Chapter 18
Regulation of Gene Expression

Quiz Questions:
1. What is a ‘housekeeping gene?’ Give an example of a housekeeping gene.
2. What is a Transcription Factor, and what does it do?
3. What is an Enhancer, and what does it do?
4. What is a ‘maternal effect,’ and what are maternal factors? (4 points)
5. What proteins make up nucleosomes, and what is their ionic charge relative to DNA?
6. What are HATs and HDACs, and how are they involved in gene transcription?
7. What are HMTases, and how are they involved in gene transcription?
8. What is the ‘Histone Code,’ and how is it involved in gene regulation? (4 points)
9. How does DNA methylation affect gene transcription?
10. Does histone acetylation increase or decrease transcription?
11. The histone proteins are rich in lysines. Why is this important?
12. Does histone deacetylation increase or decrease transcription?
13. What are ‘chromatin remodeling factors,’ and what is their function in gene transcription?
14. What are DNMTs, and what affect do they have on transcription?
15. What is the difference between an OPERON and a REGULON?
16. Which enzymes ensure that the decision to silence a gene is perpetuated (carried over) through cell divisions? (2 points)

NOTES

The term ‘gene regulation’ means the process of turning genes on or off in the right place at the right time. Recall that every human cell (for example) has two copies of every gene in the human genome. However, a heart cell only uses the genes that a heart cell needs, a kidney cell only uses the genes a kidney cell needs etc., and all of the genes that are not needed are turned off. Thus, turning on only the genes that a heart cell needs inside heart cells, while turning all other genes off is an example of gene regulation. Timing is also important. During development, many genes that determine the body plan of an organism are turned on, but are then turned off once the organisms is fully developed.

There are some genes, however, that are needed in every cell type. These are called ‘housekeeping genes.’ Examples of housekeeping genes are the Histone protein genes (because every cell needs to wrap its DNA around nucleosomes), the Actin genes (because every cell needs microfilaments to move things around in the cytoplasm), and the ribosomal RNA genes (because every cell needs ribosomes).
Gene Regulation is mediated by Promoters, Transcription Factors, post-translational histone modifications, RNA interference, and DNA Methylation.

**ENHANCERS and TRANSCRIPTION FACTORS:** Which genes are turned on or off in a specific cell is determined largely by what are called ‘tissue specific promoters.’ Recall from the previous lecture that every gene has an upstream sequence called a Promoter. Most promoters have a similar sequence (including a TATA Box). The job of the promoter is to attract an RNA Polymerase complex, consisting of RNA Pol II, and some other proteins that are needed to transcribe a gene. The RNA Polymerase complex physically interacts with the promoter by binding to it (non-covalently).

Other DNA elements called ‘Enhancers’ are located near to promoters. Other, accessory proteins called ‘Transcription Factors’ (TFs) bind to these enhancers, and increase the transcription initiation rate by stabilizing the RNA Polymerase complex. (The RNA Polymerase complex may bind to a promoter, but in the absence of TFs to stabilize the complex, it usually falls apart again.) The TFs bind directly to the Enhancer sequences in the DNA. Because the enhancers are physically part of the DNA, they are present in every cell. However, the transcription factors that bind to any given enhancer may not be produced in every cell. For example, a gene that is expressed only in brain cells will be present in non-brain cells, but the transcription factors that bind to the enhancers near the gene’s promoter, and greatly increase its transcription rate are only produced in brain cells. Thus, tissue-specific expression of genes is determined by tissue-specific expression of transcription factors.

You might very well ask “what determines whether a specific transcription factor will be turned on in a specific cell or not,” and this would be a good question! The answer is “other TFs bind to the promoters of those TFs; and THOSE TFs are regulated, in turn, by still MORE TFs!” In fact, a whole ‘cascade’ of transcription factors are set off in sequence, during development, and ultimately, the cascade is set off by the transcription factors that are put in the unfertilized egg by the mother. (So, which came first, the chicken or the egg? The answer is THE CHICKEN! The chicken was needed to put specific transcription factors into the egg before fertilization.)

**Maternal Factors and Maternal Effects:** The entire developmental ‘cascade’ is set off by transcription factors, and mRNAs that encode transcription factors that were deposited in the unfertilized egg by the mother. These TFs and mRNAs are called ‘Maternal Factors.’ The term ‘Maternal Effect’ is a term used by Geneticists to refer to the fact that what genes and gene alleles the mother is carrying are significant to the development of an organism. For example, there are many genes that encode transcription factors that, if deleted or mutated are lethal (i.e., their absence means the organism can’t develop), but only if the MOTHER carries the mutations. Thus, if you did a fruit fly genetic cross where the father carried a mutation to one of these factors, but the mother didn’t, the mother would still produce viable offspring. However, if you did the ‘reciprocal cross,’ where the mother carried the mutation, but the
father didn’t, no offspring would be produced, because the mother couldn’t deposit the needed factor into her eggs. This is an example of a ‘maternal effect.’

**The Histone Code:** Recall that DNA in the nucleus is wrapped around spool-shaped structures called nucleosomes. These nucleosomes are composed of four proteins called the Histone Proteins (designated H2A, H2B, H3 and H4). DNA is negatively charged in solution, and the histone proteins that make up nucleosomes are positively charged, allowing the DNA to wrap tightly around the nucleosomes. This presents a problem during transcription, however. In order for a gene to be transcribed, the nucleosomes have to be moved out of the way. This job is done by special proteins called ‘chromatin remodeling factors.’ This job is much harder to do if the DNA is wrapped tightly around the nucleosomes. One way to loosen the interaction between nucleosomes and DNA is to neutralize the positive charge on the histone proteins. This can be done by covalently attaching acetyl groups to the basic amino acids (usually lysines) that make up nucleosomes. The proteins that transfer acetyl groups to histone proteins, in order to neutralize their positive charge, and make it easier for other proteins to move the nucleosomes are called Histone Acetyl Transferases (abbreviated HATs). Conversely, there are other proteins whose job it is to remove the acetyl groups. They are called Histone Deacetylases (abbreviated HDACs). Thus, HATs increase the transcription rate of a gene by making it easier to move the nucleosomes out of the way, and HDACs decrease the transcription rate of a gene by making it harder to move the nucleosomes out of the way. HATs ‘activate’ genes, and HDACs ‘silence’ genes.

There are some other modifications that can be made to the histone proteins, such as transferring or removing Methyl groups and Phosphate groups, which increase or decrease the rate of transcription. Sometimes the Acetyl, Methyl or Phosphate groups are attached to the same amino acids in the histone proteins, and sometimes to different amino acids. (ie- a Methyl group can be transferred to a specific lysine in Histone H3 so that that particular lysine CANNOT be acetylated, thus silencing the gene.) Different combinations can also increase or decrease transcription. (ie-Three Acetyl groups attached to Histone H3 will activate it; but three Acetyl groups and a Methyl group will cause it to be silent. The silencing may be overridden by phosphorylating one of the serines in H3 etc.) The various combinations of post-translations to the histone proteins that can either activate or silence transcription are collectively called the ‘Histone Code.’ The Histone Code is very complicated, and we won’t cover all of it here. Just remember that HATs activate genes, while HDACs silence them. Also remember that enzymes that transfer methyl groups to the histones (called Histone Methyltransferases, abbreviated HMTases) tend to silence genes by preventing histones from being acetylated.

**DNA Methylation and Gene Silencing:** Enzymes called DNA Methyltransferases (abbreviated DNMTs) put methyl groups onto DNA (covalent bonds), making it harder for the RNA polymerase complex to transcribe the DNA. Long stretches of GGGG or CCCC base pairs tend to be methylated by DNMTs. If such stretches are located near gene promoters, and are methylated by specific DNMTs, the gene will not be transcribed. Recall that DNA is double stranded. Silencing works better if both strands are methylated. Recall, also, that DNA
replication is ‘semi-conservative,’ meaning that one strand will be methylated, but not the other. A group of DNMTs called ‘Hemimethylases’ will methylate the other strand if only one strand is methylated (ie-just after cell division). Thus, once a decision has been made to silence a gene, the silenced state will be perpetuated even if the cell divides, because hemimethylases will fully methylate a strand of DNA that initially had only one strand that was methylated. Furthermore, because these signals to silence genes through methylation can be perpetuated (carried over after cell division), it is possible that a gene that was silenced in a parent can pass the gene on in the silenced state to their offspring. This is called an ‘epigenetic inheritance.’

**Classical Genetics vs. Epigenetics:** You’ve already learned that there are dominant and recessive gene alleles, and that recessive gene alleles are often non-functioning mutant versions of a gene. If a person inherits a non-functional allele of a gene, they will show the recessive phenotype for that particular gene. However, imagine what would happen if a person inherited an allele of a gene that is functional, but happens to be permanently TURNED OFF. The result would be much the same as inheriting a non-functional mutant allele, wouldn’t it? This is the main difference between classical genetics, and epigenetics. In classical genetics, an organism shows a recessive phenotype for a certain genetic trait when they inherit a non-functional allele of a gene. In epigenetics, an organism may show a recessive phenotype for a certain genetic trait despite having functional alleles. The recessive phenotype is caused by the functional gene being inherited in a silenced state (usually the result of methylation).

**RNA Interference and Gene Silencing:** mRNA is single stranded. If DOUBLE STRANDED RNA is found in the cell, it is perceived by the cell to be either an error, or possibly the genome of an invading virus (Some viruses have double stranded RNA genomes) and is destroyed. Interestingly, many genes have TWO promoters, one located upstream of the gene, and the other downstream of it. The promoter upstream of the gene will cause a ‘sense’ (+ Strand) mRNA to be transcribed. If the promoter downstream of the gene is activated, however, an ‘antisense’ (-Strand) RNA will be produced. If the sense and antisense strands bind together they will form double stranded RNA and be destroyed. This can actually be used by the cell as an additional way of ‘fine tuning’ the level of protein production. By turning on the upstream promoter strongly, you will produce a lot of the protein that that gene encodes. If you turn on the downstream promoter by only half as much, half of the sense (forward) transcripts will combine with the antisense (backwards) transcripts and be destroyed, reducing the amount of protein translated from that gene by half.

Thus, the following things INCREASE gene transcription:
1. ‘Strength’ of the Promoter (how well it conforms to a ‘consensus’ sequence.)
2. Transcription Factors binding to Enhancers.
3. Histone Acetylation.
4. Chromatin Remodeling Complexes.

And the following things DECREASE gene transcription:
1. Histone Demethylation
2. Histone Methylation
Operons and Regulons: When two or more genes are turned on or off at the same time, they are said to be ‘Co-Regulated’ (‘co-regulation’). Sometimes several genes are physically linked together, and are being driven by a single promoter located upstream of them. When several genes are physically connected in series, and are all located downstream of the same promoter, they are collectively called an Operon. When several genes are located in different places, but are all being turned on or off by identical copies of the same promoter they are called a Regulon. Both Operons and Regulons are examples of co-regulation.

The Lac Operon: The most conspicuous example of an operon (ie-the one which all first year biology students are forced to learn about!) is called the Lac Operon. The Lac Operon is a series of three genes that are involved in Lactose metabolism in the bacteria *E. coli*. Bacteria can metabolize (eat) the sugar lactose with the help of three genes that are all located downstream of the same promoter (see Figure 18.4). When the operon is turned on, all three genes are transcribed as a single mRNA that is translated into three proteins. Another protein called the Lac Repressor is produced by another gene (the *Lac I* gene) located elsewhere on the chromosome. When the Lac Repressor is present it will bind to a DNA sequence called the Lac Operator, located just upstream of the Lac Promoter, and stop the genes from being transcribed by physically BLOCKING the RNA Polymerase complex from proceeding (like a log blocking a railroad track).

The cell is not interested in producing the three enzymes that allow it to eat lactose when lactose is not present. (It would be a waste of energy to produce enzymes that are not needed.) Thus, when no lactose is present, the Lac Repressor protein is bound to the Lac Operator sequence, preventing the genes from being transcribed. If lactose enters the cell, however, some of it will bind to the Lac Repressor, causing it to change shape (called an ‘allosteric’ change), and be RELEASED from the Lac Operator. Thus, binding of lactose to the Lac Repressor causes the repressor UN-BLOCK transcription of the genes, allowing the three enzymes to be produced. This is a classic example of ‘negative gene regulation’ because the genes are turned off by default, and something has to happen in order to turn them on. There are also examples of positive gene regulation, where genes are being transcribed, but only at a low level, and various things can happen that can increase the rate of transcription.

**PRACTICE QUESTIONS:**

**ESSAY QUESTIONS:**
1. Explain what RNAi is, and how it is involved in gene silencing. (20 points)
2. List the elements that control the Lac Operon, and how the Lac Operon works. (20 points)
3. What is meant by ‘co-regulation,’ and what is the difference between an Operon and a Regulon? (10 points)
EXTENDED MATCHING QUESTION:

A. Allosteric
B. Chromatin Remodeling Factors
C. DNMT
D. Enhancer
E. HAT
F. HDAC
G. Hemimethylases
H. HMTase
I. Lac Operator
J. Lac Repressor
K. Operon
L. Operator
M. Promoter
N. Regulon
O. Repressor
P. RNAi
Q. Transcription Factors

1. Abbreviation for an enzyme that methylates DNA.
2. A protein that binds to an area near the Lac Promoter, blocking the RNA Polymerase complex.
3. A protein that binds to an enhancer, and stabilizes the RNA Polymerase complex, thus increasing the transcription rate of a gene.
4. Two or more genes that are located on different chromosomes, but which are attached to similar promoters that are co-regulated.
5. Phenomenon of lowering the number of proteins produced from a gene by destroying double stranded mRNA.
6. A protein that binds methyl groups to nucleosomes, thus blocking histone acetylation.
7. The DNA sequence to which the Lac Repressor binds.
8. A protein that removes acetyl groups from nucleosomes.
9. Two or more genes that are physically linked (in series), and are driven by the same promoter.
10. A DNA sequence that a transcription factor binds to.
11. Term for a conformational change to a protein.
12. Proteins that move nucleosomes out of the way so that a gene can be transcribed.
13. A protein that puts a methyl group onto a non-methylated DNA strand that is paired with a methylated strand.
14. A protein that puts acetyl groups on histone proteins.
15. A DNA sequence located upstream of a gene, to which the RNA polymerase complex binds.
16. Which of the above INCREASE gene transcription?
17. Proteins that move nucleosomes out of the way, facilitating transcription.
18. Which of the above DECREASE gene transcription?

Lac Operon Question: For each set of conditions, say whether the Lac Operon enzymes will be produced or not A) in the PRESENCE of lactose, or B) in the ABSENCE of lactose.

1. Everything is normal. No mutations.
2. Lac Operator sequence is mutated so that Lac Repressor can’t bind to it.
3. Lac I gene is mutated so that no Lac Repressor is produced (ie-a deletion mutation).
4. Lac I gene is mutated so that the Lac Repressor protein produced is not capable of binding to the Lac Operator sequence.
5. Lac I gene is mutated so that the Lac Repressor protein produced is not capable of binding lactose.