

生物科學系

陳錦翠

A

JIONAL SUN YAT-SEN UNIVERSITY



THE MOLECULE BASIS OF INHERITANCE



1953, James Watson and Francis Crick 1962 Nobel prize

The search for genetic material lead to DNA

- Once T.H. Morgan's group showed that genes are located on chromosomes, the two constituents of chromosomes proteins and DNA - were the candidates for the genetic material.
- Until the **1940s**, the great heterogeneity and specificity of function of proteins seemed to indicate that proteins were the genetic material.
- However, this was not consistent with experiments with microorganisms, like bacteria and viruses.



- The discovery of the genetic role of DNA began with research by Frederick Griffith in 1928.
- Griffith called this phenomenon transformation(轉型), a change in genotype and phenotype due to the assimilation(同化) of a foreign substance (now known to be DNA) by a cell.
- For the next 14 years scientists tried to identify the transforming substance.
- Finally in 1944, Oswald Avery, Maclyn McCarty and Colin MacLeod announced that the transforming substance was DNA

 In 1952, Alfred Hershey and Martha Chase showed that DNA was the genetic material of the bacteriophage T2.





Department of Genetics





Hershey and Chase's experiment



- By 1947, Erwin Chargaff had developed a series of rules based on a survey of DNA composition in organisms.
 - He already knew that DNA was a polymer of nucleotides consisting of a nitrogenous base, deoxyribose, and a phosphate group.
 - The bases could be adenine (A), thymine (T), guanine (G), or cytosine (C).
- Chargaff noted that the DNA composition varies from species to species.
- In any one species, the four bases are found in characteristic, but not necessarily equal, ratios.

- He also found a peculiar regularity in the ratios of nucleotide bases which are known as Chargaff's rules.
- The number of adenines was approximately equal to the number of thymines (%T = %A).
- The number of guanines was approximately equal to the number of cytosines (%G = %C).
 - Human DNA is 30.9% adenine, 29.4% thymine, 19.9% guanine and 19.8% cytosine.

Watson and Crick discovered the double helix by building models to conform to X-ray data

- By the beginnings of the 1950's, the race was on to move from the structure of a single DNA strand to the threedimensional structure of DNA.
 - Among the scientists working on the problem were
 Linus Pauling, in California, Caltech
 Maurice Wilkins and Rosalind Franklin, in London.

- The phosphate(磷酸) group of one nucleotide(核苷酸) is attached to the sugar of the next nucleotide in line.
- The result is a "backbone" of alternating phosphates and sugars, from which the bases project.



- Maurice Wilkins and Rosalind Franklin used X-ray crystallography (X射線晶體學) to study the structure of DNA.
 - In this technique, X-rays are diffracted as they passed through aligned fibers of purified DNA.
- The diffraction(繞射) pattern can be used to deduce the three-dimensional shape of molecules.
- James Watson learned from their research that DNA was helical in shape and he deduced the width of the helix and the spacing of bases.



Rosalind Franklin died in the age of 38

- Watson and his colleague Francis Crick began to work on a model of DNA with two strands, the double helix (雙螺旋).
- Using molecular models made of wire, they first tried to place the sugar-phosphate chains on the inside.
- However, this did not fit the X-ray measurements and other information on the chemistry of DNA.
- The key breakthrough came when Watson put the sugarphosphate chain on the outside and the nitrogen bases on the inside of the double helix.
 - The sugar-phosphate chains of each strand are like the side ropes of a rope ladder.
 - Pairs of nitrogen bases, one from each strand, form rungs.
 - The ladder forms a twist every ten bases.





(c) Space-filling model

Watson built a model in which the backbones were **antiparallel** 反向平行 (their subunits run in opposite directions)

- The nitrogenous bases are paired in specific combinations: adenine with thymine and guanine with cytosine.
- Pairing like nucleotides did not fit the uniform diameter indicated by the X-ray data.
 - A purine-purine pair would be too wide and a pyrimidinepyrimidine pairing would be too short.
 - Only a pyrimidinepurine pairing would produce the 2-nm diameter indicated by the X-ray data.

Pyrimidine + pyrimidine: too narrow

Purine + pyrimidine: width consistent with X-ray data

- In addition, Watson and Crick determined that chemical side groups of the nitrogen bases would form hydrogen bonds, connecting the two strands.
 - Based on details of their structure, adenine would form two hydrogen bonds only with thymine and guanine would form three hydrogen bonds only with cytosine.
 - This finding explained Chargaff's rules.



- The base-pairing rules dictate the combinations of nitrogenous bases that form the "rungs" of DNA.
- However, this does not restrict the sequence of nucleotides along each DNA strand.
- The linear sequence of the four bases can be varied in countless ways.
- Each gene has a unique order of nitrogen bases.
- In April 1953, Watson and Crick published a succinct, onepage paper in *Nature* reporting their double helix model of DNA.

DNA Replication

During DNA replication, base pairing enables existing DNA strands to serve as templates (模版) for new complimentary(互補) strands





(a) The parent molecule has two complementary strands of DNA. Each base is paired by hydrogen bonding with its specific partner, A with T and G with C.

(b) The first step in replication is separation of the two DNA strands.

A T A T C G C G T A T A A T A T G C G

(c)Each parental strand now serves as a template that determines the order of nucleotides along a new complementary strand.



(d) The nucleotides are connected to form the sugar-phosphate backbones of the new strands. Each "daughter" DNA molecule consists of one parental strand and one new strand.

- In a second paper Watson and Crick published their hypothesis for how DNA replicates
- Watson and Crick's model, **semiconservative replication**



半保留,半保守

 Experiments in the late 1950s by Matthew Meselson and Franklin Stahl supported the semiconservative model, proposed by Watson and Crick, over the other two models.



- In eukaryotes, there may be hundreds or thousands of origin sites (起始點) per chromosome.
 - At the origin sites, the DNA strands separate forming a replication "bubble" with replication forks at each end. 複製叉

0.25 µm







- DNA polymerases聚合酶 catalyze the elongation of new DNA at a replication fork.
- As nucleotides align with complementary bases along the template strand, they are added to the growing end of the new strand by the polymerase.
 - The rate of elongation: is about **500 nucleotides** per second in **bacteria** and 50 per second in human cells. The raw nucleotides are nucleoside triphosphates.
- The raw nucleotides are

nucleoside triphosphates.核苷三磷酸

Copyright © 2002 Pearson Education, Inc., publishing as Benjamin Cummings



• The strands in the double helix are antiparallel.



- To summarize, at the replication fork, the leading strand is copied continuously into the fork from a single primer.
- The lagging strand (先導股) is copied away from the fork in short segments, each requiring a new primer.



A summary of bacterial DNA replication (複製).



© 2011 Pearson Education, Inc.

100-200 nucleotides

- DNA polymerases cannot initiate synthesis of a polynucleotide because they can only add nucleotides to the end of an existing chain that is base-paired with the template strand.
- To start a new chain requires a primer, a short segment of RNA.
 - The primer is about 10 nucleotides long in eukaryotes.
- Primase(引子酶), an RNA polymerase, links ribonucleotides that are complementary to the DNA template into the primer.
 - Primase starts an RNA chain from a single template strand.

- DNA polymerase III then adds deoxyribonucleotides (去氧核糖核苷酸) to the 3' end of the ribonucleotide chain.
- Another DNA polymerase I later replaces the primer ribonucleotides with deoxyribonucleotides complimentary to the template.
- After the primer is formed, DNA polymerase can add new nucleotides away from the fork until it runs into the previous Okazaki fragment.
- The primers are converted to DNA before DNA ligase (連接酶) joins the fragments together.

- Returning to the original problem at the replication fork, the leading strand requires the formation of only a single primer as the replication fork continues to separate.
- The lagging strand requires formation of a new primer as the replication fork progresses.
- In addition to primase, DNA polymerases, and DNA ligases, several other proteins have prominent roles in DNA synthesis.
- A helicase (解旋酶) untwists and separates the template DNA strands at the replication fork.

 Single-strand binding proteins keep the unpaired template strands apart during replication and protect the template from degradation.



A current model of the DNA replication complex



© 2011 Pearson Education, Inc.

- It is conventional and convenient to think of the DNA polymerase molecules moving along a stationary DNA template.
- In reality, the various proteins involved in DNA replication form a single large complex that may be anchored to the nuclear matrix (核基質).
- The DNA polymerase molecules "reel in"(捲入) the parental DNA and "extrude" newly made daughter DNA molecules.

Enzymes proofread DNA during its replication and repair damage in existing DNA

- Mistakes during the initial pairing of template nucleotides and complementary nucleotides occurs at a rate of one error per 10,000 base pairs.
- DNA polymerase proofreads (校正) each new nucleotide against the template nucleotide as soon as it is added.
- If there is an incorrect pairing, the enzyme removes the wrong nucleotide and then resumes synthesis.
- The final error rate is only one per billion (1/10⁹) nucleotides.



Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.
- The ends of eukaryotic chromosomal DNA molecules, the telomeres端粒, have special nucleotide sequences.
 - In human telomeres, this sequence is typically TTAGGG, repeated between 100 and 1,000 times.
- Telomeres protect genes from being eroded through multiple rounds of DNA replication.



Copyright © 2002 Pearson Education, Inc., publishing as Benjamin Cummings

Fig. 16.19a

- Eukaryotic cells have evolved a mechanism to restore shortened telomeres.
- Telomerase端粒酶uses a short molecule of RNA as a template to extend the 3' end of the telomere.
 - There is now room for primase and DNA polymerase to extend the 5' end.
 - It does not repair the 3'-end "overhang," but it does lengthen the telomere.



- Telomerase is *not* present in most cells of multicellular organisms.
- Therefore, the DNA of **dividing somatic cells** and cultured cells does tend to become shorter.
- Thus, telomere length may be a limiting factor in the life span of certain tissues and the organism.
- Telomerase is present in germ-line cells 生殖細胞, ensuring that zygotes have long telomeres.
- Active telomerase is also found in cancerous somatic cells 體細胞.
 - This overcomes the progressive shortening that would eventually lead to self-destruction of the cancer.

Chromatin Packing in a Eukaryotic Chromosome



© 2011 Pearson Education, Inc.

 Histones 組蛋白can undergo chemical modifications that result in changes in chromatin organization





(b) Acetylation (乙醯化) of histone tails promotes loose chromatin 染色質 structure that permits transcription 轉錄作用



- Chromatin undergoes changes in packing during the cell cycle
- At interphase (間期), some chromatin is organized into a 10-nm fiber, but much is compacted into a 30-nm fiber, through folding and looping
- Though interphase chromosomes are not highly condensed, they still occupy specific restricted regions in the nucleus





Painting Chromosomes



- Most chromatin is loosely packed in the nucleus during interphase and condenses prior to mitosis
- Loosely packed chromatin is called euchromatin (真 染色質)
- During interphase a few regions of chromatin, such as centromeres (中節) and telomeres(端粒) are highly condensed into heterochromatin (異染色質)
- Dense packing of the heterochromatin makes it difficult for the cell to express genetic information coded in these regions

The flow of genetic information.



(b) Eukaryotic cell



Reading frame 讀框



61 of 64 triplets code 三聯密碼 for amino acids Transcription can be separated into three stages: -Initiation起始 -Elongation延長 -Termination終止



Complexicity of the transcriptome (轉錄體) (The entire set of transcripts produced in a cell)



Highlight 15-1 Figure 1 *Molecular Biology: Principles and Practice* © 2012 W. H. Freeman and Company



Unnumbered 15 p515 Molecular Biology: Principles and Practice © 2012 W. H. Freeman and Company Robert Tjian (錢澤南) University of California, Berkeley

Howard Hughes Medical Institute

Dynan, W.S.

Identified the first transcription specificity protein, Sp1 Transcription factor 轉錄因子 (1983) Cell 32: 669-680. Cell 35: 79-87.

Transcription: the enzymatic production of RNA that complements to its DNA template



Molecular Biology: Principles and Practice © 2012 W. H. Freeman and Company

The RNA polymerase channels



General scheme of transcription

The transcription bubble

E coli RNA polymerase



Figure 15-6 *Molecular Biology: Principles and Practice* © 2012 W. H. Freeman and Company

Prokaryotic RNA Polymerase

RNA polymerase of *E. Coli* is a multisubunit protein

Number	Role
2	Control the frequency of polymerase
	Initiate transcription from specific promoter
1	Forms phosphodiester bond
1	Binds DNA template
1	Recognizes promoter and facilitates initiation (specificity)
	2



(a) Lactose absent, repressor active, operon off

The *lac* operon in *E. coli*: regulated synthesis of inducible enzymes



(b) Lactose present, repressor inactive, operon on



(a) Tryptophan absent, repressor inactive, operon on



(b) Tryptophan present, repressor active, operon off

Positive control of the *lac* operon by catabolite activator protein (CAP)



 The diverse functions of RNA range from structural to informational, to catalytic.

Table 17.1 Types of RNA in a Eukaryotic Cell			
Type of RNA	Functions		
Messenger RNA (mRNA)	Carries information specifying amino acid sequences of proteins from DNA to ribosomes.		
Transfer RNA (tRNA)	Transportation of amino acid into ribosome for protein synthesis		
Ribosomal RNA (rRNA)	Plays structural and catalytic (ribozyme) roles in ribosomes.		
Primary transcript	Serves as a precursor to mRNA, rRNA, or tRNA and may be processed by splicing or cleavage . In eukaryotes, pre-mRNA commonly contains introns, noncoding segments that are spliced out as the primary transcript is processed. Some intron RNA acts as a ribozyme, catalyzing its own splicing.		
Small nuclear RNA (snRNA)	Plays structural and catalytic roles in spliceosomes, the complexes of pro- tein and RNA that splice pre-mRNA in the eukaryotic nucleus.		
SRP RNA	Is a component of the signal- recognition particle (SRP), the protein-RNA complex that recognizes the signal peptides of polypeptides targeted to the ER.		

Classes of Eukaryotic RNA

(1) Ribosomal RNA (rRNA)
18 S (small subunit), 28 S (large subunit)
5.8 S (large subunit), 5 S (large subunit)

- (2) Transfer RNA (tRNA)
- (3) Messenger RNA (mRNA), microRNA (miRNA)
- (4) Heterogeneous nuclear RNA (hnRNA) (precursor of mRNA)
- (5) Small nuclear RNA (snRNA) <u>U1, U2, U3, U4, U5, U6, U7, U8, U9, U10</u>
- (6) Small cytoplasmic RNA (scRNA)7SL RNA (signal recognition particle RNA in protein synthesis)

The Human RNA Polymerases

Polymerase	Location	Product
pol I	nucleolus	18S, 28S,5.8S, rRNA
pol II	nucleoplasm	hnRNA/mRNA, miRNA U1,U2, U4, U5 snRNAs
pol III	nucleoplasm	tRNA, 5S rRNA U6 snRNA, 7SL RNA
Mitochondrial RNA pol	Mitochondria	All mitochondria RNA

An eukaryotic gene and its transcript.



UTR: untranslated region



Nuclear Architecture and Gene Expression

- Loops of chromatin extend from individual chromosomes into specific sites in the nucleus
- Loops from different chromosomes may congregate at particular sites, some of which are rich in transcription factors and RNA polymerases
- These may be areas specialized for a common function

Chromosomal interactions in the interphase nucleus.







RNA splicing is accomplished by a spliceosome.

spliceosomes consist of a variety of proteins and several small nuclear ribonucleoproteins (snRNPs).



- RNA splicing appears to have several functions.
 - First, at least some introns contain sequences that control gene activity in some way.
 - Splicing itself may regulate the passage of mRNA from the nucleus to the cytoplasm.
 - One clear benefit of split genes is to enable a one gene to encode for more than one polypeptide.
- Alternative RNA splicing gives rise to two or more different polypeptides, depending on which segments are treated as exons.
 - Early results of the Human Genome Project indicate that this phenomenon may be common in humans.

- Split genes may also facilitate the evolution of new proteins.
- Proteins often have a modular architecture with discrete structural and functional regions called domains.
- In many cases, different exons code for different domains of a protein.



Polypeptide







Structure of tRNA

•base pairing between the third base of the codon and anticodon are relaxed (called **wobble pairing**).
Atomic resolution structure of the bacterial ribosome



Tome steitz Ada Yonath Venki Ramakrishnan

2009 Nobel prize

- Each ribosome has a binding site for mRNA and three binding sites for tRNA molecules.
 - The P site holds the tRNA carrying the growing polypeptide chain.
 - The A site carries the tRNA with the next amino acid.
 - Discharged tRNAs leave the ribosome at the E site.



(a) Computer model of functioning ribosome



(b) Schematic model showing binding sites



(c) Schematic model with mRNA and tRNA

- Initiation brings together mRNA, a tRNA with the first amino acid, and the two ribosomal subunits.
 - First, a small ribosomal subunit binds with mRNA and a special **initiator tRNA**, which carries methionine and attaches to the start codon.
 - Initiation factors bring in the large subunit such that the initiator tRNA occupies the P site.



The three steps of elongation continue codon by codon to add amino acids until the polypeptide chain is completed.



- **Termination** occurs when one of the three stop codons reaches the A site.
- A release factor binds to the stop codon and hydrolyzes the bond between the polypeptide and its tRNA in the P site.
- This frees the polypeptide and the translation complex disassembles.



- Typically a single mRNA is used to make many copies of a polypeptide simultaneously.
- Multiple ribosomes, polyribosomes, may trail along the same mRNA.
- A ribosome requires less than a minute to translate an average-sized mRNA into a polypeptide.





- (a) An mRNA molecule is generally translated simultaneously by several ribosomes in clusters called polyribosomes.
- (b) This micrograph shows a large polyribosome in a prokaryotic cell (TEM).

A summary of genetic flow of an eukaryotic cell.



https://www.youtube.com/w atch?v=7Hk9jct2ozY2021 陳錦翠老師-分生.ppt

Noncoding RNAs play multiple roles in controlling gene expression

- Only a small fraction of DNA codes for proteins, and a very small fraction of the non-proteincoding DNA consists of genes for RNA such as rRNA and tRNA
- A significant amount of the genome may be transcribed into noncoding RNAs (ncRNAs), such as microRNA (miRNA)
- Noncoding RNAs regulate gene expression at two points: mRNA translation and chromatin configuration

- The phenomenon of inhibition of gene expression by RNA molecules is called **RNA interference (RNAi)**
- RNAi is caused by small interfering RNAs (siRNAs)
- siRNAs and miRNAs are similar but form from different RNA precursors

Micro RNAs and Small Interfering RNAs

- MicroRNAs (miRNAs) are small single-stranded RNA molecules that can bind to mRNA
- These can degrade mRNA or block its translation
- It is estimated that expression of at least half of all human genes may be regulated by miRNAs



- Small interfering RNAs (siRNAs) are similar to miRNAs in size and function
- The blocking of gene expression by siRNAs is called RNA interference (RNAi)
- RNAi is used in the laboratory as a means of disabling genes to investigate their function

Chromatin Remodeling by ncRNAs

- Some ncRNAs act to bring about remodeling of chromatin structure
- In some yeasts siRNAs re-form heterochromatin at centromeres after chromosome replication

Condensation of chromatin at the centromere (siRNA-protein complexes)

RNA transcripts (red) produced.

- 2 Yeast enzyme synthesizes strands complementary to RNA transcripts.
- **3** Double-stranded RNA processed into siRNAs that associate with proteins.
- 4 The siRNA-protein complexes bind RNA transcripts and become tethered to centromere region.



Condensation of chromatin at the centromere



- Small ncRNAs called piwi-associated RNAs (piRNAs) induce heterochromatin, blocking the expression of parasitic DNA elements in the genome, known as transposons
- RNA-based regulation of chromatin structure is likely to play an important role in gene regulation

The Evolutionary Significance of Small ncRNAs

- Small ncRNAs can regulate gene expression at multiple steps
- An increase in the number of miRNAs in a species may have allowed morphological complexity to increase over evolutionary time
- siRNAs may have evolved first, followed by miRNAs and later piRNAs

A program of differential gene expression leads to the different cell types in a multicellular organism

- During embryonic development, a fertilized egg gives rise to many different cell types
- Cell types are organized successively into tissues, organs, organ systems, and the whole organism
- Gene expression orchestrates the developmental programs of animals

Reprogramming of transcription (Oct4, Sox2, Klf4, and c-Myc) iPS cells



Highlight 15-2 Figure 1 Molecular Biology: Principles and Practice © 2012 W. H. Freeman and Company

Induced pluripotent stem 2012 Noble prize

Shinya Yamanaka Kyoto University, Japan

John B. Gurdon Gurdon Institute, Cambridge, United Kingdom

山中伸彌





