Transcription factors

khanacademy.org/science/biology/gene-regulation/gene-regulation-in-eukaryotes/a/eukaryotic-transcription-factors

Key points:

- **Transcription factors** are proteins that help turn specific genes "on" or "off" by binding to nearby DNA.
- Transcription factors that are **activators** boost a gene's transcription. **Repressors** decrease transcription.
- Groups of transcription factor binding sites called **enhancers** and **silencers** can turn a gene on/off in specific parts of the body.
- Transcription factors allow cells to perform logic operations and combine different sources of information to "decide" whether to express a gene.

Introduction

Do you have any transcription factors in your body? I sure hope the answer is yes, because otherwise, you're going to have a hard time keeping your cells running!

Transcription factors are proteins that regulate the transcription of genes—that is, their copying into RNA, on the way to making a protein.

The human body contains many transcription factors. So does the body of a bird, tree, or fungus! Transcription factors help ensure that the right genes are expressed in the right cells of the body, at the right time.

Transcription: The key control point

Transcription is the process where a gene's DNA sequence is copied (transcribed) into an RNA molecule. Transcription is a key step in using information from a gene to make a protein. If you're not familiar with those ideas yet, you might consider watching the central dogma video for a solid intro from Sal.
Gene expression is when a gene in DNA is "turned on," that is, used to make the protein it specifies. Not all the genes in your body are turned on at the same time, or in the same cells or parts of the body.

For many genes, transcription is the key on/off control point:

- If a gene is not transcribed in a cell, it can't be used to make a protein in that cell.

- If a gene does get transcribed, it is likely going to be used to make a protein (expressed). In general, the more a gene is transcribed, the more protein that will be made.
  [Is that always the case?]
  Not always. Sometimes, later stages of regulation can block even large quantities of mRNA from being used to make protein.

  For example, imagine that a gene is transcribed a lot, but the mRNA is "chopped up" (degraded) as soon as it leaves the nucleus. This would lead to very little protein getting made.

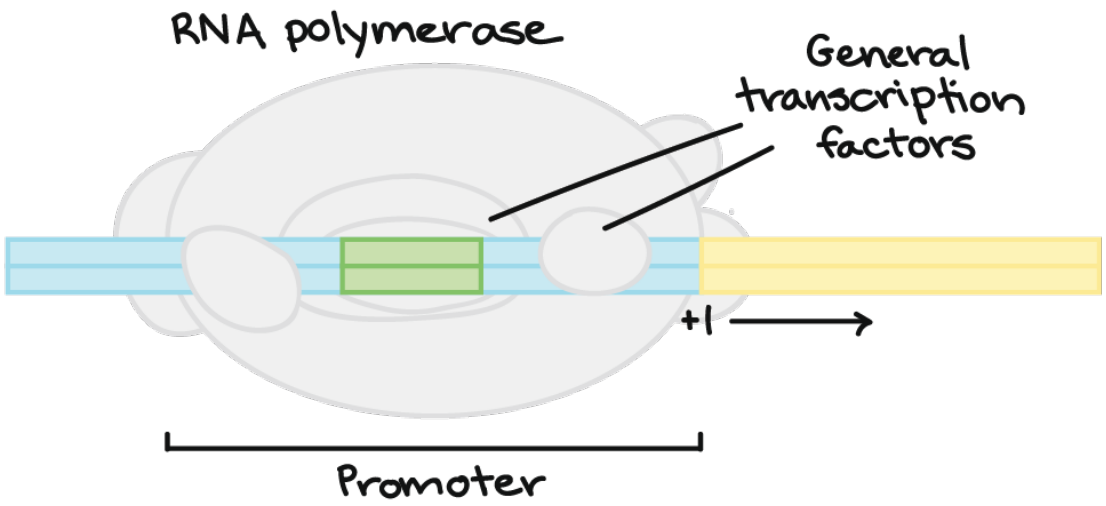
Various factors control how much a gene is transcribed. For instance, how tightly the DNA of the gene is wound around its supporting proteins to form **chromatin** can affect a gene's availability for transcription.

Proteins called **transcription factors**, however, play a particularly central role in regulating transcription. These important proteins help determine which genes are active in each cell of your body.

Transcription factors

What has to happen for a gene to be transcribed? The enzyme**RNA polymerase**, which makes a new RNA molecule from a DNA template, must attach to the DNA of the gene. It attaches at a spot called the **promoter**.

In bacteria, RNA polymerase attaches right to the DNA of the promoter. You can see how this process works, and how it can be regulated by transcription factors, in the *lac* operon and *trp* operon videos.
In humans and other eukaryotes, there is an extra step. RNA polymerase can attach to the promoter only with the help of proteins called **basal** (**general**) **transcription factors**. They are part of the cell's core transcription toolkit, needed for the transcription of any gene.



RNA polymerase binds to a promoter with help from a set of proteins called general

transcription factors.

However, many transcription factors (including some of the coolest ones!) are not the general kind. Instead, there is a large class of transcription factors that control the expression of specific, individual genes. For instance, a transcription factor might activate only a set of genes needed in certain neurons.

How do transcription factors work?

A typical transcription factor binds to DNA at a certain target sequence. Once it's bound, the transcription factor makes it either harder or easier for RNA polymerase to bind to the promoter of the gene.

Activators

Some transcription factors **activate** transcription. For instance, they may help the general transcription factors and/or RNA polymerase bind to the promoter, as shown in the diagram below.

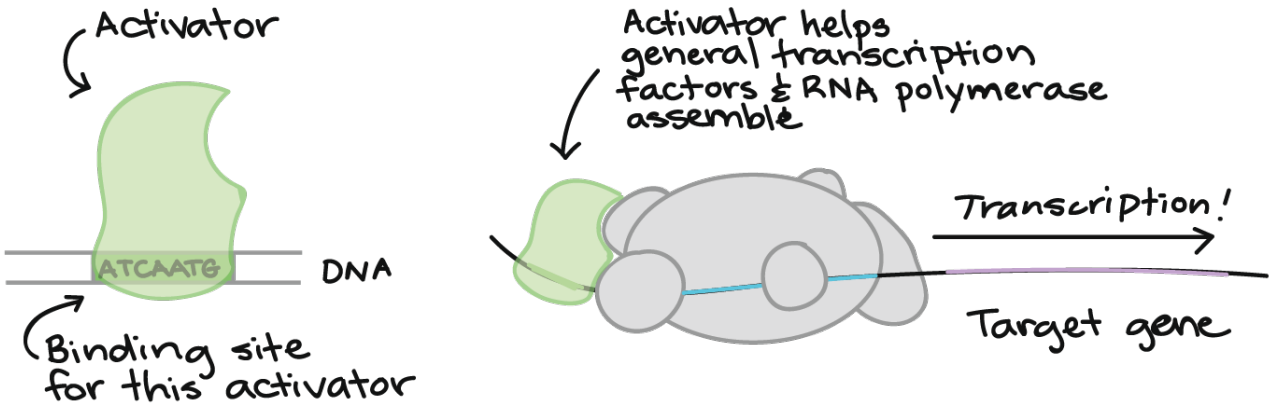


Diagram of an activator attached to a specific DNA sequence that is its binding site. The other end of the transcriptional activator (the one not bound to the DNA) interacts with general transcription factors, helping the general transcription factors and polymerase assemble tat the nearby promoter.

Repressors

Other transcription factors **repress** transcription. This repression can work in a variety of ways. As one example, a repressor may get in the way of the basal transcription factors or RNA polymerase, making it so they can't bind to the promoter or begin transcription.

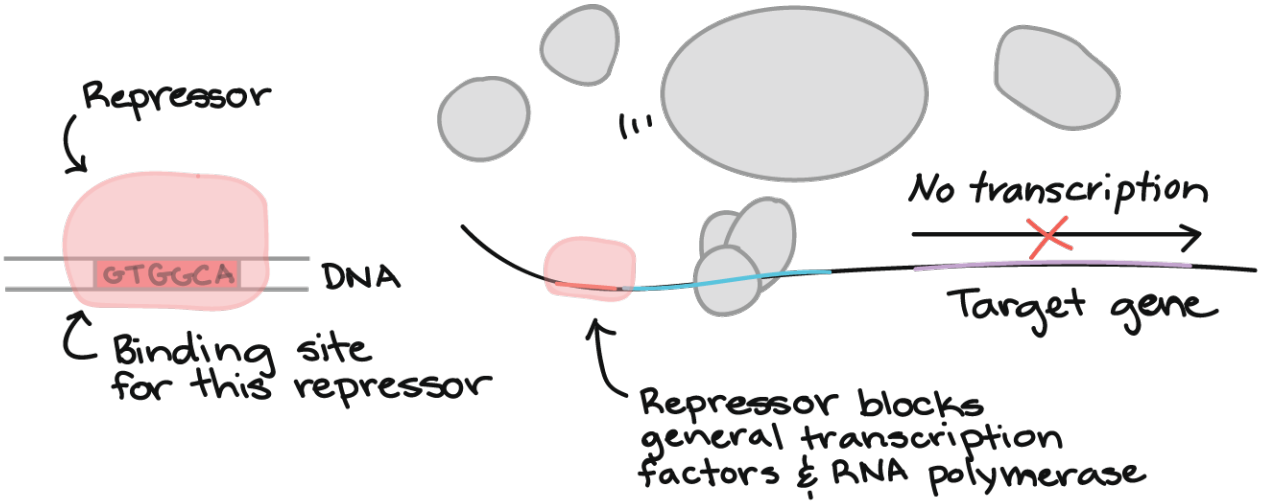
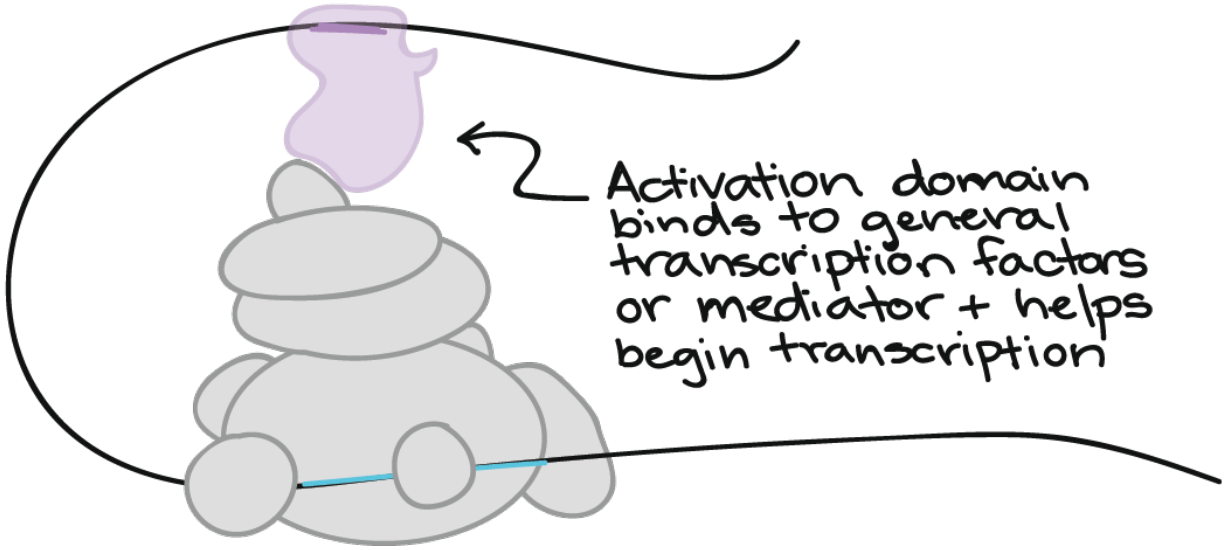


Diagram of a repressor attached to a specific DNA sequence that is its binding site. When bound to this site, the repressor blocks formation of the transcription initiation complex at the promoter of a nearby gene.

Binding sites

The binding sites for transcription factors are often close to a gene's promoter. However, they can also be found in other parts of the DNA, sometimes very far away from the promoter, and still affect transcription of the gene.



The parts of an activator protein: the DNA binding domain (which attaches to the recognition site in the DNA) and the activation domain, which is the "business end" of the activator that actually promotes transcription, e.g., by facilitating formation of the transcription initiation complex.

The flexibility of DNA is what allows transcription factors at distant binding sites to do their job. The DNA loops like cooked spaghetti to bring far-off binding sites and transcription factors close to general transcription factors or "mediator" proteins.

In the cartoon above, an activating transcription factor bound at a far-away site helps RNA polymerase bind to the promoter and start transcribing.

[Where do transcription factors come from?]
Transcription factors are proteins, so they are encoded by genes and made via gene expression (transcription and translation). In this way, they are no different from any other protein in the cell.

If you're wondering how different cell types "know" which transcription factors to make, however, that's a more complicated question. (Also a great question!)

In fact, it turns out to be the gateway to a very long rabbit hole, what might be considered the central rabbit hole of developmental biology! The set of transcription factors produced in any cell type is the result of a long series of molecular events that can be traced all the way back to the origins of the organism as a single cell.^11start superscript, 1, end superscript Developmental biologists seek to trace these long chains of causality, identifying the signals, cues, and interactions that lead to the expression of specific sets of transcription factors (and other key regulators) in particular cell types.

How is this different from *E. coli*?

So far, human and other eukaryotic transcription factors don't seem all that different from the transcription factors we've seen in bacteria. They bind DNA and make it easier or harder for RNA polymerase to do its job—just like the *lac* repressor protein of *E. coli*
In general, this is a pretty good takeaway. Proteins that control transcription tend act in similar ways, whether they're in your own cells or in the bacteria that live your nose. The main differences are mechanical—how far away regulatory sites are, whether basal transcription factors are needed, etc.

However, there are also some meaningful differences in how transcription factors are used in humans. Humans and other eukaryotes are complex: we're made up of trillions of cells organized into unique tissues and body structures. Each cell in your body must run its own "program" of gene expression.

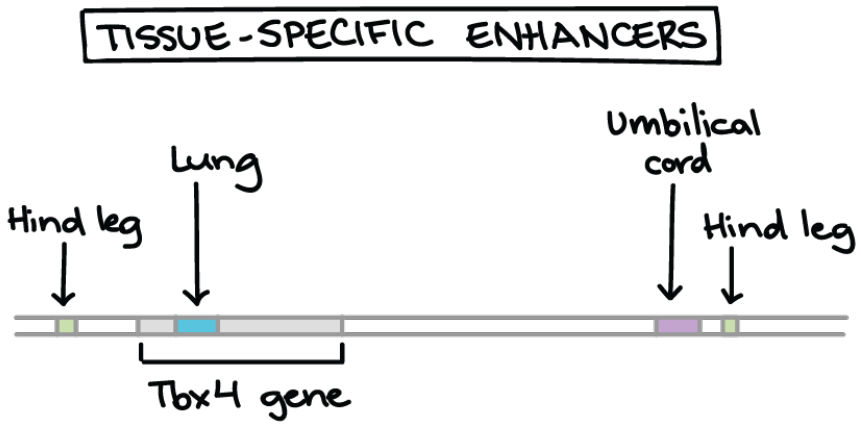
Turning genes on in specific body parts

Some genes need to be expressed in more than one body part or type of cell. For instance, suppose a gene needed to be turned on in your spine, skull, and fingertips, but not in the rest of your body. How can transcription factors make this pattern happen?

A gene with this type of pattern may have several **enhancers** (far-away clusters of binding sites for activators) or **silencers** (the same thing, but for repressors). Each enhancer or silencer may activate or repress the gene in a certain cell type or body part, binding transcription factors that are made in that part of the body.^{1,2}1,2start superscript, 1, comma, 2, end superscript

Example: Modular mouse

As an example, let's consider a gene found in mice, called *Tbx4*. This gene is important for the development of many different parts of the mouse body, including the blood vessels and hind legs^33start superscript, 3, end superscript.
During development, several well-defined enhancers drive *Tbx4* expression in different parts of the mouse embryo. The diagram below shows some of the *Tbx4* enhancers, each labeled with the body part where it produces expression.



Not drawn fully to scale. Image based on Figure 5 of Menke et al.^33start superscript, 3, end superscript

Evolution of development

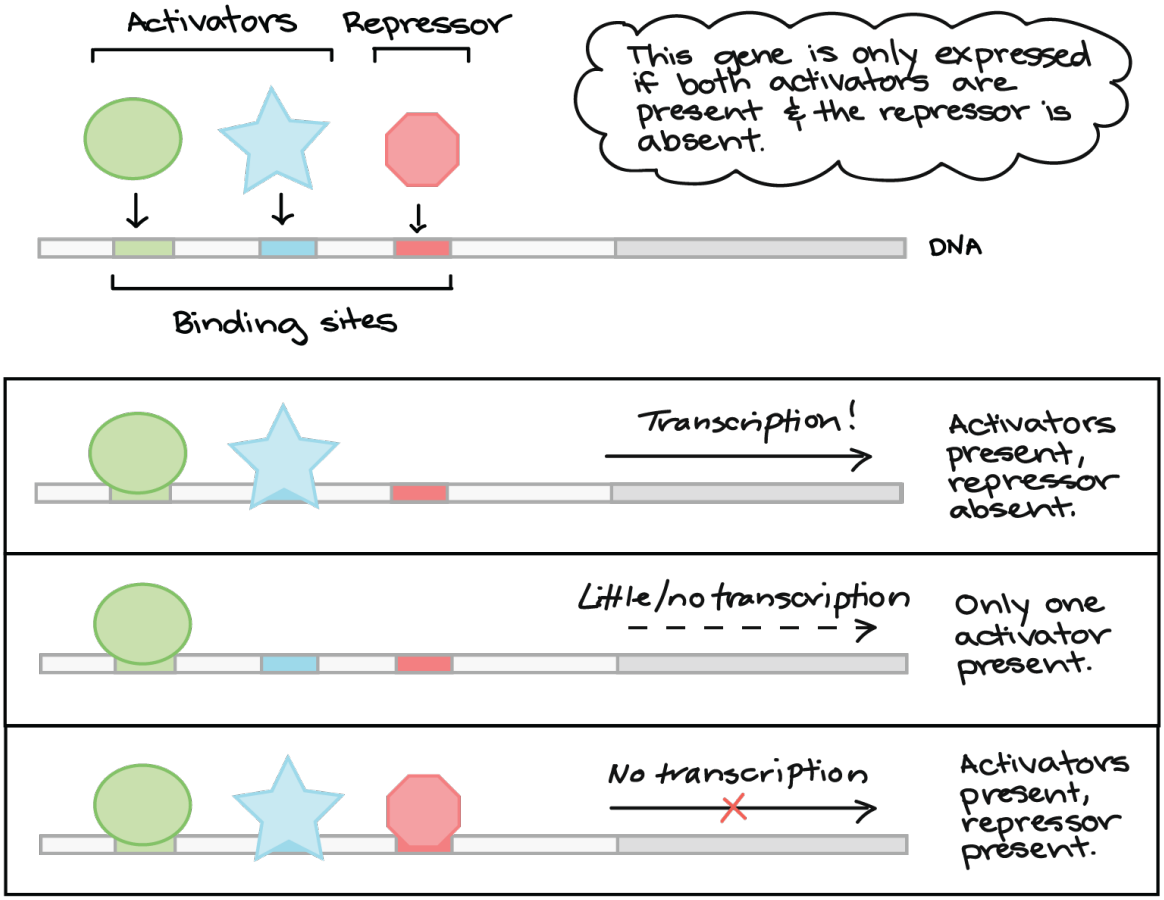
Enhancers like those of the *Tbx4* gene are called **tissue-specific** enhancers: they control a gene's expression in a certain part of the body. Mutations of tissue-specific enhancers and silencers may play a key role in the evolution of body form.^44start superscript, 4, end superscript
How could that work? Suppose that a mutation, or change in DNA, happened in the coding sequence of the *Tbx4* gene. The mutation would inactivate the gene everywhere in the body, and mouse without a normal copy would likely die. However, a mutation in an enhancer might just change the expression pattern a bit, leading to a new feature (e.g., a shorter leg) without killing the mouse.

Transcription factors and cellular "logic"

Can cells do logic? Not in the same way as your amazing brain. However, cells can detect information and combine it to determine the correct response—in much the same way that your calculator detects pushed buttons and outputs an answer.

We can see an example of this "molecular logic" when we consider how transcription factors regulate genes. Many genes are controlled by several different transcription factors, with a specific combination needed to turn the gene on; this is particularly true in eukaryotes and is

sometimes called **combinatorial regulation**.^{5,6}5,6start superscript, 5, comma, 6, end superscript For instance, a gene may be expressed only if activators A and B are present, and if repressor C is absent.



In this diagram, a gene has three binding sites. One is for a circle-shaped activator, another is for a star-shaped activator, and the third is for a repressor shaped like a stop sign (octagonal). This gene is only expressed if both activators are present and the repressor is absent.

Scenario 1: Both activators are present, the repressor is absent. In this case, transcription occurs.

Scenario 2: Only one activator is present. Little or no transcription occurs.

Scenario 3: Both activators are present, but the repressor is also present. No transcription occurs.

[Hey, that sounds like an "if statement" in programming!]
Yes, it does! If you enjoy computer programming, you may notice that this looks a lot like a conditional statement (a logic gate that determines if and when a block of code gets executed). In code, this might look something like:

```
if (activator A == TRUE and activator B == TRUE)

{
    if (repressor C == FALSE)
    {
       express target gene;
    }
}
```

Cool, huh? In fact, many regulatory systems in biology are basically logical circuits built out of biomolecules.

The use of multiple transcription factors to regulate a gene means that different sources of information can be integrated into a single outcome. For instance, imagine that:

- Activator A is present only in skin cells
- Activator B is active only in cells receiving "divide now!" signals (growth factors) from neighbors
- Repressor C is produced when a cell's DNA is damaged

In this case, the gene would be "turned on" only in skin cells that are receiving division signals and have undamaged, healthy DNA. This pattern of regulation might make sense for a gene involved in cell division in skin cells. In fact, the loss of proteins similar to repressor C can lead to cancer.
Real-life combinatorial regulation can be a bit more complicated than this. For instance, many different transcription factors may be involved, or it may matter exactly how many molecules of a given transcription factor are bound to the DNA.

[References]

Works cited:

1. Gilbert, S. F. (2000). Anatomy of the gene: Promoters and enhancers. In *Developmental biology* (6th ed.). Sunderland, MA: Sinauer Associates. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK10023/#_A751_.
2. Gilbert, S. F. (2000). Silencers. In *Developmental biology* (6th ed.). Sunderland, MA: Sinauer Associates. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK10023/#_A777_.
3. Menke, D. B., Guenther, C., and Kingsley, D. M. (2008). Dual hindlimb control elements in the Tbx4 gene and region-specific control of bone size in vertebrate limbs. *Development*, *135*, 2543-2553. http://dx.doi.org/10.1242/dev.017384.
4. Wray, Gregory A. (2007). The evolutionary significance of *cis*-regulatory mutations. *Nature Reviews Genetics*, *8*, 206-216. http://dx.doi.org/10.1038/nrg2063.

5. Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., and Jackson, R. B. (2011). Combinatorial control of gene activation. In *Campbell Biology* (10th ed., pp. 37). San Francisco, CA: Pearson.

6. Reményi, Attila, Hans R. Schöler, and Matthias Wilmanns. (2004). Combinatorial control of gene expression. *Nature Structural & Molecular Biology*, *11*(9), 812. http://dx.doi.org/10.1038/nsmb820. Retrieved from http://www.nature.com/scitable/content/Combinatorial-control-of-gene-expression-16976.

Additional references:

Body plans. (n.d.). In *National center for science education*. Accessed June 8, 2016. Retrieved from http://ncse.com/book/export/html/2585.
Chapter 7 summary. (2013). In *The cell: A molecular approach [companion website]*. Retrieved from http://sites.sinauer.com/cooper7e/.
Cooper, G. M. (2000). Regulation of transcription in eukaryotes. In *The cell: A molecular approach*. Sunderland, MA: Sinauer Associates. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK9904/.
Ernsberg, E. (2004). Regulation of neuronal gene expression by the neuron-restrictive silencer. In R. W. Davies and and B. Morris (Eds.), *Molecular biology of the neuron* (2nd ed., pp. 34-38).

Anatomy of the gene: Promoters and enhancers. In *Developmental biology* (6th ed.). Sunderland, MA: Sinauer Associates. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK10023/#_A751_.
Gilbert, S. F. (2000). Differential gene transcription. In *Developmental biology* (6th ed.). Sunderland, MA: Sinauer Associates. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK10023/.
Gilbert, S. F. (2000). Silencers. In *Developmental biology* (6th ed.). Sunderland, MA: Sinauer Associates. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK10023/#_A777_.
Griffiths, A. J. F., Gelbart, W. M., Miller, J. H. and Lewontin, R. C. (1999). Transcription: Gene regulation in eukaryotes — an overview. In *Modern genetic analysis*. New York, NY: W. H. Freeman. Retrieved from www.ncbi.nlm.nih.gov/books/NBK21346/.

Griffiths, A. J. F., Miller, J. H., Suzuki, D. T., Lewontin, R. C., and Gelbart, W. M. (2000). Transcription: An overview of gene regulation in eukaryotes. In *An introduction to genetic analysis* (7th ed.). New York, NY: W. H. Freeman. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK21780/.
Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., and Darnell, J. (2000). Molecular mechanisms of eukaryotic transcriptional control. In *Molecular cell biology* (4th ed., section 10.7). New York, NY: W. H. Freeman. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK21677/.
Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., and Darnell, J. (2000). RNA polymerase II transcription-initiation complex. In *Molecular cell biology* (4th ed., section 10.6).

New York, NY: W. H. Freeman. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK21610/.
Nelson, D. L. and Cox, M. M. (2008). Choreography of transcriptional activation. In *Lehninger principles of biochemistry* (5th ed., p. 1140). New York, NY: W. H. Freeman and Company.

OpenStax College, Biology. (2016, March 23). Eukaryotic transcription gene regulation. In _OpenStax CNX. Retrieved from http://cnx.org/contents/GFy_h8cu@10.8:7Ry3oRse@5/Eukaryotic-Transcription-Gene-.
Peirano, R. I., and Wegner, M. (2000). The glial transcription factor Sox10 binds to DNA both as monomer and dimer with different functional consequences. *Nucleic Acids Res.*, *28*(16), 3047-3055. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC108444/.
Phillips, T. (2008). Regulation of transcription and gene expression in eukaryotes. *Nature Education*, *1*(1), 199. Retrieved from http://www.nature.com/scitable/topicpage/regulation-of-transcription-and-gene-expression-in-1086.
Purves, W. K., Sadava, D. E., Orians, G. H., and Heller, H.C. (2003). Transcriptional regulation of gene expression. In *Life: The science of biology* (7th ed., pp. 290-296). Sunderland, MA: Sinauer Associates.

QIAGEN. (2012). Search result: Transcription factors regulating gene SCN2A. In *EpiTect ChIP qPCR primers*. Retrieved from http://www.sabiosciences.com/chipqpcrsearch.php?species_id=0&nfactor=n&ninfo=n&ngene=n&B2=Search&src=genecard&factor=Over+200+TF&gene=SCN2A.
Raftopoulou, M. (2005). (1987) Nucleosomes regulate gene expression: The importance of wrapping. *Nature milestones*. http://dx.doi.org/10.1038/nrm1805.
Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., and Jackson, R. B. (2011). Regulation of transcription initiation. In *Campbell biology* (10th ed., pp. 367-372). San Francisco, CA: Pearson.

Reményi, Attila, Hans R. Schöler, and Matthias Wilmanns. (2004). Combinatorial control of gene expression. *Nature Structural & Molecular Biology*, *11*(9), 812-815. http://dx.doi.org/10.1038/nsmb820. Retrieved from http://www.nature.com/scitable/content/Combinatorial-control-of-gene-expression-16976.
Schoenherr, C. J., and Anderson, D. J. (1995). The neuron-restrictive silencer factor (NRSF): A coordinate repressor of multiple neuron-specific genes. *Science*, *267*(5202), 1360-1363. http://dx.doi.org/10.1126/science.7871435.
Transcriptional regulation. (2016, May 2). Retrieved May 17, 2016 from Wikipedia: https://en.wikipedia.org/wiki/Transcriptional_regulation.
Wray, Gregory A. (2007). The evolutionary significance of *cis*-regulatory mutations. *Nature Reviews Genetics*, *8*, 206-216. http://dx.doi.org/10.1038/nrg2063.