Gene Expression: How DNA sequence information dictates form and function of organisms

DNA sequence of a genome determine phenotype through control of these properties of RNA and proteins:

- Temporal (timing of synthesis/degredation)
- Quantitative (how much = rates of synthesis – rates of degradation)
- Structural - sequence of amino acids (proteins) or bases (RNA)
  - RNA = mRNA, tRNA, and rRNA
Fundamental processes of gene expression: transcription

- **Transcription** - synthesis of RNA copies of the information in a DNA subsequence
  - mRNA is capped, extended, and spliced before being moved out of the nucleus into the cytosol where it is translated into a polypeptide according to the genetic code
  - rRNA leaves the nucleus and combines with polypeptides to form mRNA-directed protein synthesizing machines called “ribosomes”
  - tRNA are modified and become the agents that deliver the correct AA to the growing polypeptide during translation
Fundamental processes of gene expression: translation

Translation - synthesis of polypeptides whose AA sequences are colinear with their respective codons in both the mRNA and the sense strand of the corresponding DNA

Proteins often subjected to further processing to alter functionality or cellular location
Flow of information from DNA to Phenotype via mRNA and proteins

Enzymes, transcription control factors, structural components of cells
The structure of RNA

(a) The separate entities

1. The sugar: Ribose instead of deoxyribose

![Ribose and Deoxyribose structures]

2. A phosphate group

![Phosphate structure]

3. The four bases

- Uracil (U) instead of thymine (T)
- Plus Adenine, Guanine, Cytosine

(b) Assembly into a ribonucleotide

![RNA molecule]

(c) Ribonucleotides join to form a single strand of ribonucleotides
Properties of RNA

- Precursors of RNA polynucleotides are the four ribonucleosides
  - ATP, GTP, CTP, and UTP
- RNA generally exists as single strands
- Sugar-phosphate backbone similar to DNA
- RNA has 5'/3' polarity as in DNA
- RNA synthesis is 5' → 3' as in DNA
- Template for RNA synthesis is DNA, as in DNA replication
RNA is polymerized by DNA dependent (or directed) RNA polymerases in Transcription

- Similar to DNA replication polymerases, except
  - No primer needed
  - Ribonucleoside triphosphates used as building blocks
  - Only one strand of dsDNA is used as template
- RNA pol II creates mRNA
- RNA pol I and III create rRNA = ribosomal RNA
• Precursors of RNA polynucleotides are the four ribonucleosides
  • ATP, GTP, CTP, and UTP
• RNA generally exists as single strands
• Sugar-phosphate backbone similar to DNA
• RNA has 5’ / 3’ polarity as in DNA
• RNA synthesis is 5’ → 3’ as in DNA
• Template for RNA synthesis is DNA, as in DNA replication
• RNA polymerization similar to DNA replication polymerases, except
  • No primer needed
  • Ribonucleoside triphosphates used as building blocks
  • Only one strand of dsDNA is used as template
• RNA pol II creates mRNA
• RNA pol I and III create rRNA = ribosomal RNA
Transcription, part 2

(c) Termination

3' 5'
mRNA
5' 3'

Termination region

Termination signal

RNA polymerase released at terminator

(d) Information flow

DNA
5' ATCCAGGAAACGCTTAAT CTTTCAAAATCTTATAG 3'
3' TACCACTGCGTACAGTTA... TAACATAT GC

Template strand

RNA
5' AUGGAGGAAAGCGUCAAU... AUUGUUAAG 3'

Primary transcript
• NOTES ON TRANSCRIPTION (for previous two slides)- ignoring tRNA and rRNA
  - sequences immediately upstream of the transcribed sequence of DNA are recognized by RNA polymerase;
  - Initiation:
    - RNA polymerase holoenzyme (intact polymerase) and ancillary transcription co-factors (proteins) attaches to the promoter sequence (still double stranded)
    - co-factors released
  - Elongation:
    - RNA polymerase binds strongly to the template strand
    - a transcription bubble opens as the RNA polymerase creates a single stranded region in the DNA downstream from the promoter (relative to the template strand)
    - RNA is synthesized according to the base pairing rules of A:U G:C
    - RNA nucleotides added at 40+ per second to growing RNA polymer
    - region behind bubble (which travels in a 3’ - 5’ direction on the template strand) becomes dsDNA as the bubble moves
    - 3 parts now: DNA, RNA polymerase, and growing RNA polymer
    - multiple transcription bubbles can co-exist on a single gene
  - Termination:
    - specific sequences in the primary transcript cause termination of transcription, i.e., disassociation of the RNA polymerase from the DNA molecule

• POST TRANSCRIPTION PROCESSING (EUKARYOTIC GENES)
  - primary transcripts (initial product of transcription) of eukaryotic genes are processed in 3 fundamental ways:
    - the 5’ end is “Capped” (see this slide)
    - 100-200 adenines are added to the 3’ end of the primary transcript 11-30 nucleotides downstream from the sequence AAUAAA has a poly-A (adenine) tail added (next slide)
    - introns removed and exons spliced together
  - a capped, poly-A, spliced primary transcript is now called “mature mRNA” and is exported out of the nucleus for translation
poly A tail addition to the 3' end

- RNA polymerase
- AAUAAA
- Cleavage by ribonuclease
- 5' cap
- AAUAAA
- poly-A polymerase adds A's onto 3' end
- AAUAAA
- AAAAAAA...A 3'
- poly-A tail
Splicing: An overview

Splicing removes introns from a primary transcript

- Splicing: exons are retained (and then translated), introns are removed
- alternative arrangements of exons can enable one gene to generate multiple products.
Amino acids

Amino acids to polypeptides

(a) Generic amino acid structure

(b) Amino acids with nonpolar R groups

- There are 20 essential amino acids (AAs), see figure 6.19 for the structure and classification of the 20
- each has a particular shape and charge, but all have the same generic structure that enables joining of AAs into chains of polypeptides
### Amino acids to polypeptides

(b) cont'd

<table>
<thead>
<tr>
<th>Amino acids with uncharged polar R groups</th>
<th>Amino acids with basic R groups</th>
<th>Amino acids with acidic R groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine (Gly) (G)</td>
<td>Lysine (Lys) (K)</td>
<td>Aspartic acid (Asp) (D)</td>
</tr>
<tr>
<td>Serine (Ser) (S)</td>
<td>Histidine (His) (H)</td>
<td>Glutamic acid (Glu) (E)</td>
</tr>
<tr>
<td>Threonine (Thr) (T)</td>
<td>Arginine (Arg) (R)</td>
<td></td>
</tr>
<tr>
<td>Cysteine (Cys) (C)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

more amino acids

© The McGraw-Hill Companies, Inc. Permission required for reproduction or display.
Peptide bonds and polypeptides

Amino acids to polypeptides

(c) Polypeptide formation

• the bond between amino acids is a peptide bond
• 3 or more AAs joined by peptide bonds is a polypeptide
• polypeptides assume particular shapes with specific distributions of charges
  • shape and charge of a polypeptide (or a set of mutually bound polypeptides) dictate its function
  • for enzymes, this means recognizing substrates, etc.
• polypeptides have polarity, one end is the n-terminus (free amino group), and the other is the c-terminus (free carboxylic acid group)
Nucleotide sequence is translated into AA sequence according to the genetic code.
triplets of nucleotides (codons) are interpreted by the translation mechanism into specific amino acids

<table>
<thead>
<tr>
<th>First letter</th>
<th>Second letter</th>
<th>Third letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>UU</td>
<td>UG</td>
</tr>
<tr>
<td></td>
<td>UU</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>UU</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>UU</td>
<td>G</td>
</tr>
<tr>
<td>C</td>
<td>CU</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>CU</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>CU</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>CU</td>
<td>G</td>
</tr>
<tr>
<td>A</td>
<td>AU</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>AU</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>AU</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>AU</td>
<td>C</td>
</tr>
<tr>
<td>G</td>
<td>GU</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>GU</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>GU</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>GU</td>
<td>C</td>
</tr>
</tbody>
</table>

- there are 64 possible orders of the four bases in sets of three
  - 61 code for 20 amino acids (some amino acids are encoded by more than one codon)
  - 3 codons are ‘stop’ codons which terminate translation
  - each codon specifies only one amino acid ***important***
- by convention, codons are written as they are seen in mRNA
Transfer RNA’s (tRNA’s) mediate the translation of mRNA codons to amino acid chains.

- tRNA’s are short, single stranded RNA molecules 74-95 nucleotides long.
- tRNA’s are ‘charged’ with one and only one of the twenty essential amino acids by a class of enzymes called aminoacyl-tRNA synthetases.
- Each aminoacyl-tRNA synthetases catalyzes the covalent bonding of one specific tRNA to its specific amino acid.
- Aminoacyl-tRNA synthetases are therefore the true molecular translators of nucleotide sequence into protein sequence.

- Secondary structure of tRNAs appears as a cloverleaf, in 3D, tRNA’s appear as a compact letter ‘L’.
  - At one end of the ‘L’ is a 3 base “anti-codon” that will base pair with a 3-base codon in the mRNA.
  - At the other end of the ‘L’ is the amino acid attachment site (at the 3’ end of the tRNA).
aminoacyl-tRNA synthetases

• tRNA’s are ‘charged’ with one and only one of the twenty essential amino acids by a class of enzymes called aminoacyl-tRNA synthetases.
  • each aminoacyl-tRNA synthetases catalyzes the covalent bonding of one specific tRNA to its specific amino acid.
  • aminoacyl-tRNA synthetases are therefore the true molecular translators of nucleotide sequence into protein sequence.
Ribosomes: complex factories comprised of 2 subunits comprised of proteins and rRNA

- rRNA is ribosomal RNA, transcribed like other RNA’s from DNA
in eukaryotes, the small subunit binds to the 5’ methylated cap of the mRNA and then migrates to the initiation site in a 5’ to 3’ direction

initiation codon is almost always AUG, which is recognized by an initiator tRNA carrying Methionine, so all new polypeptides start with methionine

the first nucleotides of a mRNA are NOT translated

first tRNA binds to the initiation site, and large ribosomal subunit joins to form complete ribosome
Translation: elongation

• the ribosome complex migrates down the mRNA in a 5’-3’ direction, adding the appropriate AA to the growing polypeptide chain, following the rules of base pairing between the anti-codon of the charged tRNA’s and the codons in the mRNA.
• the ribosomal complex releases the polypeptide and mRNA when a stop codon is reached.