segment of DNA has the potential to code for a protein because there are no stops, just amino acids, encoded.

**c) The amino acid sequence of the protein encoded by the *ADE1* gene of the red yeast.** If you do see an ORF this indicates the potential for this sequence of bases to encode a protein. You will now determine the sequence of amino acids that make up the protein that is encoded by the ORF. The program DNA Strider will allow us to see this as well. Position the cursor between the first and second base of the sequence (between the C and the G) in the DNA Strider window. Highlight all of the bases starting from the G. Then under the “Conv” (Convert) pulldown select “Prot 5’ → 3’”. This is an abbreviation for “Protein sequence encoded by the DNA in the 5’ to 3’ direction”. A new window will appear that is filled with an amino acid sequence. This is the sequence of amino acids in the protein encoded by the red yeast *ADE1* DNA that you sequenced. The amino acids are given

in one-letter abbreviations. For example, the amino acid serine is abbreviated with the letter S and the amino acid leucine is abbreviated with the letter L. If you have succeeded then the last seven amino acids using the one letter abbreviation should be "TANKLNG".

*ii. Given what you learned in part 4(b), why was the cursor placed between the 1<sup>st</sup> and 2<sup>nd</sup> base of the sequence? Would the result have been different if you had placed the cursor between the second and third bases?[Also see example under part 4(a)].*

**5) Comparing the amino acid sequence encoded by the *ADE1* gene from the red yeast to the amino acid sequence encoded by the *ADE1* gene from the white yeast:** The above sequence of amino acids is not by itself very informative. It is more important to compare the amino acid sequence of this protein in the red yeast to the amino acid sequence in the standard white yeast. This can be done using a program that was used before, BLAST. BLAST can be used to search databases for either DNA sequences (as in step 3) or protein sequences (what we are about to do). To do this, highlight and copy the amino acid sequence in DNA Strider from step 4. Now go back to the <http://www.yeastgenome.org> web address and choose BLAST, or use your back button to back up one page. Delete any text that might be in the "Type or Paste Query Sequence" window. Next paste the amino acid sequence that you just copied into this window. You will now search the database for any protein from the standard (wild-type) white yeast whose amino acid sequence is identical or similar to the one that you just determined was encoded by the *ADE1* gene of the red yeast. Under the heading of "Choose the appropriate BLAST program" select "BLASTP". Under the next heading "Choose a Sequence Database," select "Open Reading Frames (DNA or Protein)". These steps allow you to compare the red yeast protein sequence encoded by the *ADE1* gene to all of the standard (wild-type) white yeast proteins in the database. Then select the "Run BLAST" option.

**Results of the search:** The results of this search are depicted in a way that is similar to step 3. The results are shown schematically below the query "number line" in the new window. The thick lines are color-coded to indicate the degree to which the query sequence and the target sequence match. In this case, there is only one "good" match shown in green—each blue line corresponds to an amino acid sequence whose similarity is so weak as to be of no or little significance. The green line runs the length of the "number line" indicating that the match extends for the entire length of the input sequence. The amino acid sequence of the Ade1 protein from the red yeast and the match to that sequence from the database are shown lower on the page---you can get there simply by clicking on the thick green line. At the top of this page is a title of the Subject sequence (the one that was found in the database). Notice that ADE1 is in the title (as well as a catalog number; this verifies that the sequence that has the best similarity to the Ade1 amino acid sequence from red yeast is the Ade1 protein from the white yeast). Just below the title of the Subject sequence you will see an alignment of letters, Scroll down until you see an alignment of letters, presented as sets with three lines of letters each. Again, each letter is the single letter abbreviation for an amino acid. Thus, there are 20 different letters that you could see. In this case, hash marks are not used to indicate amino acids that are identical between the two sequences. Instead the top line of each set of three is your query sequence – the protein encoded by the *ADE1* gene from the red yeast. The bottom line of each set of three is the "Sbjct" sequence, the protein encoded by the white yeast DNA in the database. The middle line indicates the amino acids that are shared by both the query and Sbjct sequences – i.e., if both Sbjct and query have a "V" (for the amino acid valine) then a

“V” will be shown in the middle line. Gaps in the middle line indicate that the two sequences do not match at this position.

*iii. **Are the two amino acid sequences identical or are there differences? If there are differences then address the following questions: what amino acid number in the chain is different? Provide the single letter abbreviation for the name of the amino acid that is present at that position in the ADE1 gene from the red yeast. In order to figure this out you need to go back to the program DNA Strider and under the pulldown "AA" select "Genetic Code". A table will appear that has the one letter designations and the corresponding amino acid. The homework that you did last week categorized mutations as "nonsense", "missense", "silent", or "frameshift". Which category does this fit into? Looking at the table of structures of the amino acids provided near the beginning of this manual, describe the difference in the amino acid that is altered—is it similar to, slightly different, or radically different in structure from the original amino acid? Thus is the mutation likely to harm the ability of the Ade1 protein to carry out its function? Explain. Is the difference in the amino acid sequence of the Ade1 protein in the red strain compared to the white strain a reflection of a difference in the DNA sequence between the ADE1 gene of the red strain compared to the white strain? Explain your answer.***

*If the two amino acid sequences are identical, is it likely that the difference in color between the red and white strains is attributable to a difference in the ADE1 gene or corresponding protein? Explain.*

**6) Examining the database for a human gene that encodes a (weakly) similar protein that carries out the same function.** We are using a lowly single-celled organism to learn about adenine biosynthesis--that is, how the critical compound adenine is made inside a cell. One reason to use such an organism is that it offers many advantages over more complex ones. For instance, it grows exceedingly fast. In 24 hours one cell will grow to more than 60,000 cells. This makes it simple to gather sufficient material on which to do experiments. In cases where one is examining different conditions for growth or absence of growth the experiment can be done in one day. Another advantage is the ability to treat the organism in a variety of ways. Consider the treatments you have done to the yeast during the last several weeks. It is unlikely that you could (or should) treat a vertebrate organism in the same way.

But do yeast cells have anything in common with cells from more complex organisms? We can make one attempt at answering this question by asking whether cells from humans have a protein that is similar to the protein encoded by the *ADE1* gene of yeast. If the answer to this question is “yes” then it seems reasonable to conclude that both yeast cells and human cells make their adenine in similar ways. Furthermore, this would then suggest that studying adenine biosynthesis in yeast is an adequate way to begin to understand adenine biosynthesis in humans.

In order to determine if human cells have a protein that is similar to the protein encoded by the *ADE1* gene of yeast we will search a database that has all of the amino acid sequences of proteins that have been determined by scientists thus far in any type of organism. In your web browser click on the link that says “SGD locus page,” just above the *ADE1* alignment. This should take you to the page titled ADE1/YAR015W that you saw last week. On the right side

of the page find the pull-down menu under “Comparison Resources.” Select “BLASTP at NCBI” and click the “View” button. A new page will load. This page will allow you to search a species-particular database for an amino acid sequence that is similar to the ADE1 protein of yeast. Like the earlier search of the yeast database, this site is using the program BLAST to find sequences similar to the Ade1 protein. Where the resulting page says “Choose search set” type “Homo sapiens” into the “Organism” window and then select the pulldown “*Homo sapiens* taxid:9606). Then click on the “Search” button. The page does a little rearranging at this point and it may take about 30 seconds for the search results to appear. The page may refresh one or more times. On the final results page most of the page background will be light blue.

Scroll down on the resulting page until you see the alignment of the yeast and the human protein. The top line of each set should be the yeast Ade1 protein sequence. The third line is the sequence of the human protein. The middle line has many blank spaces, these represent a place at which neither the human and yeast protein share similarity. Places in the middle line that have a character represent places in which there is identity between the human and yeast protein. A “+” character indicates a position in which the human and yeast protein have a similar but not identical amino acid. So the areas of the middle line with the highest density of characters are the areas that are the most similar between the yeast and human proteins. Note that the degree of similarity is not the same throughout the protein – the beginning (N-terminus), for example, is fairly similar between the two sequences (GK ++++YE+). The most similar regions are usually especially important in the function of the protein.

Examine the comparison. Near the end of the yeast protein amino acid sequence of *ADE1* there is a sequence that is DKQFLRD. There is a similar sequence in the human protein DKQSYRD. One could argue that the similarity between the yeast and human proteins is due to chance. But notice the “expected value” associated with the match between the yeast and human amino acid sequence. This value estimates the probability that such an amino acid sequence match could be found by random chance. This value is fairly low ( $4e^{-08}$ ; = one in 25 million) suggesting that there is some significance to this match.

Click on the link to the upper left of the alignment where it is written “[GENE ID: 10606 PAICS](#).” This will take you to a page describing the human protein sequence. Notice the description of this protein includes “phosphoribosylaminoimidazole succinocarboxamide synthetase.” (Also there is a summary paragraph that says that this is involved in purine biosynthesis). This is an alternative name for the enzyme SAICAR synthetase---recall from last week that the name of the enzyme encoded by the yeast *ADE1* gene is known as SAICAR synthetase but that it was also known as phosphoribosylaminoimidazole succinocarboxamide synthase. Thus we have found that the protein from humans that is similar in acid sequence to the yeast Ade1 protein also has the same function – synthesizing the intermediate SAICAR in the pathway that manufactures adenine. Thus experiments investigating the yeast enzyme might tell us important things about the human enzyme.

**You have now completed an analysis of the *ADE1* gene from the red yeast and compared it to the *ADE1* gene from the white yeast. You have identified the base sequence of the *ADE1* gene from the red yeast, you have identified the *ADE1* base sequence from the standard (wild-type) white yeast, and you have compared these two base sequences. You have also determined the sequence of amino acids that is encoded by the *ADE1* gene from the red yeast and compared it to the amino acid sequence of the protein encoded by the *ADE1* gene from the white yeast. Finally, you have seen that a protein similar to the Ade1 protein, SAICAR synthetase, also can be found in humans. *See the section at the end of the lab about what to discuss with your partner before you leave the lab.***

**Below are the objectives and procedures for studying the *ADE2* gene.**

- 1) Examining the primary data of the *ADE2* gene**
- 2) Converting the *ADE1* sequence to that on the complementary DNA strand.**
- 3) Comparing the nucleotide sequence of the *ADE2* gene from the red yeast to the *ADE2* gene from the white yeast**
- 4) Determining the amino acid sequence of the protein encoded by the *ADE2* gene**
- 5) Comparing this amino acid sequence to the sequence from the standard (wild-type) white strain**
- 6) Examining the database for a human gene that encodes a (weakly) similar protein that carries out the same function**

**1) Examining the primary data of the *ADE2* gene**

Open up the file labeled “ADE2mr.seq” using the program “Editview”. To do this, double-click on the “Editview” icon, and under the pulldown heading of “File” select “Open”. Then select the file name “ADE2mr.seq”. This computer file contains the sequence data we obtained from the PCR product generated with the *ADE2* primer set in the previous lab session. This is the sequence of bases corresponding to the *ADE2* gene from the red yeast strain. At this point you should see peaks of four different colors. There is also an X-axis that consists of a time unit of measurement, a nucleotide number, and a nucleotide (designated A, C, G, or T). There are a few things to note:

- i) at many points there is only one peak and that the color of the peak determines the nucleotide indicated on the X-axis.
- ii) at some points, there is more than one peak (this sequence actually has very few cases where this is true). This is due to technical problems and makes it difficult or impossible to predict the base that is at that position on the DNA strand. This is why there is occasionally a “N” instead of an A, C, G, or T. “N” designates that the identity of the base at this position was not unambiguously determined.
- iii) it is technically more difficult to determine the base sequence near the beginning. In this case, the sequence is unreliable for the first few bases (for example, after the A base at position number 5 there should be a C, but the blue C peak is obscured by the larger green A peak). It is also technically difficult to determine the base sequence much more than about 500 nucleotides from the start. Therefore we will limit our analysis to the base sequence between position 23 and 578.

In the bottom left corner of the window there are four icons. One of them reads "CATG". Click once on this icon. This changes the information that you were looking at into text only, rather than the multicolored peaks with the various other annotations. Notice all of the N's near the end of the text string. Highlight the characters between position 23 and 578 (from GTCCA to GGGTA) and then select "copy" from the "Edit" pull-down menu. You will now move this sequence of bases into a program that allows a search of the *Saccharomyces* database for any similar sequences.

- 2) Converting the *ADE2* sequence to that on the complementary DNA strand.** Remember that DNA has two strands of nucleotides that run in opposite directions and are complementary. The sequence you have shows only one strand of the DNA, but because of the strict base pairing rules

you can derive what the other strand must look like. For example, if the sequence of one strand is 5' CCACGT 3' then the sequence of the complementary strand must be 5' ACGTGG 3'. As luck would have it, the sequence you have received is actually the sequence of the gene on the template strand, read from the end of the gene to the beginning. In other words, the sequence is in reverse and complementary to the mRNA sequence that would be used to make a protein. In order to make the analysis of the *ADE2* gene easier, we need to turn the sequence around so that it reads like the other DNA strand (called the "non-template strand"). To do this we will use a program called DNA Strider. We could do this manually, but it would be very tedious to write out the complement of hundreds of bases.

Open DNA Strider by double-clicking on the DNA Strider icon in the "Applications" folder or on the CD. Note that it appears that nothing has happened because there is no open window but the program is open. There is now a menu bar that should not look familiar. Under the pull-down "File" select "New" and under "New" select "DNA". A window will open into which you will paste the DNA sequence of the *ADE2* gene from the red yeast. Paste what you had copied in step 1 into this new window by selecting "Paste" from the "Edit" menu or by holding down the apple and the "V" keys. You have now transferred the base sequence from the program "Editview" into the program "DNA Strider."

To convert the base sequence into the complementary base sequence, highlight the base sequence that is in the DNA Strider window. Then under the "Conv" (Convert) pulldown, select "antiparallel". Note that the new sequence that is generated is simply the sequence of the second strand of the original DNA displayed 5' -> 3'. Now select this new sequence and select "copy" from the Edit menu.

**3) Comparing the sequence of the *ADE2* gene from the red yeast to the sequence of the *ADE2* gene from the white yeast.** In the following steps you will search SGD, the database that has the DNA sequence of all genes from the standard (wild-type) white yeast that you used last week, to see if any have the same, or similar, sequence of bases as the sequence you have determined for the *ADE2* gene from the red yeast.

**a) Accessing the database via the internet**

Open Netscape and type in the following address

<http://www.yeastgenome.org>

At this website, the entire base sequence of the standard (wild-type) white strain of *Saccharomyces cerevisiae* is stored. This represents about 10 million base pairs. Nested somewhere in this 10 million base pairs is the sequence for the *ADE2* gene (only about 2000 base pairs long). At the top of this web page within the beige bar select "BLAST". This program will search the entire sequence of bases of the *Saccharomyces cerevisiae* in attempt to find some sequence that is either identical or similar to the "query" sequence (BLAST is an acronym for Basic Local Alignment Sequencing Tool). The query sequence is that which you type or paste into the window, as in the following step. In this case, the query sequence is the *ADE2* base sequence from the red yeast strain. Under the heading Choose One or More Sequence Datasets you should select "Open Reading Frames (DNA or Protein)".

**b) Searching the DNA database:** Now place the cursor into the window above which is labeled "Type or Paste a Query Sequence" and paste the sequence that you copied in step 1. Use



either the Paste command from the Edit menu or hold down the apple and “V” keys to paste. Now choose “Open Reading Frames (DNA or Protein)” under the heading Choose one or more Sequence Datasets. Next select the button “Run WU-BLAST”. The query sequence (the *ADE2* sequence from the red yeast strain) is now being compared to all 10 million base pairs in the database of the standard (wild-type) white yeast strain to find the sequence or sequences that are most similar.

- c) **Matches to the query DNA sequence:** After a short time you will see the results of this search although it will be difficult to interpret. Near the top of the page you will see what looks like a number line. This line represents the sequence you put in, with the numbers indicating the nucleotide number. Under this numbered line you will see one or more colored thick lines (actually they are arrows). Each thick line represents a match to the query sequence. In this case there is one match (there is one line), represented by a thick purple arrow. If you mouse over this line you will see a label “YOR128C ADE2 SGDID:S000005654...”. These characters are simply a database identifier tag, like a catalog number for an item in a catalog. Note that among the tags is the name “ADE2”. As the legend in the box indicates, the line is color-coded to signify how closely your query sequence matches the sequence found in the database. In addition, there is a p value in the label you just read ( $p=5.8e-122$ ). This p value is a measurement of how closely the two DNA sequences match; the higher the negative exponent the better the match. This is a very close match. In order to see how closely the two base sequences match scroll down; you should see lines of base sequence (strings of A, C, G, and T) and hash-marks between them. Note the similarity in the DNA sequence between the query (the *ADE2* gene of the red yeast) and the DNA sequence found in the database (labeled "Sbjct"; this is the *ADE2* gene of the white [wild-type] yeast). The query sequence starts with base number (coordinate) 1 and goes to coordinate 556; the "Sbjct" sequence starts at coordinate 280 and goes to coordinate 835. Note that each line of sequence alternates between the Query and the Sbjct sequence. A hash-mark between the Query and Sbjct sequences indicates that the base sequence is identical between the Query and the Sbjct at that coordinate. The top portion of the screen will look like this:

```
>YOR128C ADE2 SGDID:S000005654 Chr XV from 566192-564477, reverse complement,
Verified ORF, "Phosphoribosylaminoimidazole carboxylase, catalyzes a step in the 'de novo' purine nucleotide biosynthetic pathway; red pigment accumulates in mutant cells deprived"
[ Retrieve Sequence / ORF Map / Genome Browser / SGD Locus page ]
```

Length = 1,716

Score = 423.2 bits (2780), Expect = 5.8e-122, P = 5.8e-122  
Identities = 556/556 (100%), Frame = +1 / +1

```
Query: 1 TACCCTTCCTCCAGAAACAATCAGATTGATACAAGACAAATATATTCAAAAAGAGCATTTA 60
|||||
Sbjct: 280 TACCCTTCCTCCAGAAACAATCAGATTGATACAAGACAAATATATTCAAAAAGAGCATTTA 339
```

```
Query: 61 ATCAAAAATGGTATAGCAGTTACCCAAAGTGTTCCTGTGGAAACAAGCCAGTGAGACGTCC 120
|||||
Sbjct: 340 ATCAAAAATGGTATAGCAGTTACCCAAAGTGTTCCTGTGGAAACAAGCCAGTGAGACGTCC 399
```

- d) **But is it an exact match?** Recall that our hypothesis is that the difference between the white and red yeast strain is attributable to difference in the base sequence of the *ADE1* or *ADE2* genes. Such a difference could be fairly small so check carefully to see if the Query sequence and the Sbjct sequence are exact matches.

- i. Do you see any differences? Any difference would mean that the base sequence of the *ADE2* gene of the red yeast is different from the base sequence of the *ADE2* gene of the white yeast. If a difference was observed, what treatment of the yeast in week 1 led to this difference in the base sequence? As a reminder, the base sequence in the database (the “Sbjct”) came from the standard (wild-type) white yeast.

**Summary of Steps 1-3.** At this point you have determined the sequence of bases that make up the *ADE2* gene from the red yeast, searched a database and compared the sequence of bases of the *ADE2* gene from the white yeast to the *ADE2* gene of the red yeast. In the next step you will use the base sequence of the *ADE2* gene from red yeast to determine the amino acid sequence of the protein encoded by the *ADE2* gene from the red yeast. You will then compare the amino acid sequences from the red and white yeasts.

**4) Determining the amino acid sequence of the protein encoded by the *ADE2* gene from the red yeast.** In this section you will use the program DNA Strider to determine the amino acid sequence of the protein encoded by the *ADE2* gene.

a) What's in a DNA sequence? Recall our definition of a gene: a gene is a segment of DNA that codes for a protein. As was the case last week, the computer shows you the sequence of the non-template strand of DNA, the strand that looks like the mRNA, except for the presence of T's instead of U's. This strand can be directly analyzed using a genetic code table to decipher what amino acids are encoded. The DNA code is written in groups of three bases. Note that the "reading frame" of the sentence is quite important. For instance, if I wrote a sentence of three letter words in the midst of a random assortment of letters it would be nonsense unless you started in the correct reading frame. The three reading frames of the sentence TETHECATATETHEBIGRATTT would be:

Reading frame (1) TET HEC ATA TET HEB IGR ATT T  
(groups of three start at the first letter)

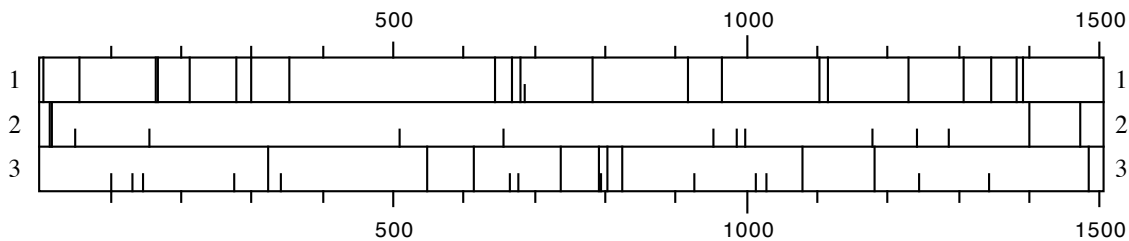
Reading frame (2) T ETH ECA TAT ETH EBI GRA TTT  
(groups of three start at the second letter)

Reading frame (3) TE THE CAT ATE THE BIG RAT TT  
(groups of three start at the third letter)

There is only one of three reading frames in which you will be able to decipher the meaning of this sentence. So in this example, unless you start reading in the "3<sup>rd</sup>" reading frame, this string of letters would have no meaning. Likewise, to decipher the amino acid sequence of a gene the information must be read in the correct "reading frame". The computer program DNA Strider will examine each of the three possible reading frames to determine which one has a "sentence". In this case a sentence means a long string of amino acids that are encoded without any "stop" codes. Recall from the homework assignment that you did last week that some 3-letter combinations represent a "stop"---that is, rather than an amino acid this stop signal results in the termination of translation. The codes for stop are TGA, TGG, and TAG.

**b) What and where is an ORF?** An "ORF" is an abbreviation for an open reading frame and is metaphorically the same as finding a sentence in the midst of a number of letters strung together. The computer program "DNA Strider" has tools designed for this purpose. Return to DNA Strider. Under the pull down "AA" (which stands for amino acids), select "ORF" and "3-phase". This operation will "decode" (translate) your DNA sequence into an amino acid

sequence by analyzing the nucleotide sequence in groups of three, noting any cases of stop codes. The map that you are currently looking at is plotting this translation in each of the three possible reading frames the presence of “periods” (stop codes; TGA, TGG, and TAG). Each of the three reading frames is indicated by a number next to a rectangle. The rectangle represents the DNA sequence. Wherever stop codes occur in the sequence they are plotted as a solid vertical line from the top to the bottom of the reading frame’s rectangle. An example of the output on a randomly chosen gene (i.e., not *ADE2*) is shown below. Your output should look similar but not identical to this:



An open reading frame would be a large area in which there are no long vertical lines. In the example provided above, the ORF is in the second reading frame. Do you see an ORF in the *ADE2* sequence? If you do, this means that this segment of DNA has the potential to code for a protein because there are no stops, just amino acids, encoded.

- c) **The amino acid sequence of the protein encoded by the *ADE2* gene of the red yeast.** If you do see an ORF this indicates the potential for this sequence of bases to encode a protein. You will now determine the sequence of amino acids that make up the protein that is encoded by the ORF. The program DNA Strider will allow us to see this as well. Position the cursor in front of the first base of the sequence (in front of the T of the very first base) in the DNA Strider window. Highlight all of the bases starting from the T. Then under the “Conv” (Convert) pulldown select “Prot 5’ → 3’”. This is an abbreviation for “Protein sequence encoded by the DNA in the 5’ to 3’ direction”. A new window will appear that is filled with an amino acid sequence. This is the sequence of amino acids in the protein encoded by the red yeast *ADE2* DNA that you sequenced. The amino acids are given in one-letter abbreviations. For example, the amino acid serine is abbreviated with the letter S and the amino acid leucine is abbreviated with the letter L. If you have succeeded then the last five amino acids using the one letter abbreviation should be "PHNSG".

*ii. Given what you learned in part4(b), why was the cursor placed in front of the first base of the sequence? Would the outcome have differed if the cursor had been placed between the first and second characters (the T and the A)? [Also see example under part 4(a)].*

- 5) **Comparing the amino acid sequence encoded by the *ADE2* gene from the red yeast to the amino acid sequence of the protein encoded by the *ADE2* gene from the white yeast:** The above sequence of amino acids is not by itself very informative. It is more important to compare the amino acid sequence of this protein in the red yeast to the amino acid sequence in the standard white yeast. This can be done using a program that was used before, BLAST. BLAST can be used to search databases for either DNA sequences (as in step 4) or protein sequences (what we are about to do). To do this, highlight and copy the amino acid sequence in DNA Strider from step 3.

Now go back to the <http://www.yeastgenome.org> web address and choose BLAST, or use your back button to back up one page. Delete any text that might be in the “Type or Paste Query Sequence” window. Next paste the amino acid sequence that you just copied into this window. You will now search the database for any protein from the standard (wild-type) white yeast whose amino acid sequence is identical or similar to the one that you just determined was encoded by the *ADE2* gene of the red yeast. Under the heading of “Choose the appropriate BLAST program” select “BLASTP”. Under the next heading “Choose a Sequence Database,” select “Open Reading Frames (DNA or Protein)”. These steps allow you to compare the amino acid sequence of the protein encoded by the *ADE2* gene of the red yeast to all of the standard (wild-type) white yeast proteins in the database. Then select the "Run WU-BLAST" option.

**Results of the search:** The results of this search are depicted in a way that is similar to step 3. The results are shown schematically below the query “number line” in the new window. The thick lines are color-coded to indicate the degree to which the query sequence and the target sequence match. In this case, there is only one “good” match shown in green (labeled YOR128C)—each blue line corresponds to an amino acid sequence whose similarity is so weak as to be of no or little significance. The amino acid sequence of the Ade2 protein from the red yeast and the match to that sequence from the database are shown lower on the page—you can get there simply by clicking on the thick green line. At the top of this page is a title of the Subject sequence (the one that was found in the database). Notice that *ADE2* is in the title (as well as a catalog number; this verifies that the sequence that has the best similarity to the Ade2 amino acid sequence from red yeast is the Ade2 protein from the white yeast). Just below the title of the Subject sequence you will see an alignment of letters presented as sets with three lines of letters each. Again, each letter is the single letter abbreviation for an amino acid. Thus, there are 20 different letters that you could see. In this case, hash marks are not used to indicate amino acids that are identical between the two sequences. Instead the top line of each set of three is your query sequence – the protein encoded by the *ADE2* gene from the red yeast. The bottom line of each set of three is the “Sbjct” sequence, the protein encoded by the white yeast DNA in the database. The middle line indicates the amino acids that are shared by both the query and Sbjct sequences – i.e., if both Sbjct and query have a “V” (for the amino acid valine) then a “V” will be shown in the middle line. Gaps in the middle line indicate that the two sequences do not match at this position.

*iii. Are the two amino acid sequences identical or are there differences? If there are differences then address the following questions: what amino acid number in the chain is different? Provide the single letter abbreviation for the name of the amino acid that is present at that position in the ADE2 gene from the red yeast. In order to figure this out you need to go back to the program DNA Strider and under the pulldown "AA" select "Genetic Code". A table will appear that has the one letter designations and the corresponding amino acid. The homework that you did last week categorized mutations as "nonsense", "missense", "silent", or "frameshift". Which category does this fit into? Looking at the table of structures of the amino acids provided near the beginning of this manual, describe the difference in the amino acid that is altered—is it similar to, slightly different, or radically different in structure from the original amino acid? Thus is the mutation likely to harm the ability of the Ade2 protein to carry out its function? Explain. Is the difference in the amino acid sequence of the Ade2 protein in the red strain compared to the white strain a reflection of a difference in the DNA sequence between the ADE2 gene of the red strain compared to the white strain? Explain your answer.*

If the two amino acid sequences are identical, is it likely that the difference in color between the red and white strains is attributable to a difference in the ADE2 gene or corresponding protein? Explain.

- 6) Examining the database for a human gene that encodes a (weakly) similar protein that carries out the same function.** We are using a lowly single-celled organism to learn about adenine biosynthesis---that is, how the critical compound adenine is made inside a cell. One reason to use such an organism is that it offers many advantages over more complex ones. For instance, it grows exceedingly fast. In 24 hours one cell will grow to more than 60,000 cells. This makes it simple to gather sufficient material on which to do experiments. In cases where one is examining different conditions for growth or absence of growth the experiment can be done in one day. Another advantage is the ability to treat the organism in a variety of ways. Consider the treatments you have done to the yeast during the last several weeks. It is unlikely that you could (or should) treat a vertebrate organism in the same way.

But do yeast cells have anything in common with cells from more complex organisms? We can make one attempt at answering this question by asking whether cells from humans have a protein that is similar to the protein encoded by the *ADE2* gene of yeast. If the answer to this question is “yes” then it seems reasonable to conclude that both yeast cells and human cells make their adenine in similar ways. Furthermore, this would then suggest that studying adenine biosynthesis in yeast is an adequate way to begin to understand adenine biosynthesis in humans.

In order to determine if human cells have a protein that is similar to the protein encoded by the *ADE2* gene of yeast we will search a database that has all of the amino acid sequences of proteins that have been determined by scientists thus far in any type of organism. In your web browser click on the link that says “SGD locus page,” just above the *ADE2* alignment. This should take you to the page titled ADE2/YOR128C that you saw last week. On the right side of the page find the pull-down menu under “Comparison Resources.” Select “BLASTP at NCBI” and click the “View” button. The NCBI database holds sequences from all types of organisms. Like the earlier search of the yeast database, this site will use the program BLAST to find sequences similar to the Ade2 protein. To limit our search for similar proteins to humans, under “Organism” in the “Choose Search Set” section type “Homo sapiens” and then select the pulldown “*Homo sapiens* (taxid:9606).” Now go to the line “Filters and Masking” and click on the box that has a check mark to deselect this option. Then hit the blue “BLAST” button at the bottom of the page. The search is now on for any amino acid sequence from *Homo sapiens* that bears a resemblance to the amino acid sequence encoded by *ADE2*. The page will refresh one or more times. The final page will have a background that is mostly light blue.

When the search is finished a new page will appear. Scroll down the page until the heading “Alignments.” If all went well you should see the top match with the identifier “Gene ID10606 PAICS”. Click on this link and you will learn about the human gene that has the highest similarity to the yeast *ADE2* gene. The title of this page includes the text “phosphoribosylaminoimidazole carboxylase.” Recall from last week that the name of the

enzyme encoded by the yeast *ADE2* gene was phosphoribosylaminoimidazole carboxylase or AIR carboxylase. Thus we have found that the protein from humans that is most similar in acid sequence to the yeast Ade2 protein also has the same function – adding a carboxyl group onto an intermediate (phosphoribosylaminoimidazole) in the pathway that manufactures adenine. Also notice under the Summary comments that this human enzyme seems to also participate in purine biosynthesis. Thus experiments investigating the yeast enzyme might tell us important things about the human enzyme.

Hit the Back button on your browser to return to the alignment page. Scroll down on the resulting page until you see the alignment of the yeast and the human protein. The top line of each set should be the query sequence – the yeast Ade2 protein sequence. The third line of each set is the sequence of the human protein. The middle line contains some letters and some plus signs. Note that the letters appear at each position where the two sequences are identical. Although it is less obvious unless you know the structures of the amino acids, a “+” appears where the amino acids are similar to each other in structure – an L and a V, for example. So the areas with the highest density of letters and + symbols are the areas that are the most similar between the yeast and human proteins.

Examine the comparison. One could argue that the similarity between the yeast and human proteins is due to chance. But notice the “expect value” associated with the match between the yeast and human amino acid sequence. This value estimates the probability that such an amino acid sequence match could be found by random chance. This value (1e-04) is fairly low suggesting that there might be some significance to this match. Also note that the degree of similarity is not the same throughout the protein – the region ++MGS-SDL near the beginning of the alignment, for example, is quite similar between the two sequences. The most similar regions are usually especially important in the function of the protein.

***Before you leave the lab, exchange data with your lab partner by filling out the sheet on the next page and having the instructor check it. Having data for both the *ADE1* and *ADE2* genes will be crucial in order to write the lab report.***

**SUMMARY: You and your partner have examined the base sequence of two genes *ADE1* and *ADE2* from the red yeast. The reason that you examined these two genes is that previous research has indicated that changes in either of these two genes can give rise to yeast cells that have a red pigment. Could this be true in the case of the red yeast that you observed after ultraviolet light treatment? You have compared the sequence of the *ADE1* gene and the *ADE2* gene from the red yeast to the corresponding genes from the white yeast. Were there differences between the red and white yeast in either *ADE1* or *ADE2*? In both *ADE1* and *ADE2*? You have also examined the sequence of amino acids encoded by the *ADE1* and *ADE2* genes from the red yeast and compared these sequences to the corresponding amino acid sequences from the white yeast. Were there amino acid sequence differences between proteins encoded by the *ADE1* and *ADE2* genes in red and white yeast?**

**QUESTIONS FOR DISCUSSION: Before you leave the lab, please meet with your lab partner to answer the following questions. Write the answers below and keep this page for your future reference after showing it to the instructor or TA.**

- 1) You should have seen some difference between the red yeast DNA and the white yeast DNA in the database. What difference did you see? Write here:

The gene that had the difference (was it *ADE1* or *ADE2*):

The position of the base that was changed and what that change was:

Base number:

Base seen at that position in white yeast DNA in the database.

Base seen at that position in red yeast DNA that you sequenced:

- 2) Was there a corresponding alteration in the protein encoded by the altered gene? What difference did you see? Write here:

The position of the amino acid that was altered and the type of alteration:

Amino acid number:

Amino acid seen at that position in white yeast in the database:

Amino acid seen at that position in the red yeast:

Nature of the mutation (silent, missense, nonsense, frameshift):

Instructor signature: \_\_\_\_\_

## Lab assignment

**Due at the start of lab in week 6**

**10 points**

Using a word processor (no handwritten assignments will be accepted), address each of the following questions. Each student will have to address these questions for both *ADE1* and *ADE2*. Because during the computer lab you only examined one of these two genes you will have to confer with your lab partner to answer all of these questions. These questions were taken from the text of the lab instructions. Look there to find the proper context of the question. This assignment is valued at 10 points and is due at the start of lab in week 6.

### For the agarose gel:

- i. Estimate the size of your DNA fragments generated by PCR. How do these sizes compare to the sizes of the *ADE1* and *ADE2* genes from white yeast that you calculated in the computer lab last week? Keep in mind that sizes estimated from gel electrophoresis are not exact, and that sizes from sequencing are exact.
- ii. Compare the distance traveled by the DNA from each of the two PCR samples. Additionally, compare these distances with the distance traveled by the DNA you recovered from the red yeast in the third week of lab. Based on the gel, are the *ADE* genes smaller or larger than the total yeast DNA pieces you isolated in the third week of lab? Does this make sense?

### For *ADE1*:

- i. Do you see any differences? Any difference would mean that the base sequence of the *ADE1* gene of the red yeast is different from the base sequence of the *ADE1* gene of the white yeast. If a difference was observed, what treatment of the yeast in week 1 led to this difference in the base sequence?
- ii. Given what you learned in part 4(b), why was the cursor placed between the 1<sup>st</sup> and 2<sup>nd</sup> base of the sequence? Would the result have been different if you had placed the cursor between the second and third bases? [Also see example under part 4(a)].
- iii. Are the two amino acid sequences identical or are there differences? If there are differences then address the following questions: what amino acid number in the chain is different? Provide the single letter abbreviation for the name of the amino acid that is present at that position in the *ADE1* gene from the red yeast. In order to figure this out you need to go back to the program DNA Strider and under the pulldown "AA" select "Genetic Code". A table will appear that has the one letter designations and the corresponding amino acid. The homework that you did last week categorized mutations as "nonsense", "missense", "silent", or "frameshift". Which category does this fit into? Looking at the table of structures of the amino acids provided near the beginning of this manual, describe the difference in the amino acid that is altered—is it similar to, slightly different, or radically different in structure from the original amino acid? Thus is the mutation likely to harm the ability of the *Ade1* protein to carry out its function? Explain. Is the difference in the amino acid sequence of the *Ade1* protein in the red strain compared to the white strain a reflection of a difference in the DNA sequence between the *ADE1* gene of the red strain compared to the white strain? Explain your answer.



If the two amino acid sequences are identical, is it likely that the difference in color between the red and white strains is attributable to a difference in the ADE1 gene or corresponding protein? Explain.

## **For ADE2**

- i. Do you see any differences? Any difference would mean that the base sequence of the ADE2 gene of the red yeast is different from the base sequence of the ADE2 gene of the white yeast. If a difference was observed, what treatment of the yeast in week 1 led to this difference in the base sequence?
- ii. Given what you learned in part 4(b), why was the cursor placed between the 1<sup>st</sup> and 2<sup>nd</sup> base of the sequence? Would the result have been different if you had placed the cursor between the second and third bases? [Also see example under part 4(a)].
- iii. Are the two amino acid sequences identical or are there differences? If there are differences then address the following questions: what amino acid number in the chain is different? Provide the single letter abbreviation for the name of the amino acid that is present at that position in the ADE2 gene from the red yeast. In order to figure this out you need to go back to the program DNA Strider and under the pulldown "AA" select "Genetic Code". A table will appear that has the one letter designations and the corresponding amino acid. The homework that you did last week categorized mutations as "nonsense", "missense", "silent", or "frameshift". Which category does this fit into? Looking at the table of structures of the amino acids provided near the beginning of this manual, describe the difference in the amino acid that is altered—is it similar to, slightly different, or radically different in structure from the original amino acid? Thus is the mutation likely to harm the ability of the Ade2 protein to carry out its function? Explain. Is the difference in the amino acid sequence of the Ade2 protein in the red strain compared to the white strain a reflection of a difference in the DNA sequence between the ADE2 gene of the red strain compared to the white strain? Explain your answer.

If the two amino acid sequences are identical, is it likely that the difference in color between the red and white strains is attributable to a difference in the ADE2 gene or corresponding protein? Explain.