1. We discussed indirect end labeling methods. Here is a problem to get you to try to figure out this method:

Direct end labeling is commonly used for instance in DNA footprinting (see Alberts textbook for description). In this method a DNA molecule is labeled at one end, a protein is bound at a specific site in the DNA fragment and an enzyme is used to cut the DNA. The enzyme used cuts naked DNA but not DNA covered by the binding of the protein, producing a “footprint” of protected DNA where the enzyme does not cut. When one runs the cut DNA on a gel, one sees a range of fragments up to the size corresponding to the beginning of the footprint, but a gap over the sizes corresponding to where the protein is bound.

In the case of trying to infer where proteins are bound or not bound in chromatin within the nucleus we would like to adapt a similar technique. There are two problems we have to solve though to be able to do this.

i) We don’t have a defined DNA fragment- ie we don’t have a defined end in the DNA to measure distances from.

Solution: We can create defined ends by using restriction enzymes to cut the DNA. The problem is that the restriction sites may be covered by chromosomal proteins and not accessible for cutting by the restriction enzyme. We solve this problem by reversing the steps in the procedure. First we “footprint” by exposing the nucleus to the nuclease used to probe chromatin structure. THEN we purify the DNA, removing all proteins from the DNA. NOW we use the restriction enzymes to cut the naked DNA at the specific restriction sites.

ii) These ends must be labeled to be able to see them on a gel. We can’t just label the whole DNA fragment. We need to label a specific end so we can measure distances from this end. But we can’t add a direct end label to the piece of DNA we want to study- it’s in the middle of a nucleus.

Solution: We use an “indirect” end labeling method. We detect the DNA fragments that arise from a specific fragment of DNA using a short probe (~100 bp) from the end of this fragment in a Southern blot. This allows us to specifically label and visualize the pieces of DNA from a specific region near a specific gene.

Let’s try this in action.
Below is the restriction map of a region near the X gene. Small arrows show EcoR1 sites. The site of transcription of gene X is marked.
The student does the following experiment:

a. Isolate Nuclei  
b. Treat for a short time with micrococcal nuclease  
c. Stop reaction of micrococcal nuclease, purify DNA from the digested nuclei  
d. Cut purified DNA with EcoR1  
e. Run gel- stain DNA with EtBr  
f. Blot DNA to membrane, do Southern Blot using a probe corresponding to the rectangular box near the start of transcription of the X gene. This probe is only about 100 bp in length.

Below is the DNA gel that is seen using EtBr for a stain.

Black bands correspond to 200, 400, 600, 800, … ladder

Southern Blot- a first band of 100 bp is observed, followed by a ladder of bands starting at 350 bp but with regular spacing of 200 bp. What I couldn’t easily draw in this picture is that beyond the first several bands of this ladder, the bands become less distinct.

The above experiment was done in a tissue where gene X is active. What is an explanation for the above results?

The first EtBr stained gel shows that the bulk chromatin in these cells is arranged in regularly spaced nucleosomes with a nucleosome repeat of 200 bp. The Southern
blot reveals a region containing regularly spaced bands separated by 200 bp (350, 550, 750, ...). Below this though there is a gap and then a 100 bp band.

The conclusion would be regularly spaced and phased nucleosomes positioned such that the linker DNA at 350 bp would be cut and then at 200 bp intervals after that. Between 100 and 350 bp there, too large to contain a single nucleosome, there is a region protected from micrococcal nuclease digestion. This implies some type of nonnucleosomal structure over part or all of this region. The 100 bp band is coming from cutting at 100 bp 3’ to the restriction site shown near the probe. This could be due to a nucleosome positioned over the restriction site but with its linker DNA at 100 bp from the site. A drawing of this is shown below:

```
    100 bp                       350 bp
```

```
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Restriction</td>
</tr>
<tr>
<td>Site</td>
</tr>
<tr>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>Phased Nucleosomes</td>
</tr>
</tbody>
</table>
```

In the restriction map drawing the restriction site near the probe is also located near the 5’ region of the gene X. Therefore the binding of some nonhistone proteins, presumably to a regulatory region, likely sets up phasing of nearby nucleosomes passively- ie because the nucleosomes now line up next to the nonnucleosomal structure.

Now the student repeats this experiment but in tissue where gene X is off. Again a similar EtBr gel is obtained but the Southern blot for the indirect end labeling experiment now shows a continuous smear instead.

What are the possible explanations now for observing a smear? How can the student distinguish between two different explanations?

The most likely explanation is that the smear is due simply to the presence of regularly spaced nucleosomes (as over the bulk chromatin) which are not phased. In theory the smear could also be due to irregularly spaced nucleosomes or no nucleosomes, although as discussed in class regularly spaced nucleosomes are the general rule over inactive genes.

One can simply leave out the restriction digest in the above procedure. Then if there are regularly spaced nucleosomes over the region surrounding the restriction site one would see a typical micrococcal nuclease ladder. If one has no nucleosomes one would see a smear with relatively rapid kinetics of digestion. If one had irregularly spaced nucleosomes one would get a smear which stopped at the pause and or limit digest.
2. You are interested in both studying gene regulation and making money but not sure how to combine these two interests into one project. You decide that the answer is cosmetics and after several years you have discovered a gene involved in hair color. You believe that this gene, if expressed at high levels, could restore grey hair to its normal color.

You clone this “Hair” gene and spend another several years identifying its promoter and getting a transformation system for hair follicle cells working. Finally, you believe you have reached the stage where you can make your fortune.

You decide to offer a gene therapy treatment to people who want to reverse their grey hair back to normal. This depends on expressing the “Hair” gene in the hair follicle cells of people with grey hair.

You need to make a DNA construct to express the “Hair” gene in hair follicle cells. You have a choice for what enhancer to use. Many people use a strong viral enhancer sequence such as from SV-40 which is often used in expression vector constructs to get high levels of expression. The other possibility is to use identify the endogenous enhancer sequence for the “Hair” gene.

a. Which choice would you make? Why?
I’d use the regular enhancer to be sure that it would work in the hair follicle cells. Some viral enhancers are notorious for not working well in nonproliferating, differentiated cells.

b. You go ahead and manage to convince your Uncle to allow you to test the “Hair” expression vector on him (he has grey hair). Within the first few weeks he gets excited because he notices a few brown hairs. Several months go by and he is less impressed. He still only has a few light brown hair. However, you pull some of his grey hair out, find some hair follicle cells still attached and do PCR. You detect DNA from your vector from most of the hairs. But if you try RT-PCR you detect very little expression of the “Hair” cell”. If you do this with the light brown hairs you detect both “Hair” gene DNA by PCR and “Hair” message by RT-PCR. What do you think is going on? Does this surprise you?

It looks like the DNA is getting stably transfected but the expression is poor, with no detectable expression in most cases and low level expression in a few hairs. This is typical for most plasmid based transgenes. Even with known promoter and enhancer sequences expression is quite low and very dependent on the chromosomal location of integration.

c. You establish an ES cell system in which you can form “skin in a dish” complete with hair follicles. You use this system to investigate the gene regulation of the “Hair” gene. You decide to do indirect end labeling but using DNase I rather than the micrococcal nuclease as described in the previous problem.
You isolate nuclei, digest with DNase I, purify DNA, cut with EcoR1, then run a gel and do a Southern blot. You use as a probe the short DNA segment (~150 bp) shown in the restriction map of the “Hair” gene shown below.

The enhancer for the Hair gene is localized just upstream of restriction site (3). The short probe for the Southern blot that you use is located just 3’ of restriction site (1).

You observe the following picture for the EtBr gel of the DNA after the EcoR1 digestion (left) and then after the Southern blot, using the probe as shown above. Where is the DNase I hypersensitive site relative to restriction site (1)?

There is a DNase I hypersensitive site (HS) located ~2.5 kbp 3’ to restriction site 1 and roughly 20 kbp upstream of the “Hair” gene promoter.

d. You now include in your construct a small region including this DNase I HS near the probe in the map above. It is far from your gene but you observe a surprising result. Now he develops lots of dark brown hair with just a few interspersed grey hair. What do you think is the explanation?
This cis regulatory region is producing both higher expression plus a resistance to position effects given that expression is now seen in most hairs. This would be the expected behavior if the identified DNase I HS site corresponded to part of or an entire locus control region for the “Hair” gene.

e. Unfortunately over a longer period of 6-12 months, his hair gradually turns back to grey. What might be happening?
*The expression is most likely being reduced over time. This is still the typical behavior of transgenes in most gene therapy trials. Presumably the complete set of regulatory sequences required for stable maintenance of high expression levels (and presumably an epigenetic state compatible with high expression) is still not present in the transgene.*

f. If you were his uncle would you go through this again or try some hair coloring agents from the drug store?