7.03, 2006, Lecture 21
Eukaryotic Genes and Genomes II

In the last lecture we considered the structure of genes in eukaryotic organisms and went on to figure out a way to identify *S. cerevisiae* genes that are transcriptionally regulated in response to a change in environment. The ability to regulate gene expression in response to environmental cues is a fundamental requirement for all living cells, both prokaryote and eukaryote. We considered how many genes each organism has, about 4,000 for *E. coli*, 6,000 for yeast and a little over 20,000 for mouse and humans. But only a subset of these genes is actually expressed at any one time in any particular cell. For multicellular organisms this becomes even more apparent…it is obvious that skin cells must be expressing a different set of genes than liver cells, although of course there must be a common set of genes that are expressed in both cell types; these are often called **housekeeping** genes.

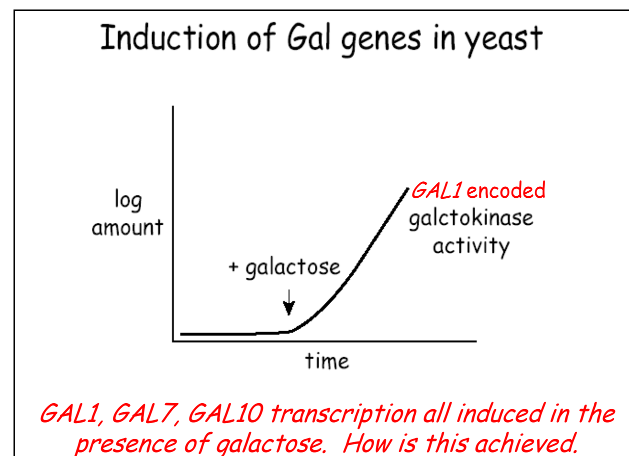
There are a number of ways that gene regulation in eukaryotes differs from gene regulation in prokaryotes.

- Eukaryotic genes are not organized into operons.
- Eukaryotic regulatory genes are not usually linked to the genes they regulate.
- Some of the regulatory proteins must ultimately be compartmentalized to the nucleus, even when signaling begins at the cell membrane or in the cytoplasm.
- Eukaryotic DNA is wrapped around nucleosomes

Today we will consider how one can use genetics to begin to dissect the mechanisms by which gene transcription can be regulated. For this we will take the example of the yeast GAL genes in *S. cerevisiae*.

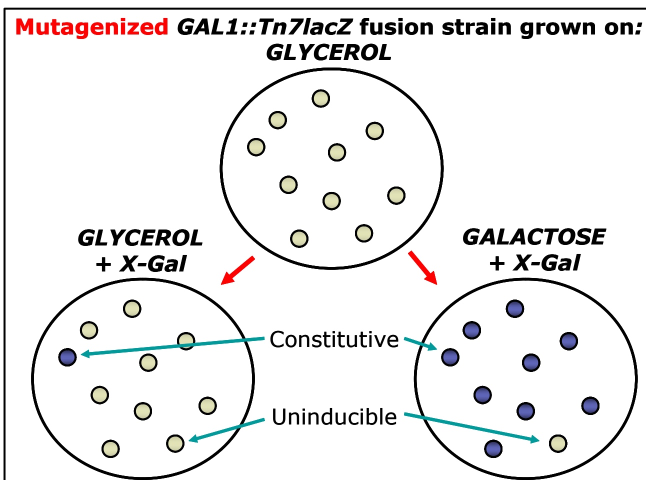
GALACTOSE METABOLISM IN YEAST

Reaction	Enzyme	Gene
D-galactose		
↓	Galactokinase	*GAL1*
D-galactose-1-phosphate		
↓	Galactose transferase	*GAL7*
UDP-D-galactose		
↓	Galactose epimerase	*GAL10*
UDP-D-glucose		
↓	UDP-glucose Phosphorylase	
D-glucose-1-phosphate		
↓	Phosphoglucomutase	
D-glucose-6-phosphate		
⋮		
GLYCOLYSIS		



GAL1, GAL7, GAL10 transcription all induced in the presence of galactose. How is this achieved.

Once a gene has been identified as being inducible under certain inducing conditions, in this case in the presence of galactose, we can begin to dissect the regulatory mechanism by isolating mutants; i.e., mutants that constitutively express the *GAL* genes even in the absence of galactose, and mutants that have lost the ability to induce the *GAL* genes in the presence of galactose. If we were studying galactose regulation today we would probably use a ***lacZ*** reporter system as we discussed in the last lecture. However, when the **Gal** regulatory system was fist genetically dissected, it was done by actually measuring the induction of **Gal1** encoded galactokiase activity, so this is how we will discuss the genetic dissection of the system.



Another approach is to simply measure galactokinase activity in the presence or absence of Galactose

Gal regulatory mutants

	galactokinase activity - galactose	galactokinase activity + galactose	Interpretation
Gal^+	−	+	
$Gal1^-$	−	−	$Gal1^-$ inactivates galactokinase
$Gal4^-$	−	−	$Gal4^-$ is uninducible
$Gal80^-$	+	+	$Gal80^-$ is constitutive
$Gal81^-$	+	+	$Gal81^-$ is constitutive

What we know is that **Gal4** mutants are uninducible and that **Gal80** and **Gal81** mutants constitutively express the **Gal1** galactokinase gene, along with the other **Gal** genes. Let's analyze each mutant in turn:

Gal4 mutant*:* It was first established that, like **Gal1**$^-$, the **Gal4**$^-$ mutant phenotype is recessive, because heterozygous diploids generated by mating **Gal4**$^-$ to wild type have normal regulation. It was then established that the mutation in the **Gal4**$^-$ strain lies in a new gene, and not simply in the *GAL1* galactokinase gene; **Gal1**$^-$ mutants don't express galactokinase activity in the presence of galactose, just as was seen for the **Gal4**$^-$ mutant. That **Gal1**$^-$ and **Gal4**$^-$ mutants have mutations in different genes was shown by complementation analysis, (diploids from mating **Mat**α **Gal4**$^-$ with **Mata Gal1**$^-$ behave like wild type) and the fact that the *GAL4* and

Analysis of Gal4

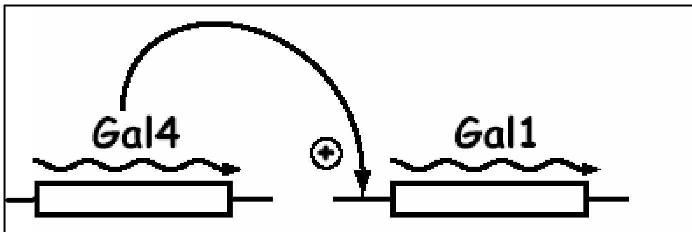
$Gal4^-$ is uninducible

$Gal4^+ / Gal4^-$ is regulated:
therefore $Gal4^-$ is recessive

MAT**a** $Gal1^-$ $Gal4^+$ x MATα $Gal1^+$ $Gal4^-$

1 :	4 :	1
Type 1	Type 2	Type 3
uninducible	uninducible	uninducible
uninducible	uninducible	uninducible
uninducible	uninducible	regulated (wt)
uninducible	regulated (wt)	regulated (wt)

GAL1 genes are unlinked was established by **tetrad analysis**. You should think about what the tetrads from the aforementioned diploids would look like.



Put together the simplest model is that **Gal4** is a positive regulator of **Gal1** (and the other Gal genes). The + sign indicates that Gal4 increases Gal expression, but does not indicate whether this is direct or indirect.

Gal80 mutant*:* The next useful regulatory mutant isolated was **Gal80⁻**, in which the **Gal1** encoded galactokinase is expressed even in the absence of galactose and is not further induced in its presence. Again, heterozygous diploids (**Gal80⁻/wt**) showed that **Gal80⁻** is recessive, **Tetrad analysis** showed that **Gal80** is not linked to **Gal1**, **Gal4** or any of the **Gal** genes.

Analysis of Gal80

Gal80⁻ is constitutive

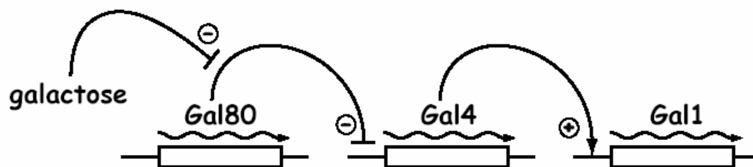
Gal80⁺ / Gal80⁻ is regulated: therefore Gal80⁻ is recessive

MATa Gal1⁻ Gal80⁺ x MATα Gal1⁺ Gal80⁻

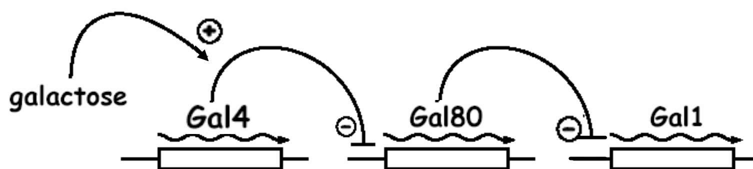
1	:	4	:	1
Type 1		Type 2		Type 3
uninducible		uninducible		uninducible
uninducible		uninducible		uninducible
constitutive		constitutive		regulated (wt)
constitutive		regulated (wt)		regulated (wt)

If a mutant **Gal80** results in constitutive **Gal1** expression, the simplest model is that **Gal80** negatively regulates the **Gal** genes. Since **Gal4** positively regulates, and **Gal80** negatively regulated **Gal1** expression, we have to figure out how these two gene products work together to achieve such regulation. Assuming that **Gal4** and **Gal80** act in series there are two formal possibilities:

Model 1



Model 2



Model 1 is that **Gal4** positively regulates **Gal1**, and that **Gal80** negatively regulates **Gal4**; the presence of galactose somehow inhibits **Gal80** function thus releasing **Gal4** to positively activate **Gal1** expression.

Model 2 is that **Gal80** negatively regulates **Gal1**, and **Gal4** negatively regulates **Gal80**; here the presence of galactose positively activates **Gal4** which in turn negatively regulates **Gal80**, thus relieving inhibition of **Gal1** expression.

We can distinguish between these two models by doing what's called an **epistasis** test to establish the epistatic relationship between **Gal4⁻** and **Gal80⁻**. This involves making a double **Gal4⁻** / **Gal80⁻** mutant strain. The phenotype of the double mutant will indicate which of the two models is most likely to be true....take a look at the two models to predict what phenotype the double mutant should have. For Model 1 the double mutant would become **uninducible**, for Model 2 the double mutant should be **constitutive**.

We could make the **Gal4⁻** / **Gal80⁻** double mutant strain using molecular biological approaches...but an easier way is to let yeast meiosis do the job for you. If we mate the **Gal4⁻** / **Gal80⁺** haploid strain with the **Gal4+** / **Gal80⁻** haploid strain (we know these two genes are unlinked) we should obtain double mutants among the tetratype and non-parental ditype tetrads that result from this cross.

Using double mutants to order Gal4 and Gal80

MAT**a** Gal4⁻ Gal80⁺ x MATα Gal4⁺ Gal80⁻

Type 1	Type 2	Type 3
uninducible	uninducible	uninducible
uninducible	uninducible	uninducible
constitutive	constitutive	regulated (wt)
constitutive	regulated (wt)	regulated (wt)

Parental Ditype	Tetratype	Non Parental Ditype
Gal4- Gal80+	Gal4- Gal80+	**Gal4- Gal80-**
Gal4- Gal80+	**Gal4- Gal80-**	**Gal4- Gal80-**
Gal4+ Gal80-	Gal4+ Gal80-	Gal4+ Gal80+
Gal4+ Gal80-	Gal4+ Gal80+	Gal4+ Gal80+

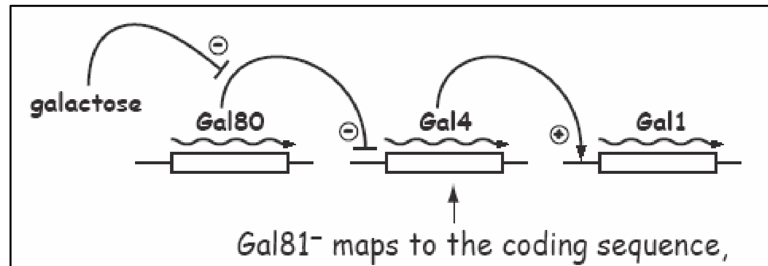
These results clearly support Model 1, i.e., because the double mutant is uninducible rather than constitutive, **Gal4** liklely behaves as a positive activator of **Gal1** expression, and in the absence of galactose **Gal80** somehow prevents **Gal4** from activation Gal1 expression. When galactose is present Gal80 can no longer prevent **Gal4** from activating **Gal1** expression.

Now lets consider a new class of mutant that turned out to be quite informative, **Gal81⁻**. **Gal81⁻** mutants, like **Gal80⁻** mutants are consititutive for **Gal1** expression, but unlike **Gal80⁻**, **Gal81⁻** is dominant. (**Gal81⁻**/ **Gal80⁻** diploids are constitutive).

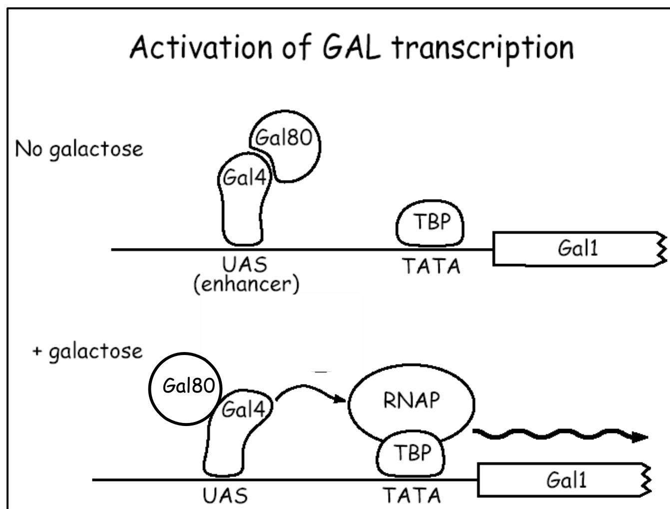
An obvious question is whether **Gal81⁻** mutants are still constitutive in a **Gal4⁻** background, since it was already established that **Gal4** positively regulates **Gal1** (and the other Gal genes).

Mata **Gal81⁻ Gal4⁺** **X** Matα **Gal4⁻ Gal81⁺**

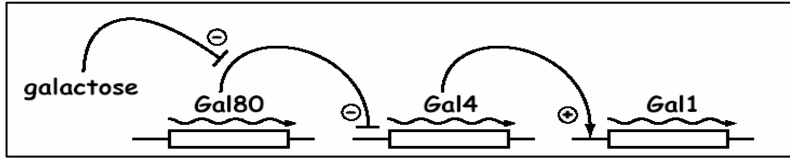
The surprising finding was that from this cross all the tetrads were of the **parental ditype**; in other words there were no tetratypes or nonparental ditypes, indicating that **Gal81⁻** and **Gal4⁻** are very tightly linked. Indeed, it turns out that the **Gal81⁻** mutation maps to the **Gal4** gene, in the coding region. The **Gal81⁻** mutation was redesignated as **$Gal4^{81}$**. Essentially **$Gal4^{81}$** behaves as a super-activator that is impervious to the negative effects of **Gal80; $Gal4^{81}$** thus activates independently of galactose and **Gal80**.



So, how do all these genetic facts fit into a molecular model? Upstream of the **Gal1** gene (and other **Gal** genes) two cis-acting elements are needed for transcriptional activation. First the TATA-binding protein (**TBP**) binds to the TATA-concensus site, and provides a landing pad for a very large RNA ploymerase complex (**RNAP**). However, just binding to TBP does not enable transcription, the complex must be activated by a transcriptional activator, in this case the **Gal4** protein. The **Gal4** protein sits at another cis-acting element in the **Gal1** promoter region, the upstream activator sequence (**UAS**) that tethers **Gal4** to the promoter. In the absence of galactose, **Gal80** physically prevents **Gal4** from recruiting and activating **RNAP**. In the presence of galactose the **Gal80** protein changes conformation and binds to a different region of **Gal4**, unveiling the ability of **Gal4** to recruit and activate **RNAP.**



(Note that the mutation in the **Gal4[81]** allele interferes with **Gal80** binding allowing **Gal4** to recruit and activate **RNAP** all the time, even in the absence of galactose.)



One final comment about the model for induction of the **Gal** genes by galactose. For many years it was assumed that galactose (or a derivative of galactose) actually binds directly to the **Gal80** protein, thus preventing it from inhibiting the **Gal4** protein from activating **Gal1** transcription. However, it now seems that one extra protein involved in this chain of events. The **Gal3** protein turns out to be directly bound by galactose (or a derivative); this allows **Gal3** to move from the cytoplasm into the nucleus, the galactose/**Gal3** moiety binds to **Gal80** to facilitate moving **Gal80** to a different site on the **Gal4** protein, thus allowing **Gal4** to activate transcription. Thus while the model as written in this figure does not actually include **Gal3**, the models are still formally correct.

In the next lecture we will be looking at promoter elements in eukaryotic genes in more detail.