Chapter 10
The Structure and Function of DNA

Dec. 04, 2015
© Jong B. Lee, Ph.D.

DNA: STRUCTURE AND REPLICATION

- DNA was known to be a chemical in cells by the end of the nineteenth century
  - Has the capacity to store genetic information
  - Can be copied and passed from generation to generation

DNA Structure

- Nucleotides are joined by a sugar-phosphate backbone.

DNA and RNA Structure

- DNA and RNA are nucleic acids.

DNA와 RNA의 차이점
1. deoxyribose or ribose
2. thymine or uracil
3. double strand vs single strand
4. long vs short chain
5. DNA is much stable than RNA (the reason that DNA is adopted as a genetic material)
Watson and Crick’s Discovery of the Double Helix

- James Watson and Francis Crick determined that DNA is a double helix.

Watson and Crick used X-ray crystallography data to reveal the basic shape of DNA. (Rosalind Franklin collected the X-ray crystallography data)

- DNA 구조의 규칙성 (uniform diameter)
  - AA pair (purine pair) would be almost twice as wide as a CC pair (pyrimidine pairs)
  - A double-ringed base on one strand must always be paired with a single-ringed base on the opposite strand

Complementarity (상보성)

- A=T (G=C) or A=C (G=T)?
- Hint by Chargaff’s rule: the amount of adenine (A) in the DNA of any species was equal to the amount thymine (T) and that the amount of guanine (G) was equal to that of cytosine (C).
- 화학적 구조(수소결합 가능함)를 바탕으로 한 수소결합능력을 고려할 때 A=T (G=C) pair가 올바르다는 것을 알아냄

The model of DNA is like a rope ladder twisted into a spiral.
Anti-parallel: two sugar-phosphate backbones are oriented in opposite directions

DNA Replication

- Watson and Crick’s model for DNA structure suggested to them that DNA replicates by a template mechanism (주행): each DNA strand can serve as mold, or template, to guide reproduction of the other strand
- Both two strands of parental DNA are used as template
- New complementary strands are synthesized using the template strand in accordance with the base-pairing rules
- If you know the sequence of bases in one strand of the double helix, you can determine the sequence of bases in the other strand by applying the base-pairing rules

Template model for DNA replication
- The two strands of parental DNA separate
- Each becomes a template for the assembly of a complementary strand
- The nucleotides are lined up one at a time along the template strand in accordance with base-pairing rule
- Enzymes link the nucleotides to form the new DNA strands

DNA polymerases
- The enzymes that make the covalent bonds between the nucleotides of a new DNA strand
- The process is both fast (50 nt/sec) and amazingly accurate (1/10^9 error rate)
- Are also involved in repairing damaged DNA
• DNA replication
  – Begins at specific sites on a double helix
  – Multiple origins for eukaryotes
  – Proceeds in both directions (bidirectional synthesis: bubbles
    형성)
  – Semi-conservative replication (반보존적 복제)

• DNA can be damaged by ultraviolet light.
  – The enzymes and proteins involved in replication can repair the damage.

The Flow of Genetic Information from DNA to RNA to Protein

• An organism’s genotype and phenotype
  – The genotype is genetic makeup (= the sequence of nucleotide bases in DNA)
  – The phenotype is the organism’s specific traits, which arise from the actions of a wide variety of proteins.

• What is the connection between the genotype and phenotype? : A gene does not build a protein directly, but rather dispatches instructions in the form of RNA, which in turn programs protein synthesis

• DNA specifies the synthesis of proteins in two stages:
  – Transcription: the transfer of genetic information from DNA to an RNA molecule
  – Translation: the transfer of the information from RNA to a protein
Connection between genotype and phenotype

• Gene과 Protein의 관계는 1909년 A. Garrod에 의해 유전병 관찰을 통해서 제안되었다.
  예: alkaptonuria : alkapton이라는 chemical 때문에 소변에 검붉은 색을 뜨 괴로이 때문에 걸리는 유전병.
  - 유전병이 특정 효소를 만드는 능력이 없어졌을 때 생기는 것이라는 가설을 세웠다.
  - 이러한 병을 “inborn errors of metabolism (선천적 대사 질환)”라 언급했다.

→ 결론: 유전자는 화학적 과정을 촉매하는 효소 (단백질)를 통해 표현형으로 나타난다고 처음으로 제안

G. Beadle와 E. Tatum의 orange bread mold 연구

- There were mutants of this mold made by UV.
The mutant molds were not able to grow on the usual simple growth medium (단순배지).
- Each of the mutant turned out to lack an enzyme in a metabolic pathway that produced some molecule the mold needed, such as an amino acid: 배지에 아미노산 보충하면 자란
- Each mutant was defective in a single gene
- They hypothesized that the function of an individual gene is to dictate the production of a specific enzyme

→ The one gene–one enzyme hypothesis
The one gene–one protein hypothesis

최종결론: The one gene–one polypeptide hypothesis

From Nucleotide Sequence to Amino Acid Sequence:
An Overview

• When DNA is transcribed, the result is an RNA molecule (the same language of nucleotides).
The nucleotide bases of the RNA molecule are complementary to those on the DNA strand.
(RNA 분자는 DNA에 상보적이다 → 일대일 대응)
• RNA is then translated into a sequence of amino acids in a polypeptide
• Translation is the conversion of the nucleic acid language into the polypeptide language.
  → 4 대 20 (아미노산 종류) 대응을 해야한다. How?

• The sequence of nucleotides of the RNA molecule dictates the sequence of amino acids of the polypeptide
  (Nucleotides: 4 kinds, Amino acids: 20 kinds)

How to dictate?: 어떻게 4가지의 nucleotides 가 20가지의 아미노산을 지정할 수 있을까?
Solution: using combination (조합)
  최소조합: 3가지 조합 : 4 x 4 x 4 = 64 개의 조합

- Triplets of Bases → Specify all the amino acids

The three adjacent nucleotides or bases = triplet bases = codon
(따라서 두 개 이상의 codons이 하나의 amino acid을 중복해서 지정할 수 있음)
A typical gene consists of thousands of nucleotides, and a single DNA molecule may contain thousands of genes.

The genetic code is the set of rules relating nucleotide sequence to amino acid sequence.

- AUG: codes for the amino acid methionine (Met), a signal for the start of a polypeptide chain.
- UAA, UAG & UGA: stop codons.
- Redundancy in the codon but no ambiguity.
- Although codons UUU & UUC both specify Phe (redundancy), neither of them ever represents any other amino acid (no ambiguity).
- No gap, no punctuation separating codons.
- Reading frame: open reading frame (ORF).

The genetic code is shared by all organisms.

- The universality of genetic vocabulary suggests that the genetic code arose very early in evolution and was passed on over the long period to all the organisms.
- Because the code is the same in different species, genes can be transcribed and translated after transfer from one species to another, even when the organisms are as different as a bacterium and a human.
Transcription: From DNA to RNA

- In transcription
  - DNA must **first separate at the place** where the process will start
  - An RNA molecule is transcribed from a DNA template
  - Only one of the DNA strands serves as a template for the newly forming molecule
  - Synthesized by RNA polymerase

RNA Elongation

- The second phase of transcription is elongation
  - The RNA grows longer
  - Synthesis continues
  - DNA strands separate
  - The process repeats

Termination of Transcription

- When RNA polymerase reaches a sequence of DNA bases called a terminator, RNA polymerase detaches from the DNA.

Initiation of Transcription

- The “start transcribing” signal is a nucleotide sequence called a promoter
  - Located in the DNA at the beginning of the gene
  - A specific place where RNA polymerase attaches
  - RNA polymerase attaches to the promoter, RNA synthesis begins

The Processing of Eukaryotic RNA

- The eukaryotic cell processes the RNA after transcription whereas the prokaryotic cells undergo both transcription and translation simultaneously

  1. Adding a cap and tail
  2. Removing introns
  3. Splicing exons together

  장소: Nucleus
• RNA splicing
  – Noncoding stretches of nucleotides that interrupt the nucleotides that actually code for amino acids are eliminated by this process
  – intron: internal noncoding regions, exon: coding regions
  – Primary transcript: exons and introns are transcribed from DNA into RNA.
  – Before the RNA leaves the nucleus, the introns are removed, and the exons are joined to produce an RNA molecule with continuous coding sequence. This process is called RNA splicing.

– RNA splicing play a significant role in humans in allowing 25,000 genes to produce many thousand more polypeptides (by alternative splicing by which selective exons are joined in a final RNA molecule)

Translation: The Players

• Translation
  – Is the conversion from the nucleic acid language to the protein language.
  – Three players are required
    1. mRNA the first ingredient for translation
    2. tRNA,
    3. ribosome

Transfer RNA (tRNA)

• Acts as a molecular interpreter (분자해독기)
• Two components
  1. Amino acid attachment site: carries amino acids using
  2. Anticodon: attaches amino acids with codons in mRNA
**Transfer RNA (tRNA)**

- Is made of a single strand of RNA (about 80 nucleotides). About 50 tRNAs in the cell.
- The chain twists and folds upon itself to form “L” shape as a 3-dimensional structure.
- **Anticodon**, a special triplet of bases, is complementary to a codon on mRNA.
- During translation, the anticodon on the tRNA recognizes a particular codon on the mRNA by base-pairing rules.
- At the other end of tRNA is a site where an amino acid can attach.

**Ribosomes**

- Are the organelles that coordinate the functioning of the mRNA and tRNA and actually make polypeptides.
- Are made up of two subunits (Large and small subunits).
- Each subunit contains a few ribosomal RNAs (rRNAs) and a few dozen ribosomal proteins (r-proteins).
- A fully assembled ribosome has a binding site for mRNA on its small subunit and binding sites for tRNAs on its large subunit.

---

**Translation: The Process**

- Translation is divided into three phases:
  - Initiation
  - Elongation
  - Termination
**Initiation**

- The first phase brings together:
  - The mRNA
  - The first amino acid with its attached tRNA
  - The two subunits of the ribosome
- An mRNA molecule has a cap and tail that help it bind to the ribosome

**Elongation**

- **Step 1: codon recognition**
  - The anticodon of an incoming tRNA pairs with the mRNA codon
- **Step 2: peptide bond formation**
  - The ribosome catalyzes bond formation between amino acids:
    - The peptidyl transferase (RNA enzyme = ribozyme)
- **Step 3: translocation**
  - A tRNA leaves the P site of the ribosome.
  - The ribosome moves down the mRNA, the remaining tRNA (which carries the growing polypeptide) to the P site.

**Termination**: Elongation continues until the ribosome reaches a stop codon
• In eukaryotic cells
  – Transcription occurs in the nucleus
  – Translation occurs in the cytoplasm

• In prokaryotic cells
  – Transcription and translation occur simultaneously

Types of Mutations

• Mutations within a gene
  – Can be divided into two general categories.
  – Can result in changes in the amino acids in proteins.
• **Base substitution**

Based on how a base substitution is translated, it can result in no change in the protein, in an insignificant change, or in a change that might be crucial to the life of the organism

- **Silent mutation**: no change in protein $\rightarrow$ due to redundancy of the genetic code

- **Missense mutations**: some have little or no effect on protein function and some cause changes in protein function (hemoglobin gene in sickle cell)

- **Nonsense mutation**: amino acid codon $\rightarrow$ stop codon; the result will be prematurely terminated, which will not function properly

• **Insertion and Deletion**

- Can have disastrous effects

- Because mRNA is read as a series of nucleotide triplets during translation, adding or subtracting nucleotides may alter the reading frame of the genetic message

- All the nucleotides that are “downstream” of the insertion or deletion will be grouped into different codons

- The altered polypeptide is likely to be nonfunctional

---

**Mutagens**

• Mutations may result from **Errors in DNA replication** or recombination (spontaneous mutation)

• **Physical mutagens**: high energy radiation such as X-rays and UV light $\rightarrow$ mutations의 10%를 차지함

• **Chemical agents**: chemicals that are similar to normal DNA bases but that base-pair incorrectly when incorporated into DNA (agents that cause cancer = **carcinogens**)

• **Oxidative damage**: the most important source of mutagenic alterations in DNA $\rightarrow$ hydrogen peroxide ($H_2O_2$), hydroxyl radicals (OH$^\cdot$), superoxide radicals ($O_2^\cdot$) from irradiation or byproduct of aerobic respiration $\rightarrow$ cells have defense systems: catalase, superoxide dismutase (SOD)

• **Mutagens**

  - Although mutations are often harmful,

  - They are the source of the rich diversity of genes in the living world.

  - They contribute to the process of evolution by natural selection

  - Essential tools for geneticists
Viruses: Genes in Packages

- Viruses sit on the fence between life and nonlife.
  - A virus is lifelike in having genes and a highly organized structure.
  - But it differs from a living organism in not having cellular structure or able to reproduce on its own.
  - Nothing more than packaged genes (genes in a box), a bit of nucleic acid wrapped in a protein coat.
  - Can survive only by infecting a living cell with genetic material that directs the cell’s molecular machinery to make more viruses.

Bacteriophages

- Bacteriophages, or phages, attack bacteria.
- Phages have two reproductive cycles: lytic and lysogenic cycles.

Lysogenic cycles of lambda (λ) phage can switch to lytic cycles.
- Occasionally a phage leaves its chromosome; this event may be triggered by environmental conditions such as exposure to a mutagen (UV).
- Once separated, the lambda DNA usually switches to the lytic cycle, which results in the production of many copies of the virus and bursting of the host cell.
- Sometimes the few prophage genes active in a lysogenic bacterial cell can cause medical problems (diphtheria, botulism, and scarlet fever). Certain genes direct the bacteria to produce the toxins responsible for making people ill.
Plant Viruses

- Viruses that infect plants
  - Can stunt growth and diminish plant yields.
  - Can spread throughout the entire plant.

**TMV (tomato mosaic virus):**
- Rod-shape with a spiral arrangement of proteins (2130 identical coat proteins) surrounding the nucleic acid (RNA).

Genetic engineering methods

- Have been used to create virus-resistant plants
- Papaya is Hawaii’s second largest crop.
- The spread of papaya ringspot potyvirus (RPSV) by aphids had wiped out the papaya in certain island regions.
- Since 1998, farmers have been able to plant a newly engineered PRSV-resistant strain of papaya, and papaya are now being reintroduced into their old habitats.

Animal Viruses

- Virus that infect animal cells are common causes of disease.

- Influenza virus (flu) – enveloped virus
  - Outer envelope made of phospholipid membrane, with projecting spikes of protein

- Other RNA viruses: common cold, measles, mumps, polio, AIDS 유발
Life cycle of enveloped virus (flu virus)

- On viral entry to cell, spike of the virus binds to receptor on cell surface.
- A envelope protein plays a role in fusion with the cell’s membrane, allowing the protein-coated RNA to enter the cytoplasm.
- Enzymes then remove the protein coat.
- Finally, the viruses leave the cell by cloaking themselves in plasma membrane.

Not all animal viruses reproduce in the cytoplasm

- Herpes viruses are enveloped DNA viruses that reproduce in a cell’s nucleus and they get their envelopes from the cell’s nuclear membrane.
- Copies of the herpes virus DNA usually remain behind as mini-chromosomes in the nuclei of certain nerve cells.
- There they remain latent until some sort of physical stress, such as a cold or sunburn, or emotional stress triggers the herpes virus DNA to begin producing the virus, resulting in unpleasant symptoms.
- Once acquired, herpes infections may flare up repeatedly throughout a person’s life: Over 75% of American adults carry herpes simplex 1 (causes cold sores) and over 20% carry simplex 2 (causes genital herpes).

HIV, the AIDS Virus

- HIV is a retrovirus.
- A retrovirus is an RNA virus that reproduces by means of a DNA molecule.
- It copies its RNA to DNA using reverse transcriptase (RT).
- RT is a DNA polymerase with high error rate (= no proofreading) (1/10⁹ vs 1/10⁴)
- Mutation rate: extremely high → rapid evolution.

How HIV reproduces inside a cell
AIDS (Acquired immune deficiency syndrome)
- The disease caused by HIV infection.
- Kill white blood cells
- Loss of such cells causes the body to become susceptible to other infections that it would normally be able to fight off
- Secondary infections cause the syndrome eventually kills AIDS patients
- Treated with the drug AZT