**From Genes to Protein**

**Metabolism Teaches Us About Genes**
- **Metabolic defects**
  - studying metabolic diseases suggested that genes specified proteins
    - alkaptonuria (black urine from alkapton a.k.a. homogentisic acid)
    - PKU (phenylketonuria)
  - each disease is caused by non-functional enzyme

Genotypes create phenotype

**1 Gene – 1 Enzyme Hypothesis**
- **Beadle & Tatum**
  - Compared mutants of bread mold, *Neurospora* fungus
    - created mutations by X-ray treatments
      - X-rays “break” DNA
      - inactivate a gene
    - wild type grows on “minimal” media
      - sugars + required precursor nutrient to synthesize essential amino acids
    - mutants require added amino acids
      - each type of mutant lacks a certain enzyme needed to produce a certain amino acid
      - non-functional enzyme = broken gene

**Beadle & Tatum’s Neurospora Experiment**

**Beadle & Tatum**

**George Beadle**

**Edward Tatum**
So... What is a Gene?

- One gene = one enzyme
  - all genes code for enzymes
  - but there are proteins that are not enzymes coded
- One gene = one protein
  - each gene codes for a chain of amino acids
  - but many proteins are composed of several polypeptides (many chains of amino acids)
- One gene = one polypeptide
  - but many genes have the code for only RNA
- One gene = one product
  - but many genes can code for more than one product …

Defining a Gene...

“Defining a gene is problematic because... one gene can code for several protein products, some genes code only for RNA, two genes can overlap, and there are many other complications.”

---

The “Central Dogma”

- How do we move information from DNA to proteins?

DNA \rightarrow RNA \rightarrow protein

From nucleus to cytoplasm...

- Where are the genes?
  - genes are on chromosomes in nucleus
- Where are proteins synthesized?
  - proteins made in cytoplasm by ribosomes
- How does the information get from nucleus to cytoplasm?
  - messenger RNA

RNA

- ribose sugar
- N-bases
  - uracil instead of thymine
  - U : A
  - C : G
- single stranded
- mRNA, rRNA, tRNA, siRNA....

Transcription

- Transcribed DNA strand = template strand
- untranscribed DNA strand = coding strand
- Synthesis of complementary RNA strand
- transcription bubble
- Enzyme that facilitates the building of RNA:
  - RNA polymerase
Transcription in Prokaryotes

- **Initiation**
  - RNA polymerase binds to **promoter sequence** on DNA

- **Role of promoter**
  1. Where to start reading = starting point
  2. Which strand to read = template strand
  3. Direction on DNA = always reads DNA 3’→5’

- **Elongation**
  - RNA polymerase unwinds DNA ~20 base pairs at a time
  - reads DNA 3’→5’
  - builds RNA 5’→3’ (the energy governs the synthesis!)

- **Termination**
  - RNA polymerase stops at **termination sequence**
  - mRNA leaves nucleus through pores

**No proofreading**
- 1 error/10^6 bases
- many copies
- short life
- not worth it!

Transcription in Eukaryotes

- **3 RNA eukaryotic polymerase enzymes**
  - RNA polymerase I
    - only transcribes rRNA genes
  - **RNA polymerase II**
    - transcribes genes into mRNA
    - RNA polymerase III
      - only transcribes tRNA genes
    - each has a specific promoter sequence it recognizes
Transcription in Eukaryotes

- Initiation complex
  - transcription factors bind to **promoter region** upstream of gene
  - proteins which bind to DNA & turn on or off transcription
  - **TATA** box binding site
  - only then does RNA polymerase bind to DNA

Eukaryotic Post-transcriptional Processing

- **Primary transcript**
  - eukaryotic mRNA needs work after transcription

- **Protect mRNA**
  - from RNA-ase enzymes in cytoplasm
    - add 5’ G cap
    - add polyA tail

- **Edit out introns**

Prokaryote vs. Eukaryote Genetics

- **Prokaryotes**
  - DNA in cytoplasm
  - circular chromosome
  - naked DNA
  - no introns

- **Eukaryotes**
  - DNA in nucleus
  - linear chromosomes
  - DNA wound on histone proteins
  - introns vs. exons

From Gene to Protein

- **DNA**
  - **mRNA**
    - mRNA leaves nucleus through nuclear pores
  - proteins synthesized by ribosomes using instructions on mRNA

Prokaryote vs. Eukaryote Genetics

- Differences between prokaryotes & eukaryotes
  - time & physical separation between processes
  - RNA processing
Translation in Prokaryotes

- Transcription & translation are simultaneous in bacteria
  - DNA is in cytoplasm
  - no mRNA editing needed

How Does DNA Code for Proteins

- DNA: TACGCACATTACGTACGC GG
- mRNA: AUGCGUGUAAAUGCAUGCGCC
- Protein: Met Arg Val Asn Ala Cys Ala

How can you code for 20 amino acids with only 4 nucleotide bases (A,U,G,C)?

Cracking the Code

- Nirenberg & Matthaei
  - determined 1st codon–amino acid match
    - UUU coded for phenylalanine
  - created artificial poly(U) mRNA
  - added mRNA to test tube of ribosomes, tRNA & amino acids
    - mRNA synthesized single amino acid polypeptide chain

Translation

- Codons
  - blocks of 3 nucleotides decoded into the sequence of amino acids

mRNA Codes for Proteins in Triplets

- DNA: TACGCACATTACGTACGC GG
- mRNA: AUGCGUGUAAAUGCAUGCGCC
- Protein: Met Arg Val Asn Ala Cys Ala

mRNA codes for proteins in triplets... CODONS!
The Code!
- For ALL life!
  - strongest support for a common origin for all life
- Code is redundant
  - several codons for each amino acid
- Start codon
  - AUG
  - methionine
- Stop codons
  - UGA, UAA, UAG

How are Codons Matched to Amino Acids?
- DNA
  - TACGCACTTTACGTACGCGG
- mRNA
  - AUGCUUGUAAGUCAUGGCC
- tRNA
  - Met
  - Arg
  - Val

Amino acid attachment end

anti-codon

Amino acid

nucleus

cytoplasm

Loading tRNA
- Aminoacyl tRNA synthetase
  - enzyme which bonds amino acid to tRNA
  - endergonic reaction
    - ATP \rightarrow AMP
  - energy stored in tRNA-amino acid bond
    - unstable
  - so it can release amino acid at ribosome

Ribosomes
- Facilitate coupling of tRNA anticodon to mRNA codon
  - organelle or enzyme?
- Structure
  - ribosomal RNA (rRNA) & proteins
  - 2 subunits
    - large
    - small
**Ribosomes**

- **A site (aminoacyl-tRNA site)**
  - Holds tRNA carrying next amino acid to be added to chain
- **P site (peptidyl-tRNA site)**
  - Holds tRNA carrying growing polypeptide chain
- **E site (exit site)**
  - Empty tRNA leaves ribosome from exit site

**Building a Polypeptide**

- **Initiation**
  - Brings together mRNA, ribosome subunits, proteins & initiator tRNA
- **Elongation**
- **Termination**

**Elongation: Growing a Polypeptide**

1. Codon recognition
2. Peptide bond formation
3. Translocation

**Termination: Release Polypeptide**

- **Release factor**
  - "release protein" bonds to A site
  - Releases ribosome subunits, mRNA, and polypeptide

**Can you tell the eukaryotic story?**

- DNA
- RNA polymerase
- Pre-mRNA
- Mature mRNA
- 5' cap
- PolyA tail
- Large subunit
- Ribosome
- Small subunit
- Polypeptide
- 'Charged' tRNA

**Putting it all together...**

- The linear sequence of the gene is transcribed into RNA, which is then processed in the nucleus.
- The mature mRNA is transported to the cytoplasm, where it binds to the ribosome.
- The ribosome translates the mRNA into a polypeptide chain, following the genetic code.
- The polypeptide chain is then folded into its functional 3D structure, a process that involves the coordination of various enzymes and chaperones.
Mutations

- Code is redundant
  - several codons for each amino acid
  - "wobble" in the tRNA
  - "wobble" in the aminoacyl-tRNA synthetase enzyme that loads the tRNA

Universal Code

- Code is redundant
  - several codons for each amino acid
  - "wobble" in the tRNA
  - "wobble" in the aminoacyl-tRNA synthetase enzyme that loads the tRNA

Mutations

- Point mutations
  - single base change
  - base-pair substitution
    - silent mutation
      - no amino acid change
      - redundancy in code
    - missense
      - change amino acid
    - nonsense
      - change to stop codon

When do mutations affect the next generation?

A Mutation Leads to Sickle Cell Anemia

- What kind of mutation?

Wild-type hemoglobin DNA

Mutant hemoglobin DNA

Sickle Cell Anemia

- Normal red blood cells and the primary structure of normal hemoglobin
- Sickle red blood cells and the primary structure of sickle-cell hemoglobin
Mutations
- Frameshift
  - shift in the reading frame
  - changes everything “downstream”
- Insertions
  - adding base(s)
- Deletions
  - losing base(s)

Any Questions?