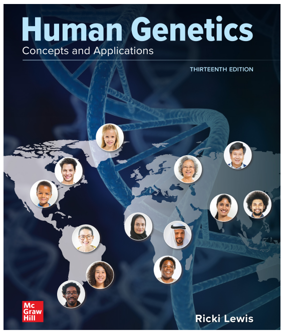
Because learning changes everything.®

Chapter 11

Gene Expression and Epigenetics

Thirteenth Edition

Ricki Lewis



Learning Outcomes

1. Define *epigenetic*.
2. Explain how development of the placenta, globin chain switching, development of organs, and the types of proteins cells make over time illustrate gene expression.
3. Explain how small molecules binding to histone proteins control gene expression by remodeling chromatin.
4. Explain how microRNAs and microproteins control transcription.
5. Explain how division of genes into exons and introns maximizes the number of encoded proteins.
6. Discuss how viral DNA, noncoding RNAs, and repeated DNA sequences account for large proportions of the human genome.

Gene Expression through Time and Tissue 1

A genome is like an orchestra.

- Just as not all of the musical instruments play with the same intensity at every moment, not all genes are expressed continually at the same levels.

Some genes are always expressed in all cells

- These are called “housekeeping” genes

Other genes are more specialized

- Become active as a cell differentiates or challenges arise

The collection of mRNAs in a cell constitutes its **transcriptome**.

Gene Expression through Time and Tissue 2

Changes in gene expression may occur over time and in different cell types

May occur at the molecular, tissue, or organ/gland level

Epigenetic changes

- Changes to the chemical groups that associate with DNA that are transmitted to daughter cells after cell division

Gene Expression in the Placenta

The physiology of the placenta at various times is being investigated using mRNA patterns

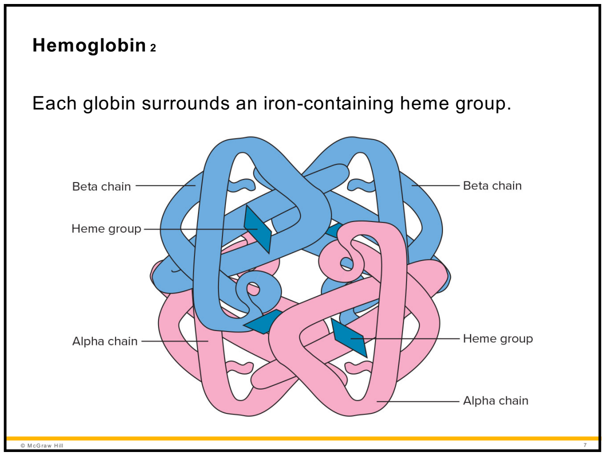
The genes expressed in human placentas fall into five groups.

- Only one represents the woman
  - The mRNAs indicate that the genes in that group dampen her immune response, which might otherwise reject the fetus.
- The other four groups of genes encode proteins that invade the uterine lining, build the blood supply, and release hormones necessary to maintain the pregnancy.

Hemoglobin 1

Adult hemoglobin has four globular polypeptide chains:

- Two alpha (α) chains = 141 amino acids
  - Encoded on chromosome 11
- Two beta (β) chains = 146 amino acids
  - Encoded on chromosome 16



Hemoglobin 2

Each globin surrounds an iron-containing heme group.

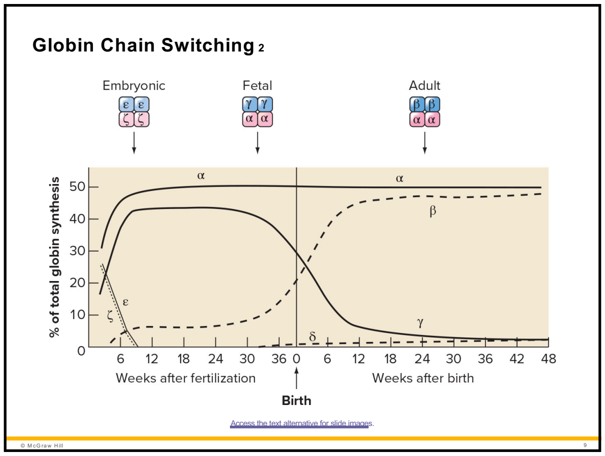
Globin Chain Switching 1

Subunits change in response to oxygen levels.

Subunit makeup varies over lifetime.

- Embryo—Two epsilon (ε) + two zeta (ζ)
- Fetus—Two gamma (γ) + two alpha (α)
- Adult—Two beta (β) + two alpha (α)

Adult type is about 99% of hemoglobins by four years of age.



Globin Chain Switching 2

Access the text alternative for slide images.

Building Tissues and Organs

A solid gland or organ is much more complex than blood.

- It has a distinctive three-dimensional form and is constructed from specialized types of cells and commonly more than one type of tissue.

In all tissues and organs, genes are turned on and off during development, as stem cells self-renew and yield more specialized daughter cells.

- The organization of the pancreas is simple enough to reveal how its cells specialize.

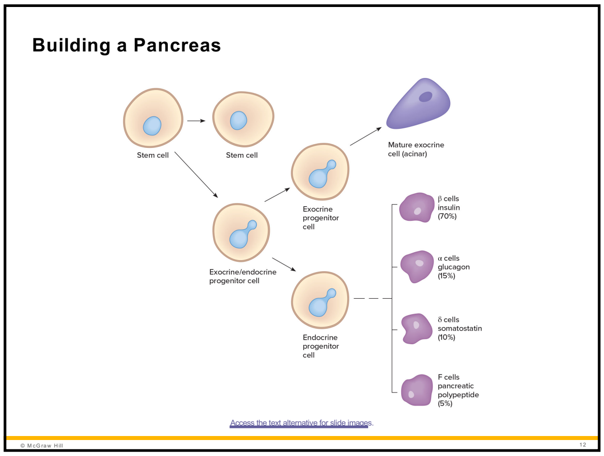
Pancreas

Dual gland

- Exocrine part releases digestive enzymes into ducts
- Endocrine part secretes polypeptide hormones directly into the bloodstream

Differential gene expression produces either endocrine or exocrine cells.

- If a transcription factor is activated, some progenitor cells follow the exocrine pathway.
- Other progenitor cells respond to different signals and yield daughter cells that follow the endocrine pathway.



Building a Pancreas

Access the text alternative for slide images.

Transcriptomics and Proteomics

Transcriptomics identifies all the mRNA molecules made in a specific cell under specific circumstances.

Proteomics identifies all the proteins a cell manufactures under specific conditions.

Transcriptomics and proteomics are useful in better understanding basic biology.

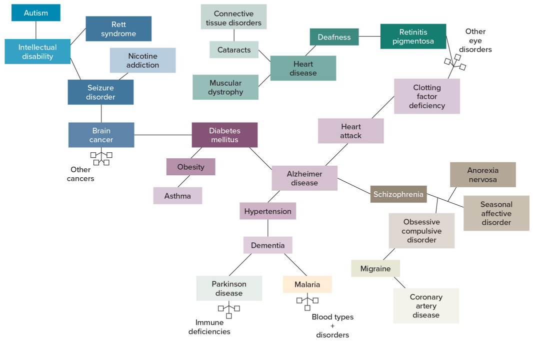
- Proteins can be charted based on the relative abundance of each class at different stages of development.

The Diseasome 1

It connects diseases that share genes that have altered expression.

- Some of the links and clusters are well known, such as obesity, hypertension, and diabetes.
- Others are surprises, such as Duchenne muscular dystrophy and heart attacks.
  - Following the discovery of this link using gene expression information, researchers are now testing heart drugs on boys with DMD.

The Diseasome 2



Access the text alternative for slide images.

Control of Gene Expression 1

Protein-encoding gene contains some controls over its own expression level.

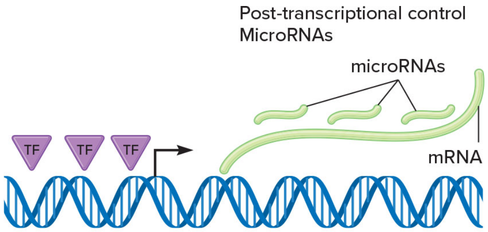
- Promoter sequence (mutations)
- Extra copies of gene

Much of the control of gene expression occurs in two general processes.

- **Chromatin remodeling**—On/off switch
- **microRNAs**—Dimmer switch

Proteins also control each other via **microproteins**

Control of Gene Expression 2



TF = Transcription factor

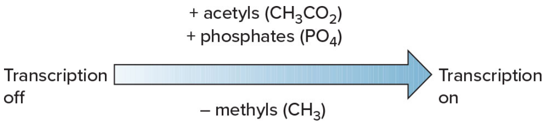
Chromatin Remodeling

Histones play major role in gene expression.

- Expose DNA when and where it is to be transcribed and shield it when it is to be silenced

Major types of small molecules that bind to histones.

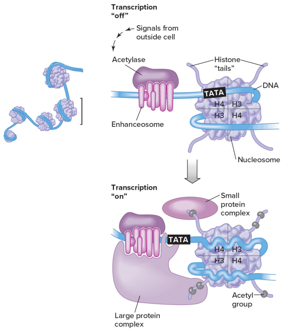
- Acetyl group
- Methyl groups
- Phosphate groups



Acetylated Histones Allow Transcription to Begin

Acetyl binding can subtly shift histone interactions in a way that eases transcription.

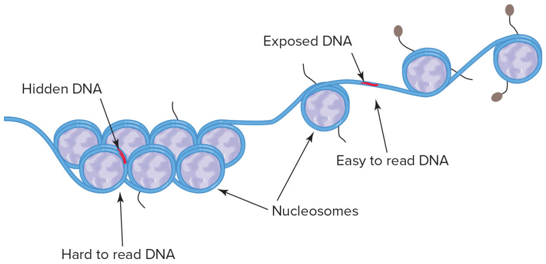
- The TATA box is now exposed, and transcription factors and RNA polymerase can now bind to it



Access the text alternative for slide images.

Closed Chromatin

In most cell types, only about 2 percent of the chromatin is "open" and "easy to read"



MicroRNAs 1

- Belong to a class of molecules called noncoding RNAs
- Just 21 to 22 bases long
- Human genome encodes at least 2,555 distinct sequences of microRNAs that regulate at least 1/3rd of the protein-encoding genes
- A typical human cell has from 1,000 to 200,000 individual microRNA molecules.
- When a microRNA binds to a target mRNA, it prevents translation

MicroRNAs 2

Practical application of microRNAs

- Cancer treatment
  - Certain microRNAs are more or less abundant in cancer cells than in healthy ones.
- Concussion markers
  - MicroRNAs are transported between cells in bubble-like exosomes.
  - 15 microRNAs differ in levels in the cerebrospinal fluid between people with concussion who go on to develop the complication and others with concussion who do not.

Microproteins

Microproteins are fewer than 100 amino acids long.

- They can fit inside the nooks and crannies of large proteins, such as cell surface receptors and ion channels, affecting their functioning.

The genes that encode microproteins may lie within other genes,

- But the microproteins are not just parts of larger proteins, nor are they encoded by microRNAs.

One type of microprotein is like a garbage disposal, mopping up mRNA molecules no longer needed.

- In mice, microproteins control muscles

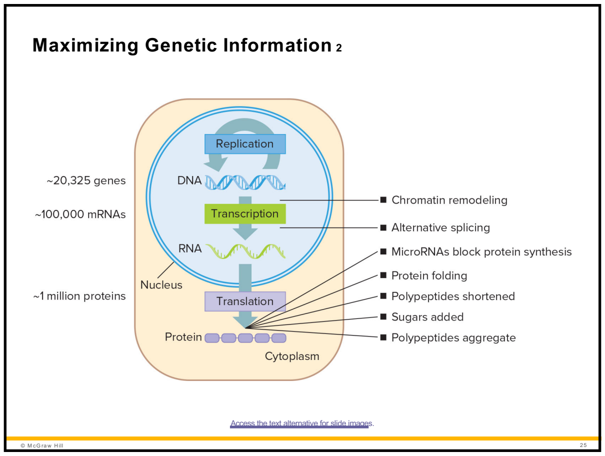
Maximizing Genetic Information 1

Human genome contains about 20,325 genes.

- Encode about 100,000 mRNAs, which in turn specify more than a million proteins

Several events account for the fact that proteins outnumber genes.

- Alternative splicing
- Use of introns
- Protein modification
- Cutting of a precursor protein into two proteins

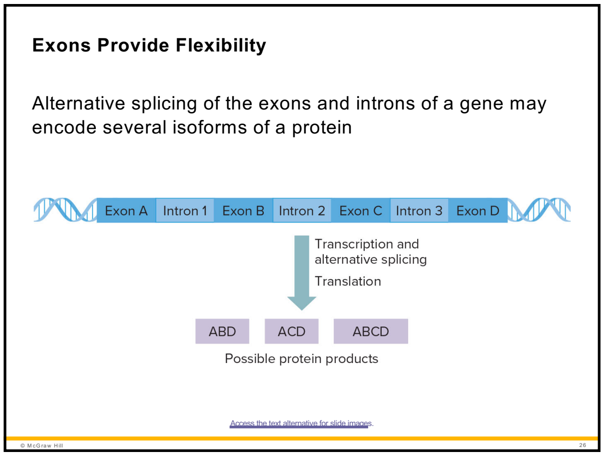


Maximizing Genetic Information 2

Access the text alternative for slide images.

25

25



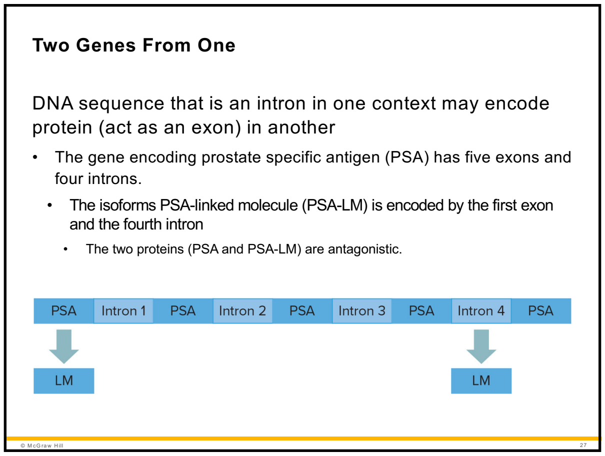
Exons Provide Flexibility

Alternative splicing of the exons and introns of a gene may encode several isoforms of a protein

Access the text alternative for slide images.

26

26



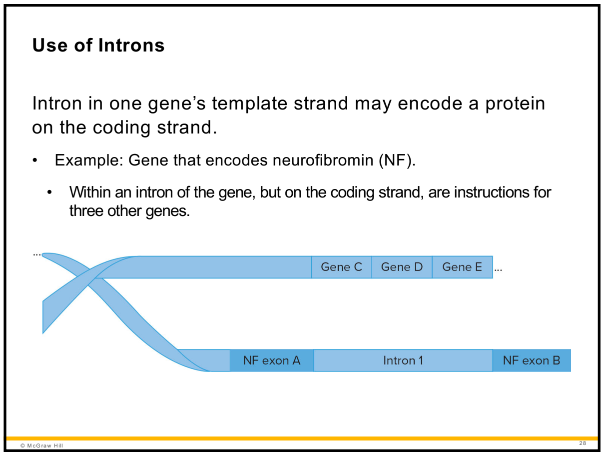
Two Genes From One

DNA sequence that is an intron in one context may encode protein (act as an exon) in another

- The gene encoding prostate specific antigen (PSA) has five exons and four introns.
  - The isoforms PSA-linked molecule (PSA-LM) is encoded by the first exon and the fourth intron
    - The two proteins (PSA and PSA-LM) are antagonistic.

27

27



Use of Introns

Intron in one gene's template strand may encode a protein on the coding strand.

- Example: Gene that encodes neurofibromin (NF).
  - Within an intron of the gene, but on the coding strand, are instructions for three other genes.

28

28

Post-translational Modifications

Information is also maximized when a protein undergoes post-translational modifications.

- Addition of sugars and lipids to create glycoproteins and lipoproteins

Many proteins on cell surfaces have these additions, marking them as part of a particular tissue, organ, and individual.

29

29

Sculpting Two Genes From One 1

Another way that one gene can encode more than one protein is if the protein is cut to yield two products

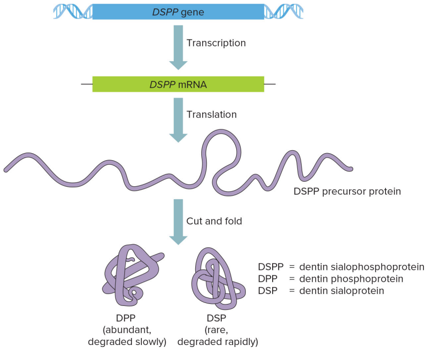
Happens in dentinogenesis imperfecta

- Caused by a deficiency in the two proteins DPP and DSP
- Both are translated from a single mRNA molecule as the precursor protein dentin sialophosphoprotein (DSPP)
  - Then the two proteins are separated

30

30

Sculpting Two Genes From One 2



Access the text alternative for slide images.

Mechanisms That Maximize Genetic Information

Mechanism	Example
Alternative splicing	Alternative splicing
An intron encoding one isoform is an exon in another	Prostate specific antigen (PSA) and PSA-linked molecule (PSA-LM)
An intron on one DNA strand is an exon on the other	Neurofibromin and three other genes
Adding sugars to form glycoproteins or lipids to form lipoproteins	Cell surface molecules important in cell-cell recognition
Precursor protein is cut to yield two proteins	Dentinogenesis imperfecta

Source: Data from the National Institute of Mental Health

Most Human Genomes Do Not Encode Protein

Only 1.5% of human DNA encodes protein

Rest of genome includes:

- Viral DNA
- Noncoding RNAs
- Introns
- Promoters and other control sequences
- Repeated sequences

Most of the genome is transcribed—a DNA sequence is not "junk" if it does not encode protein or we have not discovered a function.

Viral DNA

About 8% of our genome is derived from RNA viruses called **retroviruses**.

- Evidence of past infection
- Sequences tend to increase over time
- The genetic material in the human chromosomes are called human endogenous retroviruses, or HERVs
  - Most HERV sequences have exchanged parts (recombined) and mutated to the extent that they no longer make us sick.
    - Intriguingly, people with certain diseases (forms of AML, MS, and melanoma) have overexpressed HERVs

Noncoding RNAs 1

Nearly all of our genome can be transcribed, and much of it is in the form of noncoding RNA (ncRNAs).

- Includes all RNAs other than mRNAs

tRNAs = Our 500 or so types of tRNA genes account for 0.1 percent of the genome.

rRNAs = 243 types of rRNA genes are grouped on six chromosomes

Hundreds of thousands of other ncRNAs exist

- Transcribed from **pseudogenes**
  - Not translated into protein

Noncoding RNAs 2

Another class of RNA molecules are the 12,000 **long non-coding RNAs**

- 200 or more nucleotides
- Are transcribed from exons, introns, and regions between genes
  - Associate with chromatin and control gene expression
- A third of the long noncoding RNAs are found only in primates, and many only in the brain
  - Thus, they may hold clues to what makes us human

Repeats

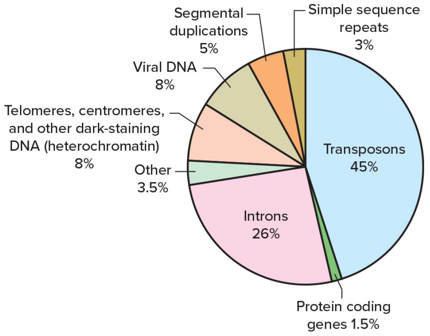
Human genomes are riddled with highly repetitive sequences that hold a different type of information than a protein's amino acid sequence

Transposons are the most abundant type of repeat.

- Sequences that jump about the genome
- Alu repeats can copy themselves
  - Comprise about 2 to 3% of the genome

Rarer classes of repeats are the telomeres, centromeres, and copies of pseudogenes.

Parts of a Human Genome



Access the text alternative for slide images.

Some Non-Protein-Encoding Parts of the Human Genome 1

Type of Sequence	Function or Characteristic
Viral DNA	Evidence of past infection
Noncoding RNA genes	
tRNA genes	Connect mRNA codons to amino acids
rRNA genes	Parts of ribosomes
Long noncoding RNAs	Control of gene expression
Pseudogenes	DNA sequences very similar to known genes that are not translated
Piwi-interacting RNA (piRNA)	Keeps transposons out of germline cells
Large intergenic noncoding RNAs	Between genes
Small nucleolar RNAs (snoRNAs)	Process rRNAs in nucleolus
Small nuclear RNAs (snRNAs)	Parts of spliceosomes
Telomerase RNA	Adds bases to chromosome tips
Xist RNA	Inactivates one X chromosome in cells of females

Some Non-Protein-Encoding Parts of the Human Genome 2

Type of Sequence	Function or Characteristic
Introns	Parts of genes that are cut out of mRNA
Promoters and other control sequences	Guide enzymes that carry out DNA replication, transcription, or translation
Small interfering RNAs (siRNAs)	Control translation
MicroRNAs (miRNAs)	Control translation of many genes
Repeats	
Transposons	Repeats that move around the genome
Telomeres	Protect chromosome tips
Centromeres	Largest constrictions in chromosomes, providing attachment points for spindle fibers
Segmental duplications	Repeated DNA sequences of >1,000 bases long
Simple sequences (Alu and others)	Functions unknown

Mc Graw Hill

Because learning changes everything.®

www.mheducation.com