Gene Expression: Transcription and Translation, Chapters 5 and 6
Gene Expression

• Genes “control” phenotype
• Cellular environment affects gene expression
• Involves extensive intracellular signaling
• Tissue, development, disease specific
THE CENTRAL DOGMA

DNA $\rightarrow$ RNA $\rightarrow$ PROTEIN

Which part is gene transcription? translation?
Transcription – Copy DNA to make mRNA

DNA + RNA polymerase + nucleotides → mRNA

template enzyme building blocks transcript

nucleotides in RNA _______
mRNA is _____ stranded
Note that synthesis is 5’ -> 3’
I. TRANSCRIPTION (prokaryotes)

A. INITIATION of Transcription

RNA polymerase recognizes gene promoter

holoenzyme
**Promoters** = regulatory elements upstream of coding region of gene (+1) recognized by RNA polymerase

Prokaryotes: -10 box (Pribnow) and -35 box
• **Consensus sequence** of promoters
• If promoter has a slightly different sequence, it will not be as “strong”
• **sigma factor** (σ) of RNA pol binds promoter

• core polymerase locked on to form **holoenzyme**

• A variety of sigma factors
The **template strand** of ds DNA is “read”

\[
\begin{align*}
5' \text{GGTATAACCTACC} & \quad 3' \text{ coding/sense} \\
3' \text{CCATATGGATGG} & \quad 5' \text{ template} \\
\text{mRNA} & \quad 5' \text{G}
\end{align*}
\]

RNA polymerase works 5’ \rightarrow 3’
practice

5’ GTTAACCCCATATGCTAGGGGGG 3’

mRNA transcript?
TRANSCRIPTION: Initiation and Elongation

http://www.youtube.com/watch?v=41_Ne5mS2ls&feature=related
B. ELONGATION of mRNA transcript

- RNA pol - synthesizes mRNA from template DNA
  - DNA exposed one base at a time (transcription bubble)
  - Template strand read to make mRNA

http://www.scientificpsychic.com/fitness/transcription.gif
Once RNA pol has moved, another RNA pol molecule can initiate to make another mRNA transcript. What does this mean?

**Genetics institute** – promoter and termination sites
C. TERMINATION

Termination sequences

RNA pol loosens

mRNA dissociates

http://www.biog1105-1106.org/demos/106/unit02/media/termination-signal.jpg
Polymerase fame

Severo Ochoa
1/2 of the prize

Arthur Kornberg
1/2 of the prize
II. Gene expression in eukaryotes

List the differences
contrast prokaryotes with eukaryotes

PROKARYOTES
• no nucleus
• 1 RNA polymerase
• cotranslation
• Polycistronic mRNA
• One promoter shared by a few genes
• All of gene is coding = no modification of mRNA
Eukaryotes - differences

1. **3 RNA polymerases**
   - RNA pol I – transcribes rRNA genes
   - RNA pol II – transcribes mRNA and snRNAs
   - RNA pol III – transcribes tRNA genes

2. **Split genes**
   - Genes have coding and non-coding sequences
   - Non-coding removed from mRNA

3. **transcription and translation in separate compartments**

4. **Different regulatory elements**
Split genes in eukaryotes

- **exons** - coding (expressed sequences)
- **introns** = intervening sequences
  - non-coding but transcribed into mRNA

*Takifugu rubripes* has few introns and far less repetitive DNA

400 million bp genome but same # genes as humans
Split gene fame

- Philip Sharp and Richard Roberts (independent) Nobel 1993 for discovery of split genes
Chick collagen = 37kb with 50 introns

mRNA/DNA hybrid shows introns as loops
Eukaryotic Promoters
TATA promoter at -25

Where is the transcription start site of the gene (base #1)?
What are the black zig zag lines?

TATAAAA
β-globin

SV40, early

thymidine kinase

histone H2B

TATA box = (TATAAAA)

GC box = (GGGCGGG)

CAAT box = (GGCCCAATCT)

Octamer = (ATTTGACAT)
**Promoter (s)** = recognized by RNA polymerase II

**UTR** = untranslated region (transcribed, not translated)

**Start** – start site for transcription

**Introns** = non coding sequences within gene

**Exons** = coding regions
4 types of genes

- **protein coding** -> mRNAs -> protein
- tRNA genes -> transfer RNAs
- r RNA genes -> ribosomal RNAs
- snRNA genes -> small nuclear RNAs
• TRANSCRIPTION OF PROTEIN CODING GENES
Initiation of transcription in eukaryotes

2 types of proteins required:

1. **Transcription factors** allow RNA pol II to bind promoter

2. RNA polymerase II
Stages of transcription

Initiation, Elongation, Termination

VIRTUAL CELL

http://vcell.ndsu.nodak.edu/animations/transcription/first.htm

TATA box, enhancer
TFs in general
RNA pol enzyme
mRNA

Also  http://www.youtube.com/watch?v=bk7PW1FKMTI

enhancer, 7G, poly A tail, TF, spliceosome
Pre-mRNA must be processed to mature mRNA
Pre-mRNA is modified:

- At 5’ and 3’ ends
- Remove introns
The European sea bass DAX1 gene structure
Martins et al. Reproductive Biology and Endocrinology 2007 5:19
A. Addition of 7 methyl guanosine cap

• added to 5’ end of mRNA transcript
• increases half life of mRNA (stability)
• recognition site for ribosome binding
A. Addition of 7 methyl guanosine cap

- Added to 5’ end of mRNA transcript
- Increases half life of mRNA (stability)
- Recognition site for ribosome binding
B. Addition of the Poly A tail

- Poly A tail of ~200 A’s added to end of transcript
- Increases the half life of the mRNA
Capping and tailing animation

- [http://vcell.ndsu.nodak.edu/animations/mrna processing/movie.htm](http://vcell.ndsu.nodak.edu/animations/mrna processing/movie.htm)

  Transcription → 7mG cap at 5’ end → termination → cleavage factors → poly A tail at 3’ end → splicing
C. Removal of introns

- DNA Genome
- Primary RNA Transcript
- 5' Cap
- INTRON
- EXON
- Mature mRNA
- Export to cytoplasm

Steps:
1. Addition of 5'-cap
2. Addition of poly [A] tail
3. Removal of 5 introns (splicing)
4. Removal of 2 introns (splicing)
5. 1,872 bases
6. Export to cytoplasm
• What is the size in bases of the fully processed mature mRNA?
Expression of the b-globin gene is a typical process. This gene contains two introns and three exons. Interestingly, the codon of the 30th amino acid, AGG, is separated by an intron. As a result, the first two nucleotides AG are in one exon and the third nucleotide G is in another exon.
• Splicing excises introns at specific exon/intron junction sites. Joins exons.

GU……AG
Splice sites
Intron information

• Variable in size
• May regulate gene expression
• Accumulate mutations
• Alternate splicing = more than one protein from a gene
• GU AG consensus sequences
• Some genes do not contain introns
Intron removal

• [http://vcell.ndsu.nodak.edu/animations/mrna_splicing/movie.htm](http://vcell.ndsu.nodak.edu/animations/mrna_splicing/movie.htm)
Components of the spliceosome

- snRNPs + mRNA = spliceosome

“Snurps” are sn RNAs + proteins

- small nuclear RNAs (U1, U2, U4, U5, U6 ~ 100 bases)
RNA world hypothesis

• RNA can possess enzymatic (protein) activity
• RNA-based life would be able to engage in enzymatic activities and store genetic information
Gene mutations and beta thalassemia

• Beta thalassemia
  – Mutation beta globin gene
  – Inherited blood disorder (autosomal recessive)
  – Can cause mild → severe anemia, poor appetite, slow growth, bone abnormalities, enlarged heart
  – Nucleated (immature) red blood cells
  – More than 200 mutations identified

About 20 new cases each year in US
Thalassemia

- $B^0 \rightarrow$ no beta globin produced $\rightarrow$ no Hb
- $B^+ \rightarrow$ low Hb production
Single base change (point) mutations

Missense Mutations

<table>
<thead>
<tr>
<th>ATG</th>
<th>GAA</th>
<th>GCA</th>
<th>CGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>Glu</td>
<td>Ala</td>
<td>Gly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ATG</th>
<th>GAC</th>
<th>GCA</th>
<th>CGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>Asp</td>
<td>Ala</td>
<td>Gly</td>
</tr>
</tbody>
</table>

Silent Mutations

<table>
<thead>
<tr>
<th>ATG</th>
<th>GAA</th>
<th>GCA</th>
<th>CGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>Glu</td>
<td>Ala</td>
<td>Gly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ATG</th>
<th>GAG</th>
<th>GCA</th>
<th>CGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>Glu</td>
<td>Ala</td>
<td>Gly</td>
</tr>
</tbody>
</table>

http://members.cox.net/amgough/Fanconi-genetics-genetics-genetics-primer.htm#Nonsense
Single base change (point) mutations

Frameshift Mutation

ATG | GAA | GCA | CGT
Met | Glu | Ala | Gly

ATG | AAG | CAC | GT
Met | Lys | His

Nonsense Mutations

ATG | GAA | GCA | CGT
Met | Glu | Ala | Gly

ATG | TAA | GCA | CGT
Met | STOP
Normal beta chain
ATG GTG CAC CTG ACT CCT GAG GAG AAG TCT GCC GTT ACT GCC CTG TGG GCC AAG GTG AAC GTG GAT GAA GTT
  Val His Len Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Len Trp Gly Lys Val Asn Val Asp Glu Val

HbS Sickle cell (missense)
ATG GTG CAC CTG ACT CCT GTG GAG AAG TCT GCC GTT ACT GCC CTG TGG GCC AAG GTG AAC GTG GAT GAA GTT
  Val His Len Thr Pro Val Glu Lys Ser Ala Val Thr Ala Len Trp Gly Lys Val Asn Val Asp Glu Val

HbC (missense)
ATG GTG CAC CTG ACT CCT AAG GAG AAG TCT GCC GTT ACT GCC CTG TGG GCC AAG GTG AAC GTG GAT GAA GTT
  Val His Len Thr Pro Lys Glu Lys Ser Ala Val Thr Ala Len Trp Gly Lys Val Asn Val Asp Glu Val

HbThalassemia (nonsense)
ATG GTG CAC CTG ACT CCT GAG GAG AAG TCT GCC GTT ACT GCC CTG TGG GCC TAG GTG AAC GTG GAT GAA GTT
  Val His Len Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Len Trp Gly Stop

HbThalassemia (frameshift) -AA
ATG GTG CAC CTG ACT CCT GAG GAG GTC TGC CGT TAG TGC CCT GTG GGG CAA GGT GAA CGT GGA TGA AGT TGG
  Val His Len Thr Pro Glu Glu Val Cys Arg Tyr Cys Pro Val Gly Gln Gly Gln Arg Ala Stop

Stop codon
Frameshift mutation
Missense mutation
Normal intron removal

AACAUGCUUCUGACACAAACUGUGUUCACUCAGCACCUCACAAACAGACACCCGUGCACCCUGACCUCCUUGAGAGAAG
(met) val his leu thr pro glu glu lys

UCUGCCGUUACUGCCCGUGUUGGCGCAAGGUGAAACGUGGAGAAGUUGGUGGAGGCCUUGGCAAGGUUGUAUCAAAG
ser ala val thr ala leu trp gly lys val asn val asp glu val gly gly glu ala leu gly arg

GUUACAAAGACAGGGUUUAGGGACACCAAUAGAAACUGGGCAUGUGGAGACAGAGAAGACUCUUUGGGUUUCUGAUAGGC

ACUGACUCUCUCUGCCUAUUGGUCUAUUUCCACCCCUUAAGGCUGCGUGGUUGGUCUACCCUUGGACCCCAAGGUGUUCUUU
leu leu val val tyr pro trp thr glu arg phe phe

GAGUCCUUUGGGAUCUGUCCACUCCUGAUGCUGUUAUGGGAACCCUAAGGUAGGCUCACUGGCAGAAAGUG
glu ser phe gly asp leu ser thr pro asp ala val met gly asn pro lys val lys ala his gly lys lys val

45
50
55
60
65

CUCGGUGCCUUUAGUGAUGCUGGCGCCUGACACCUGGACACCCUAAAGGCGACCACCUUGCCACACUGAGUGAGCUGCAC
leu gly ala phe ser asp gly leu ala his leu asp aln leu lys gly thr phe ala thr leu ser glu leu his

70
75
80
85
90

UGUGACAAAGCGUGCAUGGAAUCUGAGAACUAAGGUGAGUCUAGGACCUCUUUGGAUGUUUCUUUCUCCCUUCCUUU

cys asp lys leu his val asp pro glu asp phe arg
Loss of splice site
Cis elements and trans factors

• **Cis elements**
  – DNA sequences near gene that are required for gene expression
  – Example:

• **Trans factors** – proteins that bind to DNA
  – Example:
DNA (cis factor) or protein (trans factor)?

Promoter
Transcription factors
Enhancer
TATA, CAAT, and GC
-10 and -35 box
RNA polymerase II

Which are found in prokaryotes? Eukaryotes? Both?
VII. TRANSLATION

mRNA  \rightarrow  protein

Nucleotides  \rightarrow  amino acids

Translation
http://vcell.ndsu.nodak.edu/animations/translation/movie.htm
Translation end product = PROTEINS

a. Amino acid – building block of protein
   • 20 in nature
b. Peptide – short stretch of amino acids

c. Polypeptide – the fully translated message

d. Protein – functional, 3D shape
The genetic code

• Nirenberg and Khorana 1965 → Nobel Prize

1. 4 nucleotide bases encode 20aa
   → TRIPLET CODE
   64 possible codons
What is this amino acid sequence?
2. Non-overlapping
3. Degenerate
   More than 1 codon per particular amino acid
4. Universal
   A few differences in mitochondria, yeast
Most polymorphisms are in the 3rd position of codon — Why?
5. 61 sense codons
   start codon AUG (methionine)
   • Initiation of translation
   • All proteins start with met, in some removed post translation

6. 3 stop codons UAA, UAG, UGA
Which part is the 5’ UTR of the mRNA?
Example of the genetic code

A partial mRNA is as follows:

5’ AGGAGGCUCGAACAUGUCAAAUAUGCUUUGUCCUGACGC 3’

• start site for translation?
• open reading frame?
C. Ribosomes translate the mRNA to amino acid sequence

- 100,000s /cell
- protein + rRNA
D. Translation steps

1. Initiation
   • mRNA read 5’ -> 3’ initiating at first AUG
   • IFs involved
2. Elongation

- “Charged” tRNAs bring amino acid to ribosome
Anticodon: 3' CGG 5'
Codon: 5' GCC 3'
tRNA cloverleaf structure – how?
• 2 codons can “sit” in a ribosome at once

Peptide bond between aa
- wobble \( \rightarrow \) 30 tRNAs for 61 codons
DNA coding  5’ATG  ACT  AGC  TGG GGG ATG  TAC  TTT  TAG  3’
DNA template

mRNA

tRNA

anticodon

aa
3. **Termination** of translation

- Stop codon
- RFs
Polysome = many ribosomes can translate a single message to produce many copies of a protein
• polysomes
2. Protein structure

Primary structure – sequence of aa

- Secondary structure – spontaneous folds
• **Tertiary structure**
  – 3D

• **Quaternary structure**
  – Some proteins have multiple subunits
  – Example: hemoglobin
Try it

• **Translate a sequence Utah**

  Translation tool

The anticodon is 5’ ACG 3’

What is the codon and the amino acid coded for?

Amino acid?

The anticodon is 3’ GGC 5’

Amino acid?

Codon  CGU
Amino acid  arginine

Codon  CCG
Amino acid  alanine
VIII. Post translational modification - to the endoplasmic reticulum for processing
IX. CONTROL OF GENE EXPRESSION
1. Hormonal control of gene expression

![Diagram of Hormonal Control of Gene Expression](image)

- **Induction of Specific Gene Expression by a Steroid Hormone**
- **Stimulated** (e.g., conalbumin in the hen oviduct)
- **Not Stimulated** (e.g., glucose-6-phosphate dehydrogenase in the hen oviduct)

- **Gene Expression**
- **Time Course**

- **Addition of hormone** (e.g., an estrogen)
3. Splicing of mRNA

alternate splicing to form protein isoforms

The fetal acetylcholine receptor, has variable kinetics during postnatal development. RT-PCR revealed, in addition to the full-length mRNA, three new forms lacking exon 4. One also lacks 19 nucleotides from exon 5, with a 43 residues shorter N-terminus. A third one lacking the complete exon 5 predicts a subunit without transmembrane segments. These forms, generated by alternative splicing, may account for the kinetic variability of the acetylcholine receptor channel.
4. mRNA half life

poly A tail and 5’ cap. The longer the mRNA is in the cytoplasm, the more protein
5. RNA interference (RNAi)

- Silencing of mRNA by tiny homologous dsRNA
- → mRNA degradation
- RNAi can selectively turn off gene expression
RNAi – dsRNA complementary to red eye pigment mRNA microinjected into *Drosophila* embryos
6. Promoter strength
   – Number of transcripts made

end