Gene expression

- Gene function at the level of traits
- Gene function at the molecular level

- Two levels tied together since the molecular level affects the structure and function of cells which determines what traits are expressed
One gene – one enzyme hypothesis has been modified.

Enzymes are only one category of cellular proteins.

More accurate to say one gene encodes a polypeptide.

- Hemoglobin composed of 4 polypeptides required for function.

One gene – one polypeptide theory.
Central dogma

- Transcription
  - Produces an RNA copy or transcript of a gene
  - Structural genes produce messenger RNA (mRNA) that specifies the amino acid sequence of a polypeptide

- Translation
  - Process of synthesizing specific polypeptide on a ribosome

- Eukaryotes have additional intervening step called RNA processing where pre-mRNA is processed into functionally active mRNA
(a) Molecular gene expression in prokaryotes

Transcription: DNA is transcribed into an RNA copy.

Translation: mRNA is translated into a polypeptide at the ribosome.

(b) Molecular gene expression in eukaryotes

Transcription

RNA processing

Entry into cytosol

Translation

Cytosol
Nucleus
DNA
Pre-mRNA
mRNA

Polypeptide
Ribosome
Genes contain the information necessary to make the organism and allow it to favorably interact with its environment

- Structural genes code for polypeptides
- Polypeptide becomes a unit of function or protein
- Activities of proteins determine structure and function of cells
- Traits or characteristics of organism based on cellular activities
The normal gene is transcribed into a functional mRNA molecule.

The mRNA is translated into pigment precursor transport protein.

Colorless pigment precursor molecules are transported into the cell.

Intracellular enzymes convert colorless pigment precursor molecules into red pigment.

Mutation may prevent transcription of functional mRNA.

Mutation may prevent translation of functional pigment precursor transport protein.

Without a functional transporter, colorless pigment precursor molecules are not transported into the cell.
Transcription

- A gene is an organized unit of DNA sequences that enables a segment of DNA to be transcribed into RNA and ultimately results in the formation of a functional product.
- Over 90% of all genes are structural.
- Other genes code for the RNA itself:
  - Transfer RNA (tRNA) – translates mRNA into amino acids.
  - Ribosomal RNA (rRNA) - part of ribosomes.
DNA

Promoter: Site for RNA polymerase binding; signals the beginning of transcription.

Regulatory sequence: Site for the binding of regulatory proteins. The role of regulatory proteins is to influence the rate of transcription. Regulatory sequences can be found in a variety of locations.

Transcribed region: Contains the information that specifies an amino acid sequence.

Terminator: Signals the end of transcription.
Three stages

1. Initiation
2. Elongation
3. Termination
Three stages

1. Initiation
   - Recognition step
   - In bacteria, sigma factor causes RNA polymerase to recognize promoter region
   - Catalytic portion of RNA polymerase has similar structure in all species
2. Elongation

- RNA polymerase synthesizes RNA after release of sigma factor
- Open complex 10-15 base pairs long
- Template or coding strand used for RNA synthesis
  - Noncoding strand is not used
- Synthesized 5’ to 3’
- Uracil substituted for thymine
- Behind open complex, DNA rewinds
3. **Termination**

- RNA polymerase reaches termination sequence
- Causes it and newly made RNA transcript to dissociate from DNA
1. **Initiation:**
   The promoter functions as a recognition site for sigma factor. RNA polymerase is bound to sigma factor, which causes it to recognize the promoter. Following binding, the DNA is unwound into a bubble known as the open complex.

2. **Elongation/synthesis of the RNA transcript:**
   Sigma factor is released and RNA polymerase slides along the DNA in an open complex to synthesize RNA.

3. **Termination:**
   When RNA polymerase reaches the terminator, it and the RNA transcript dissociate from the DNA.

(a) **Stages of transcription**

(b) **Structure of a bacterial RNA polymerase**
- Direction of transcription and DNA strand used varies among genes
- In all cases, synthesis of RNA transcript is 5’ to 3’ and DNA template strand reads 3’ to 5’
Eukaryotic transcription

- Basic features identical to prokaryotes
- However, each step has more proteins
- 3 forms of RNA polymerase
  - RNA polymerase II – transcribes mRNA
  - RNA polymerase I and III – transcribes nonstructural genes for rRNA and tRNA
- RNA polymerase II requires 5 general transcription factors to initiate transcription
RNA processing

- Bacterial mRNAs can be translated into polypeptides as soon as they are made
- Eukaryotic mRNAs are made in a longer pre-mRNA form that requires processing into mature mRNA
- Introns - transcribed but not translated
- Exons - coding sequence found in mature mRNA
- Splicing - removal of introns and connection of exons
- Other modifications also occur – addition of tails and caps
1. Mature mRNA binds to template strand, which causes the intron DNA to loop out.

2. The coding strand then binds to the template strand, but it loops out where the RNA is already bound.

(a) Schematic drawing of mature mRNA binding to DNA containing an intron

(b) Results after electron microscopy

- Introns found in many eukaryotic genes
- Widespread in complex eukaryotes
  - Human dystrophin gene has 79 exons and 78 introns
- Introns are rare in all prokaryotes
- Introns removed from eukaryotic pre-mRNA using a spliceosome composed of snRNPs (snurps)
  - Nucleur RNA and a set of proteins
Intron RNA is defined by particular sequences within the intron and at the intron-exon boundaries
- 5’ splice site, branch site, 3’ splice site

Spliceosome subunits recognize intron sequences

Binding causes intron to loop outward

Two exons brought closer together

Alternative splicing - function of spliceosome regulated so that single gene can encode 2 or more polypeptides

Introns in rRNA and tRNA are self-splicing
1. First two subunits bind to 5' splice site and branch site.

2. Additional subunits bind, creating a loop.

3. 5' splice site is cut. 3' end of intron is connected to the branch site. Two subunits are released.

4. 3' splice site is cut. Exon 1 is connected to exon 2. The intron (in the form of a loop) is released along with the rest of the subunits and degraded.
- 5’ cap and capping
  - 7-methylguanosine covalently attached to 5’ end
  - Needed for proper exit of mRNA from nucleus and binding to ribosome

- 3’ poly A tail
  - 100-200 adenine nucleotides added to 3’ end
  - Increases stability and lifespan in cytosol
  - Not encoded in gene sequence
(a) Cap structure at the 5′ end of eukaryotic mRNA

(b) Addition of a poly A tail at the 3′ end of eukaryotic mRNA
Translation

- Genetic code – sequence of bases in an mRNA molecule
- Read in groups of three nucleotide bases or codons
- Most codons specify a particular amino acid
  - Also start and stop codons
- Degenerate - more than one codon can specify the same amino acid
<table>
<thead>
<tr>
<th>First position</th>
<th>Second position</th>
<th>Third position</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>UU UU Leu</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>UUC</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>UUA UUG</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>CUU CCU Leu</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>CUC CCC</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>CUA CCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CUG CCG</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>AUU ACU Ile</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>AUC ACC Thr</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>AUA ACA Met/stop</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUG ACG</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>GUU GCU Val</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>GUC GCC</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>GUA GCA</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>GUG GCG</td>
<td>G</td>
</tr>
</tbody>
</table>

*Exceptions to the genetic code are sporadically found among various species. A few examples are as follows: AUA encodes methionine in yeast and mammalian mitochondria; UGA encodes tryptophan in mammalian mitochondria; CUU, CUA, CUC, and CUG encode threonine in yeast mitochondria; AGA and AGG act as stop codons in ciliated protozoa and in yeast and mammalian mitochondria; and UAA and UAG encode glutamine in ciliated protozoa.
- Bacterial mRNA
- 5’ ribosomal-binding site
- Start codon usually AUG
- Typical polypeptide is a few hundred amino acids in length
- 1 of 3 stop codons
  - Termination or nonsense codons
  - UAA, UAG or UGA
Reading frame

- Start codon defines reading frame

5’ –AUAAGGAGGGUUACG(AUG)(CAG)(CAG)(GGC)(UUU)(ACC) – 3’
  Met   -Gln   -Gln   -Gly   -Phe   -Thr

- Addition of a U shifts the reading frame and changes the codons and amino acids specified

5’ –AUAAGGAGGGUUACG(UCA)(GCA)(GGG)(CUU)(UAC)C – 3’
  Met   -Ser   -Ala   -Gly   -Leu   -Tyr
DNA sequence of gene transcribed into mRNA

mRNA
- Codons
- T of DNA substituted for U of RNA

tRNA
- Anticodon allows binding of tRNA to mRNA codon
tRNA

- Different tRNA molecules encoded by different genes
- $\text{tRNA}^\text{ser}$ carries serine
- Common features
  - Cloverleaf structure
  - Anticodon
  - Acceptor stem for amino acid binding
Aminoacyl-tRNA synthetase catalyzes the attachment of amino acids to tRNA

One for each of 20 different amino acids

2 reactions result in tRNA with amino acid attached or charged tRNA or aminoacyl tRNA

Ability of aminoacyl-tRNA synthetase to recognize appropriate tRNA has been called the second genetic code
1. A specific amino acid and ATP bind to the aminoacyl-tRNA synthetase.

2. AMP is covalently bound to the amino acid thereby activating it, and pyrophosphate is released.

3. The correct tRNA binds to the synthetase. The amino acid is covalently attached to the tRNA. AMP is released.

4. The “charged” tRNA is released.
Ribosomes

- Prokaryotes have one kind
- Eukaryotes have distinct ribosomes in different cellular compartments
  - Focus on cytosolic ribosomes
- Composed of large and small subunits
- Structural differences between prokaryotes and eukaryotes exploited by antibiotics to inhibit bacterial ribosomes only
(a) Bacterial ribosomes

(b) Eukaryotic ribosomes
- Overall ribosome shape determined by rRNA
- Discrete sites for tRNA binding and polypeptide synthesis
- P site - peptidyl site
- A site - aminoacyl site
- E site - exit site
Stages of Translation

1. Initiation
   - mRNA, first tRNA and ribosomal subunits assemble

2. Elongation
   - Synthesis from start codon to stop codon

3. Termination
   - Complex disassembles at stop codon releasing completed polypeptide
1. **Initiation:** mRNA, tRNA, and the ribosomal subunits form a complex.

2. **Elongation:** The ribosome travels in the 5’ to 3’ direction and synthesizes a polypeptide.

3. **Termination:** The ribosome reaches a stop codon and all of the components disassemble, releasing a completed polypeptide.

Translation: mRNA is translated at the ribosome into a polypeptide.
Initiation

- mRNA, first tRNA and ribosomal subunits assemble
- Requires help of ribosomal initiation factors
- Also requires input of energy (GTP hydrolysis)
Bacteria
- mRNA binds to small ribosomal subunit facilitated by ribosomal-binding sequence
- Start codon a few nucleotides downstream
- Initiator tRNA recognizes start codon in mRNA
- Large ribosomal subunit associates
- At the end, the initiator tRNA is in the P site
1. Small ribosomal subunit binds to mRNA.

2. Initiator tRNA binds to start codon in mRNA.

3. Large ribosomal subunit binds.
Eukaryotic differences in initiation

- Instead of a ribosomal-binding sequence, mRNAs have 7-methylguanosine cap at 5’ end
  - Recognized by cap-binding proteins
- Position of start codon more variable
  - In many cases, first AUG codon used as start codon
Elongation

1. Aminoacyl tRNA brings a new amino acid to the A site
   - Binding occurs due to codon/anticodon recognition
   - Elongation factors hydrolzye GTP to provide energy to bind tRNA to A site
   - Peptidyl tRNA is in the P site
   - Aminoacyl tRNA is in the A site
2. A peptide bond is formed between the amino acid at the A site and the growing polypeptide chain
   - The polypeptide is removed from the tRNA in the P site and transferred to the amino acid at the A site – peptidyl transfer reaction
3. Movement or translocation of the ribosome toward the 3’ end of the mRNA by one codon

- Shifts tRNAs at the P and A sites to the E and P sites
- The next codon is now at the A spot
- Uncharged tRNA exits from E spot
1. **tRNA entry**: A charged tRNA binds to the A site.

2. **Peptidyl transfer reaction**: A bond forms between the polypeptide chain and the amino acid in the A site. The polypeptide is transferred to the A site.

3. **Translocation of ribosome and release of tRNA**: The ribosome translocates 1 codon to the right. The uncharged tRNA is released from the E site. This process is repeated again and again until a stop codon is reached.
Termination

- When a stop codon is found in the A site, translation ends
- 3 stop codons- UAA, UAG, UGA
- Recognized by release factors
Completed polypeptide attached to a tRNA in the P site and stop codon in the A site
1. Release factor binds to stop codon at the A site
2. Bond between polypeptide and tRNA hydrolyzed to release polypeptide
3. Ribosomal subunits and release factors disassociate
1. A release factor binds to the A site, where a stop codon is found.

2. The polypeptide is released from the tRNA in the P site. The tRNA is then released.

3. The ribosomal subunits, mRNA and release factor dissociate.
- Polypeptide chain has directionality
- Parallels 5’ to 3’ orientation of mRNA
- N-terminus or amino terminus is first amino acid
- C-terminus or carboxyl terminus is last amino acid