VARIATIONS OF DNA

Ho Huynh Thuy Duong

University of Science

What is the meaning of DNA variations in living world ?

Negative effects	Positive effects
☙ At the scale of the living world, high DNA variations would prevent the maintenance of species ☙ For a living organism, they usually have lethal effects	☙ Without these variations, the Evolution of Life on Earth would not be possible, new species could not be created and the existing biodiversity would not be observed

Thus, challenging problem for the living world is how to maintain a relative stability of genetic material while allowing variations at a certain level.

This is actually realized by repair systems which, however, do not perfectly reverse all variations on DNA

VARIATIONS OF DNA INCLUDE :

- **DNA MUTATIONS**
- **DNA RECOMBINATION – HOMOLOGOUS AND SITE-SPECIFIC RECOMBINATION**
- **DNA TRANSPOSITION**

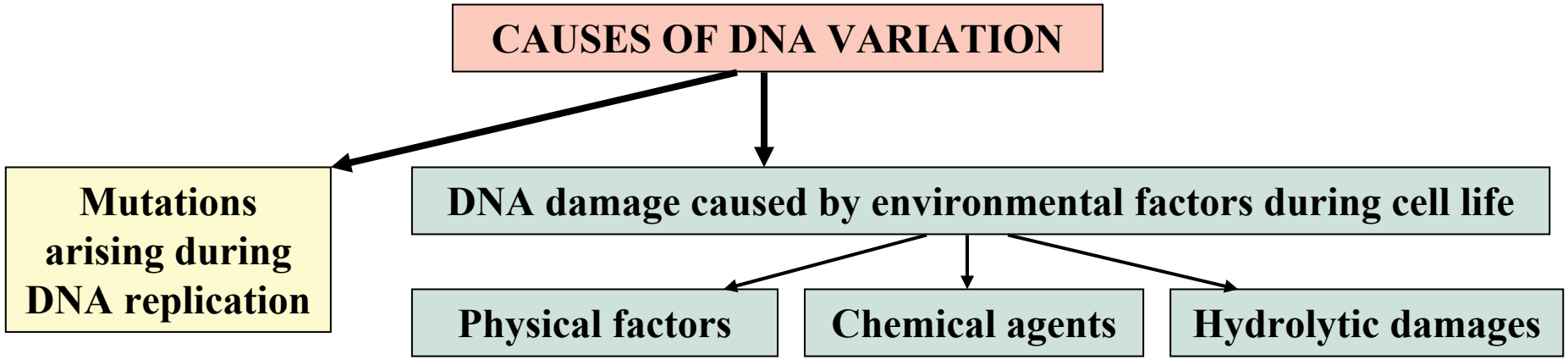
DNA MUTATIONS

WHAT KIND OF MUTATIONS DOES DNA UNDERGO DURING ITS EXISTENCE?

1. **Point mutations** : insertion or deletion of one base and base substitutions including :
   - Transitions : pyrimidine → pyrimidine, e.g T→C ; purine → purine, e.g A→G
   - Transversions : pyrimidine → purine, e.g T →G/A ; purine → pyrimidine, e.g A→C/T

2. **Insertions, deletions** of large DNA fragments and **rearrangements** of the chromosomes

What are the causes of DNA mutations ?



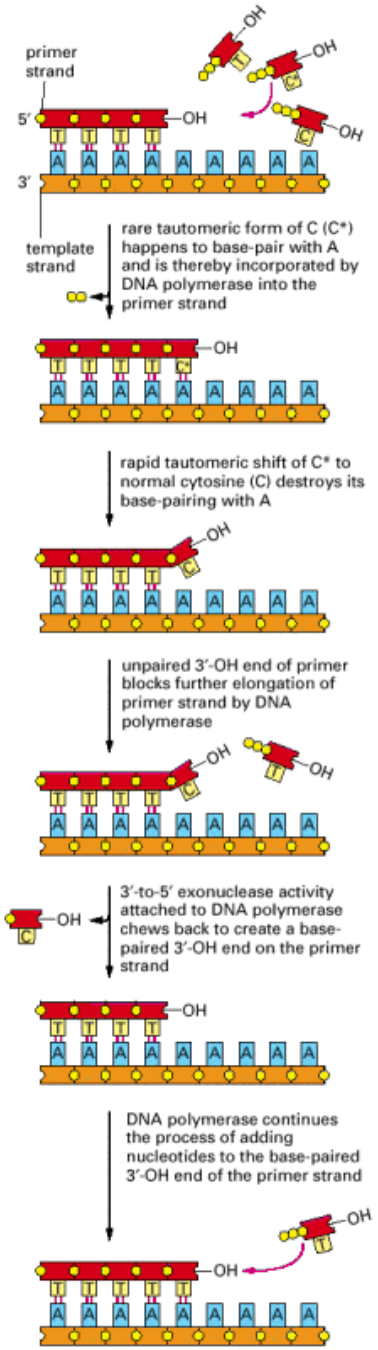
What are the consequences of DNA variations ?

1. Blocking DNA replication and transcription → cell death
2. Altering bases → mispairing → mutations → cell mutated

HOW DOES CELL MAINTAIN ITS DNA INTEGRITY ?

MAINTENANCE OF DNA INTEGRITY IS ASSURED BY REPAIR SYSTEMS WHICH OPERATE DURING DNA REPLICATION OR BETWEEN TWO REPLICATION ROUNDS

REPAIR OF REPLICATION ERRORS



During replication, misincorporation of nucleotides can occur.

These errors can be repaired in two ways :

- **Many DNA polymerases have proofreading function which uses 3'-5' exonuclease activity to eliminate the wrongly incorporated nucleotide during replication process.**

- **Some of the misincorporated nucleotides which escape proofreading could be recognized and eliminated just after finishing replication by the mismatch repair system.**

The mismatch repair system has two functions :

1. Rapidly detects the mismatches before they being fixed at the next round of replication

2. Accurately corrects the mismatches, that is always replacing the mismatched nucleotide in the daughter strand, never in the parental strand

DNA polymerase proofreading : a misincorporated nucleotide blocks the replication ; DNA polymerase 3'-5' exonuclease activity hydrolyzes the unmatched nucleotideand allows synthesis process to continue.

MISMATCH REPAIR SYSTEM

In *E. coli*, the mismatch repair system is composed of many components - MutS, MutL, MutH : (1) MutS detects the mismatch through distortion it causes to the DNA structure and binds to the region containing the mismatch, (2) MutL activates MutH, (3) activated MutH causes a nick on one strand near the region containing the mismatch, (4) A helicase, UvrD, unwinds the DNA toward the site of the mismatch, (5) An exonuclease digests the displaced single strand including the site of the mismatch, (6) DNA polymerase III fills in the single-stranded gap, (7) A ligase seals the nick.

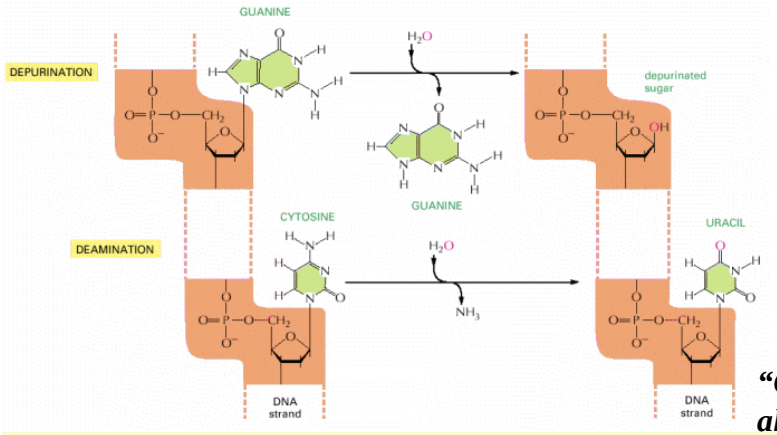
How does the mismatch repair system work accurately through distinguishing the parental from the daughter strand ?

All the A residues belonging to the sequence 5' GATC 3' throughout *E. coli* genome are methylated. Just after the replication fork passes, each daughter strand is hemi-methylated, that is they are only methylated on the initial parental strand. Thus, before Dam methylase methylates the newly synthesized strand, there is a moment when the repair system can recognize the newly synthesized strand and corrects the mismatch on it.

Eukaryotes have a homolog system to the *E. coli* mismatch repair system composed of MSH (MutS homologs) and MLH (MutL homologs)/PMS. But they lack MutH and the capacity of recognizing the newly synthesized strand based on hemi-methylation.

The repair process is based on the lagging strand sequence composed of many Okazaki fragments separated by nicks which are considered as nicks created by MutH on the newly synthesized strand in *E. coli*.

DNA DAMAGES CAUSED BY ENVIRONMENTAL FACTORS



HYDROLYTIC DAMAGES

Water can cause DNA damages through : (1) deamination transforming C to U, 5-methyl C to T and (2) depurination which gives rise to abasic sites in the DNA (*figure*).

CHEMICAL DAMAGES

- **Mutagens such as alkylating agents, nitrous acid can alter DNA structure through ethyl/methylation or deamination.**
- **Base analogs are derivatives of normal bases with altered base pairing properties, e.g the enol tautomer of 5-bromouracil base pairs with G instead of A.**
- **Intercalating agents (ethidium, acridine) which bind to purine and pyrimidine bases of DNA generate insertions and deletions.**

PHYSICAL DAMAGES

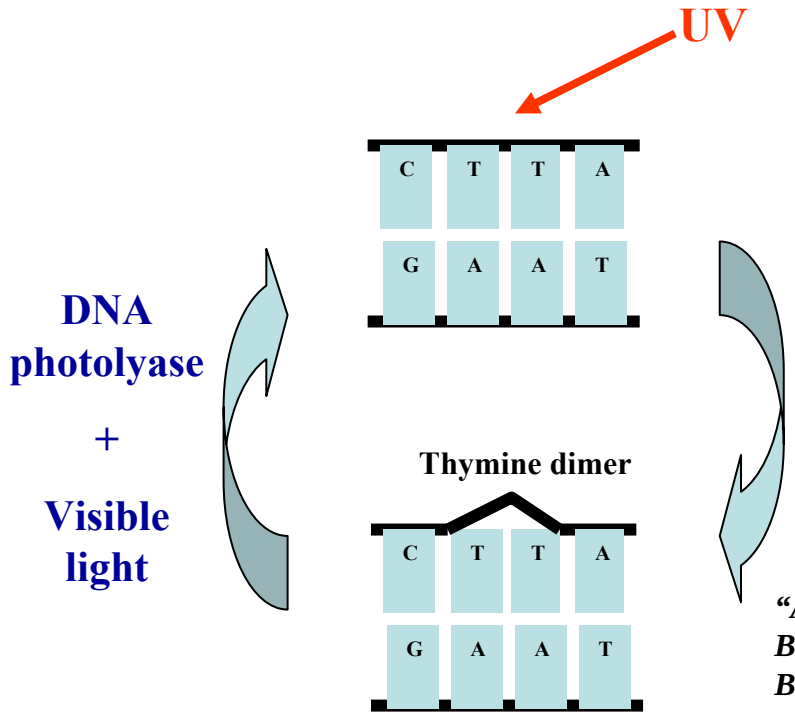
- **Ionizing radiations (X-rays, γ-rays) generate free radicals. The free radicals transform G to oxoG which can base pairs with either C and A. Ionizing radiations can also cause double strand breaks.**
- **Non-ionizing radiation such as ultra violet light supply energy to create new chemical bonds. The most important form of DNA damage is the formation of pyrimidine dimers.**

REPAIR SYSTEMS FOR DNA DAMAGES

Four categories of repair systems are involved in DNA damage repair : Direct reversal repair, Excision repair, Recombinational repair, Translesion synthesis

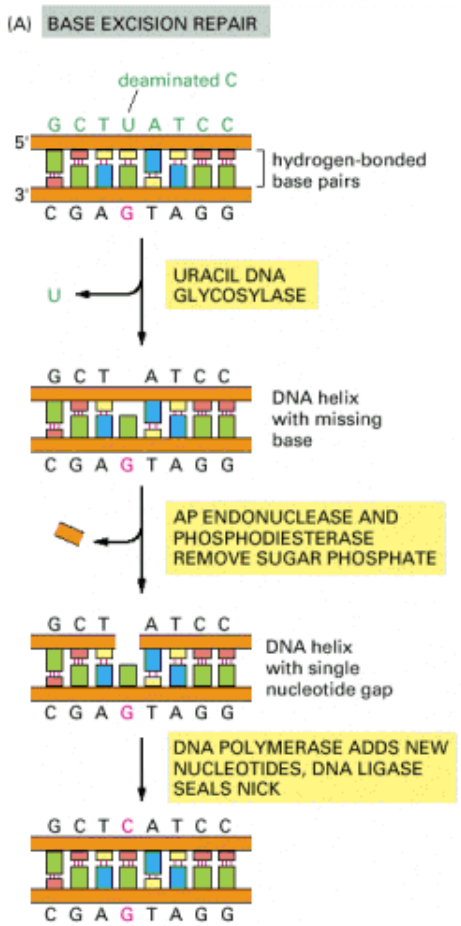
DIRECT REVERSAL REPAIR SYSTEMS

Repair enzymes directly reverse the damage. In photoreactivation, the enzyme DNA photolyase uses energy from visible light to break bonds forming pyrimidine dimers (*figure*). The enzyme methyltransferase reverses the damage caused by alkylation by removing and retaining the methyl group from altered bases.



"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.248, fig 9.11. Benjamin Cummings., CSHL Press"

EXCISION REPAIR SYSTEMS - BASE EXCISION REPAIR



Excision repair systems include the Base excision repair and the Nucleotide excision repair

1. Base excision repair :

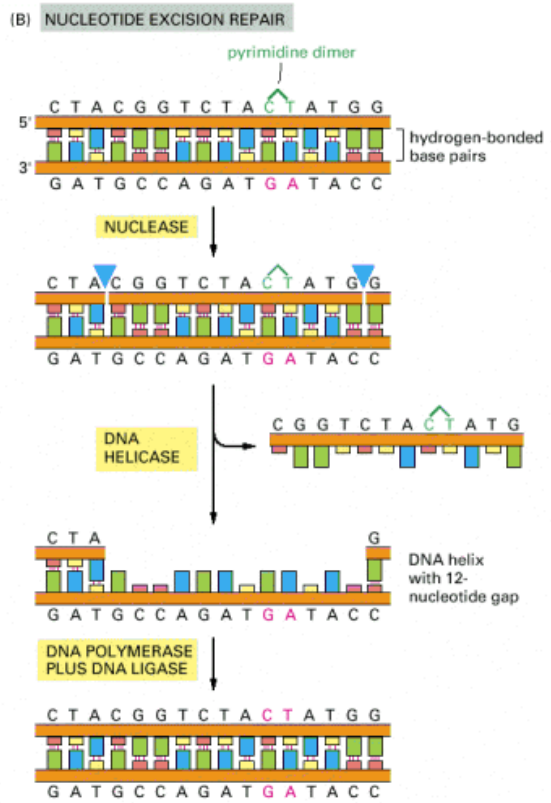
A DNA glycosylase removes the wrong base. The remaining nucleoside is then removed by an endonuclease. The gap is filled in by a repair DNA polymerase using the intact strand as template and is finally sealed by DNA ligase

(A) Base excision repair. An uracil DNA glycosylase removes a deaminated C. After that, the remaning sugar phosphate is cut out by the action of AP endonuclease and a phosphodiesterase. The gap of a single nucleotide is then filled by DNA polymerase and DNA ligase.

EXCISION REPAIR SYSTEMS - NUCLEOTIDE EXCISION REPAIR

Functions of NER repair proteins:	*E. Coli – 4 proteins*	*Eukaryote – more than 25 proteins*
Detect the distortion of DNA due to mutation.	**UvrA**	**XPC**
Separate the two strands and recruit UvrC	**UvrB**	**XPA, XPD**
Cut the single strand bearing the mutation at two sites : at 8 nucleotides upstream and four nucleotides downstream of the site of mutation.	**UvrC**	**ERCC1-XPF, XPG**
Helicase, releases the cut segment	**UvrD**	

Finally, DNA polymerases fills in the gap and DNA ligase seals the nick.



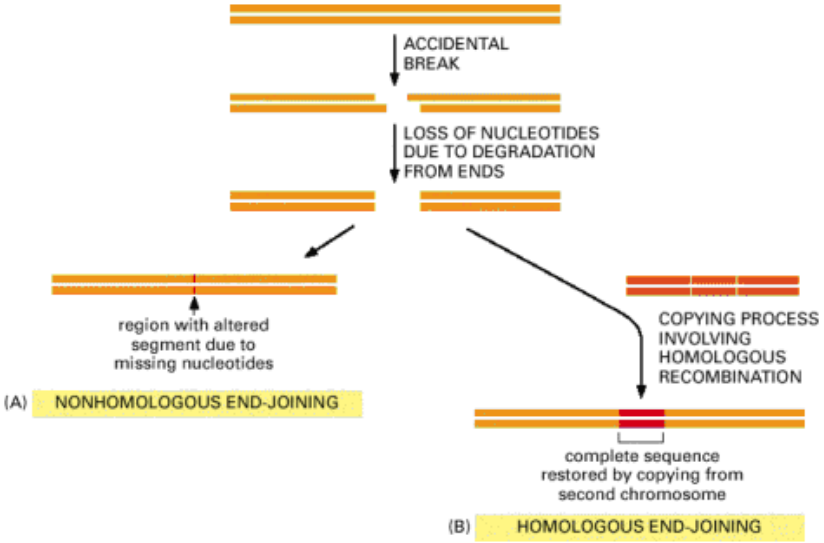
"Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC."

(B) Nucleotide excision repair. Distortion of DNA backbone due to thymine dimer is recognized, cuts are made on each side of the lesion, and a DNA helicase removes the cut segment

RECOMBINATION REPAIR SYSTEMS

RECOMBINATIONAL REPAIR SYSTEMS INCLUDE DOUBLE STRAND BREAK (DSB) REPAIR & NON HOMOLOGOUS END JOINING (NHEJ)

These two repair systems are activated when the double helix is broken and no single strand can be used as template to correct the other strand.



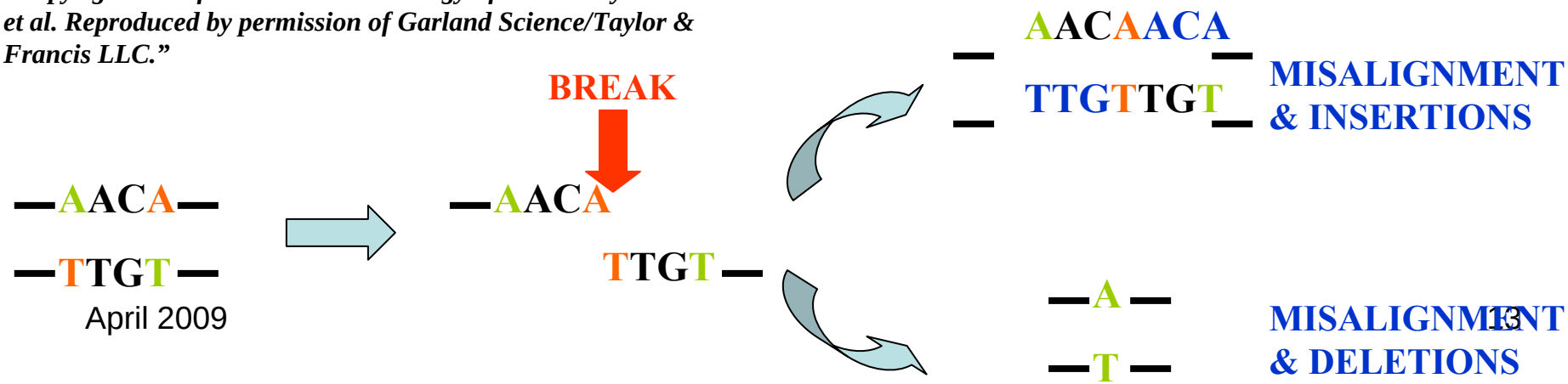
"Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC."

DSB REPAIR

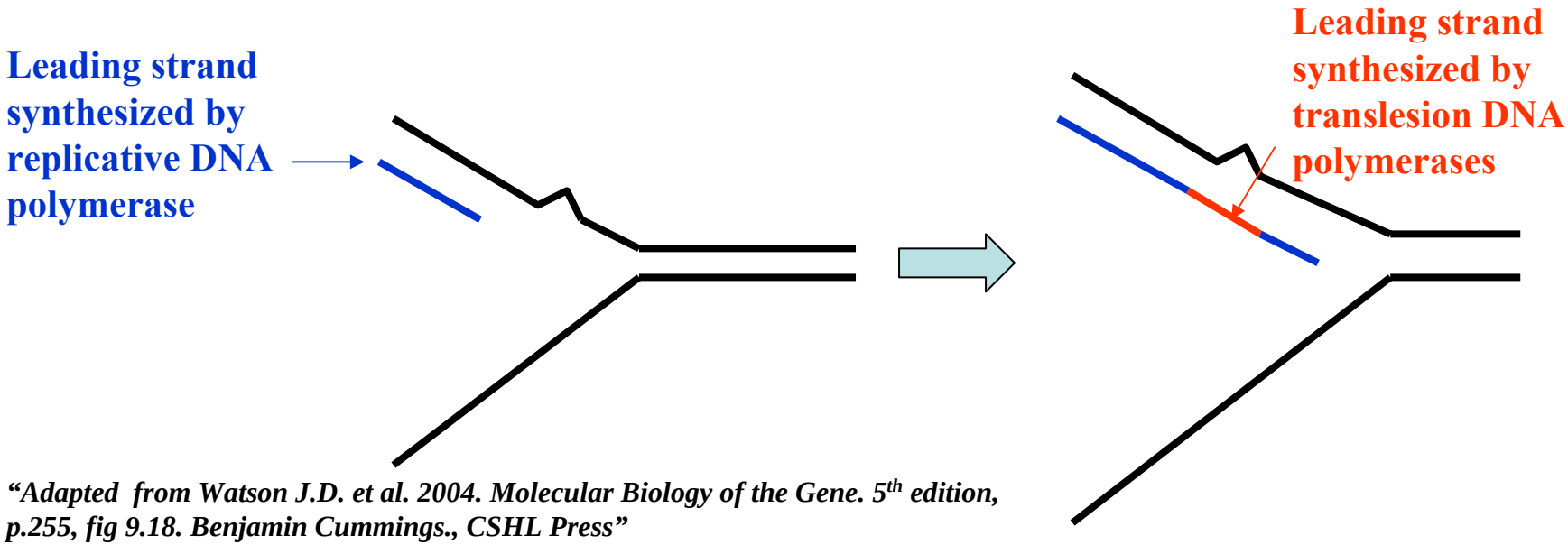
Sister chromosome is used as template to synthesize the broken double-stranded DNA through homologous recombination

NHEJ REPAIR

If DSB arises before sister chromosome is made, the broken DNA could be religated or is repaired by NHEJ systems as follows :



TRANSLESION DNA SYNTHESIS



"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.255, fig 9.18. Benjamin Cummings., CSHL Press"

When DNA synthesis is blocked by a non repaired mutation, a translesion DNA polymerase will replace the replicative DNA polymerase.

The translesion DNA polymerases in *E. coli* are UmuC and UmuD'. They can incorporate nucleotides independently of base-pairing propriety.

Activities of translesion DNA polymerases : (1) allow cell survival by bypassing the lesion which blocks DNA replication, (2) induce important mutagenesis

This phenomenon is called "SOS response" and is only used as the last resort.

SOME HUMAN DISEASES CAUSED BY DEFECTS IN DNA REPAIR SYSTEMS

Table 5-2. Inherited Syndromes with Defects in DNA Repair

NAME	PHENOTYPE	ENZYME OR PROCESS AFFECTED
MSH2, 3, 6, MLH1, PMS2	colon cancer	mismatch repair
Xeroderma pigmentosum (XP) groups A–G	skin cancer, cellular UV sensitivity, neurological abnormalities	nucleotide excision-repair
XP variant	cellular UV sensitivity	translesion synthesis by DNA polymerase δ
Ataxia–telangiectasia (AT)	leukemia, lymphoma, cellular γ-ray sensitivity, genome instability	ATM protein, a protein kinase activated by double-strand breaks
BRCA-2	breast and ovarian cancer	repair by homologous recombination
Werner syndrome	premature aging, cancer at several sites, genome instability	accessory 3′-exonuclease and DNA helicase
Bloom syndrome	cancer at several sites, stunted growth, genome instability	accessory DNA helicase for replication
Fanconi anemia groups A–G	congenital abnormalities, leukemia, genome instability	DNA interstrand cross-link repair
46 BR patient	hypersensitivity to DNA-damaging agents, genome instability	DNA ligase I

DNA RECOMBINATION

HOMOLOGOUS RECOMBINATION

All DNA is recombinant DNA (*)

ꕤ Homologous recombination (general recombination): Exchange of homologous regions between two DNA molecules.

ꕤ The main biological roles of homologous recombination :

(1) Inducing genetic variability through formation of recombinant DNA. During meiosis in eukaryotes, homologous recombination is needed for correct pairing of chromosomes ; it creates new combinations of genes which are transmitted to the descendants

(2) Controlling DSB repair system

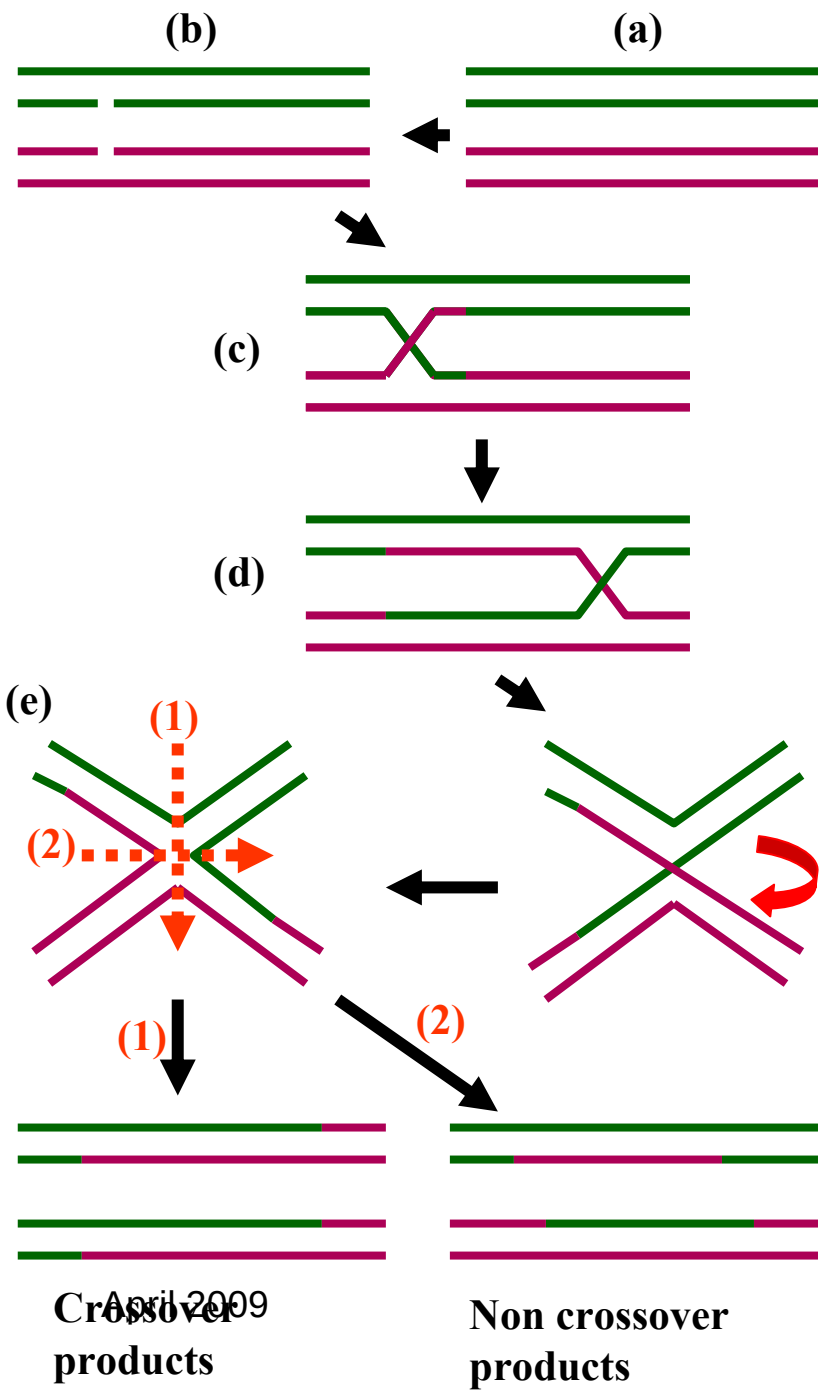
(3) Regulating gene expression by transferring genes between "dormant" or "active" sites in the genome, e.g genes controlling mating types in yeast (MAT locus)

ꕤ Homologous recombination is the basis of many methods used to create "transgenic" or "knock-out" organisms.

ꕤ Homologous recombination can be explained by Holliday model

(*) Watson, J.D, Baker, T.A., Bell, S.P., Gann, A., Levine, M., Losick, R. Molecular Biology of the Gene, 5th ed, Cold Spring Harbor Laboratory Press. San Francisco.

HOLLIDAY MODEL



(a) Alignment of two homologous DNA

(b) Creation of break in a single strand of each DNA molecule

(c) Base pairing between a single-stranded region of one DNA molecule with its complementary strand on the homologous DNA molecule. This step is called strand invasion and results in a structure called Holliday junction

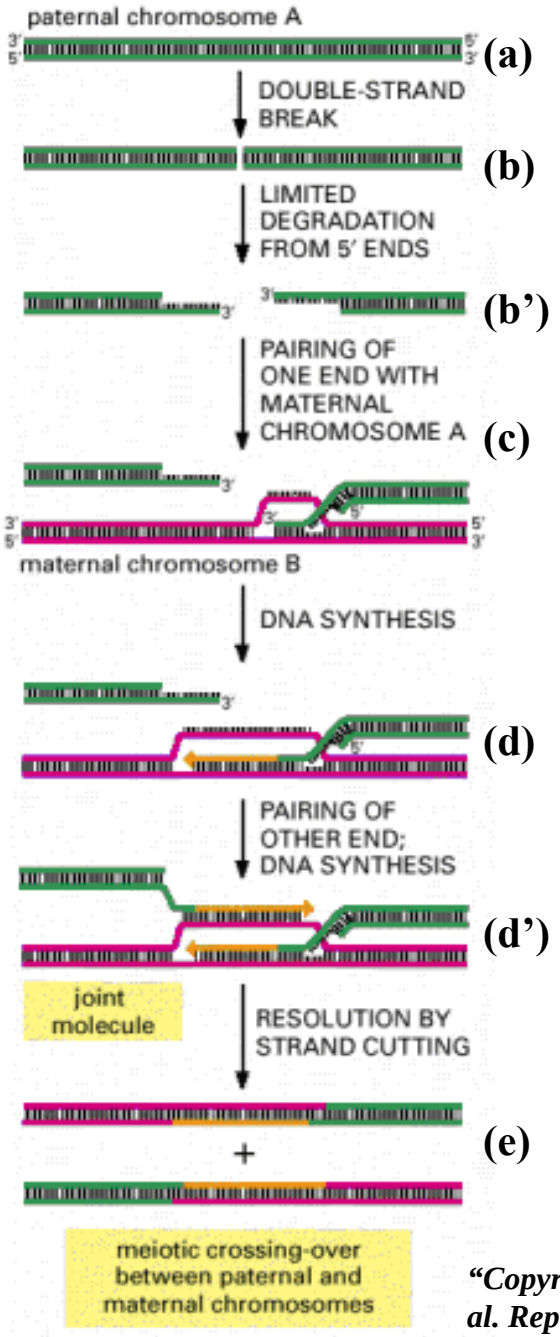
(d) The Holliday junction moves along the DNA.This step is called branch migration. This process increases the length of exchanged DNA regions.

(e) Cleavage of the Holliday junctions. This step is called resolution.

Resolution of Holliday junction could be done in two ways : the cuts marked (1) result in crossover products whereas the cuts marked (2) give rise to non crossover products containing a fragment of "hybrid DNA"

"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, fig 10.1 and 10.2. Benjamin Cummings., CSHL Press"

HOMOLOGOUS RECOMBINATION PROCESS THROUGH DSB



In reality, homologous recombination is often initiated by DSB. Meiotic crossing over is a result of this process including many steps :

(a) Alignment of two homologous DNA, e.g two alleles of a gene.

(b) Creation of double strand break in one DNA molecule

(b') Formation of single-stranded ends of broken DNA by nuclease activities.

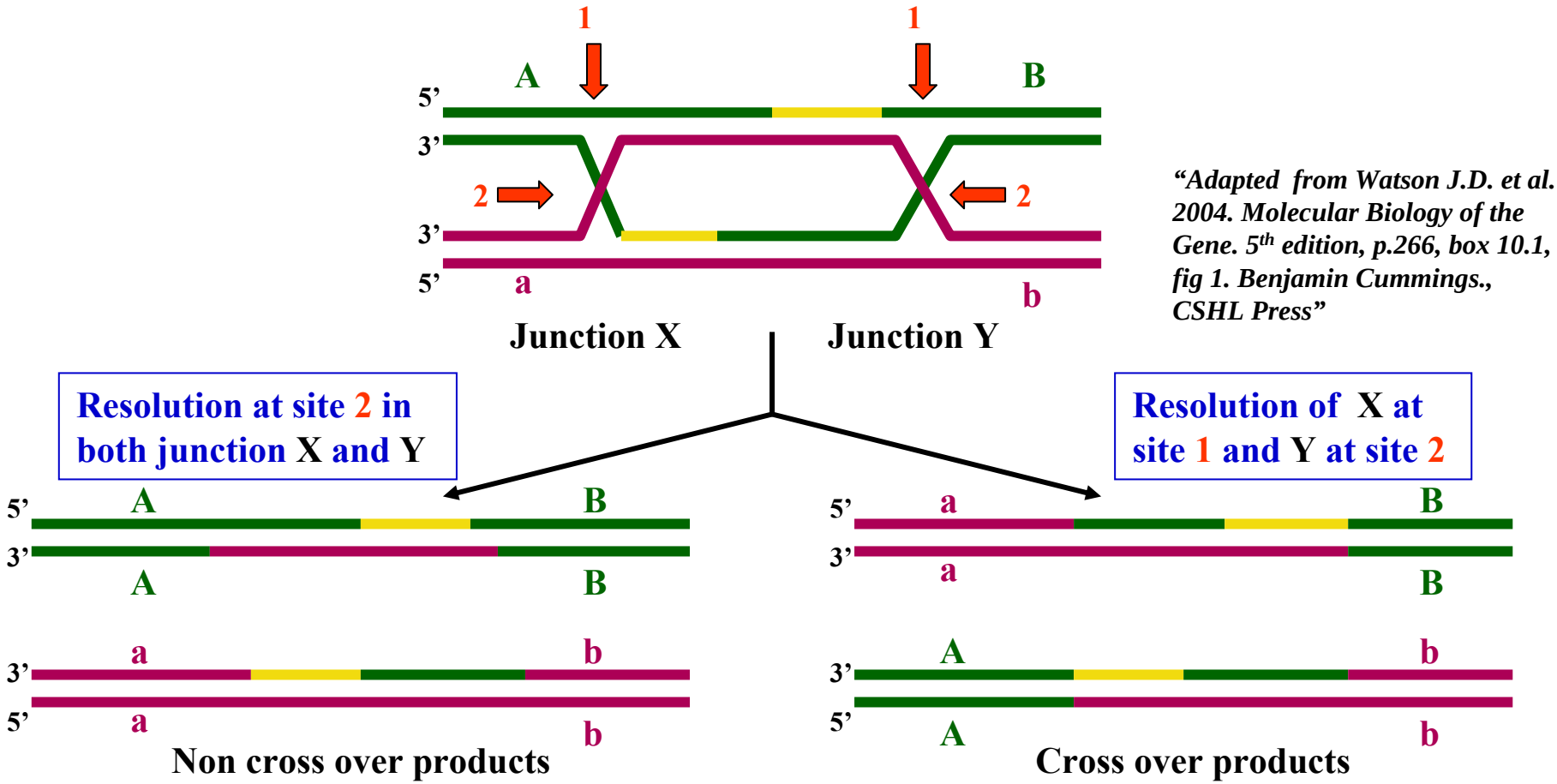
(c) Strand invasion : Base pairing between a single-stranded region of one DNA molecule with its complementary strand on the homologous DNA molecule to form the Holliday junction

(d) Branch migration : The Holliday junction moves along the DNA. The invading strands (*in* green) serve as primers to synthesize new DNA fragment (*in yellow*) to replace the region cleaved at (*b')*

(d') Second Holliday junction is formed and subsequent branch migration

(e) Resolution of the Holliday junctions.

RESOLUTION OF HOLLIDAY JUNCTIONS CREATED BY DSB



"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.266, box 10.1, fig 1. Benjamin Cummings., CSHL Press"

There are many ways to resolve Holliday junctions created by DSB. A rule determines whether cross over or non cross over products are made.

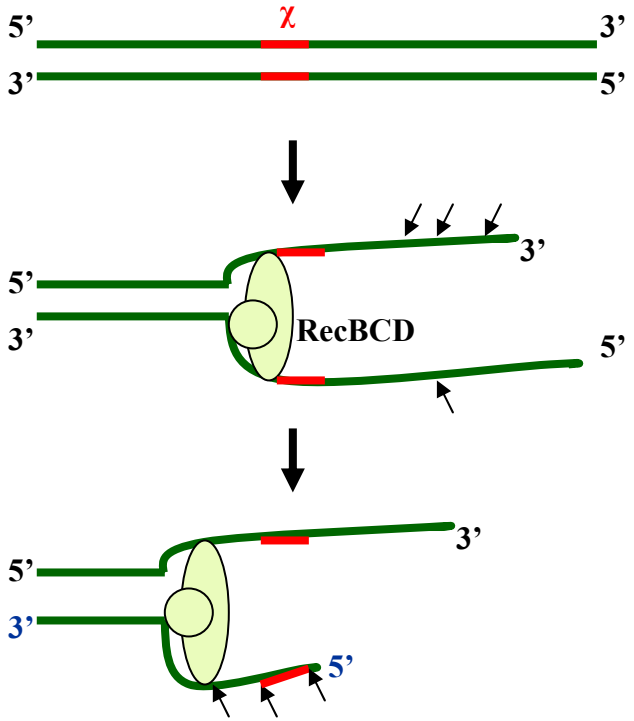
- **If the two junctions are resolved in the same way (both at site 1 or 2), non cross over products will be made**
- **If the two junctions are resolved in different ways (one at site 1, the other at site 2 and *vice versa*), cross over products will be made**

HOMOLOGOUS RECOMBINATION MACHINERY IN *E. COLI*

Proteins involved in homologous recombination in *E. coli* are listed by order of their participation in the process as follows :

RECOMBINATION STEP	RECOMBINATION FACTORS
Creation of DSB on one DNA molecule	No specific protein. DSB in prokaryotes usually results from DNA damage
Generating single-stranded 3' end from DSB	RecBCD complex
Strand invasion : base pairing and strand exchange between homologous DNA, resulting in the Holliday junction	RecA
Branch migration	RuvAB
Resolution of the Holliday junction	RuvC

PROKARYOTIC HOMOLOGOUS RECOMBINATION FACTORS



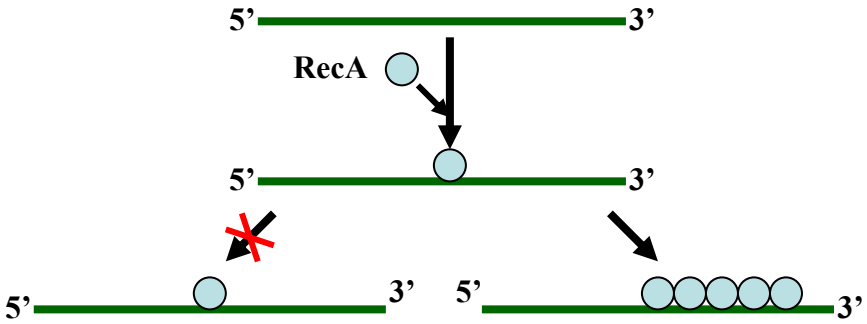
“Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.270, fig 10.5. Benjamin Cummings., CSHL Press”

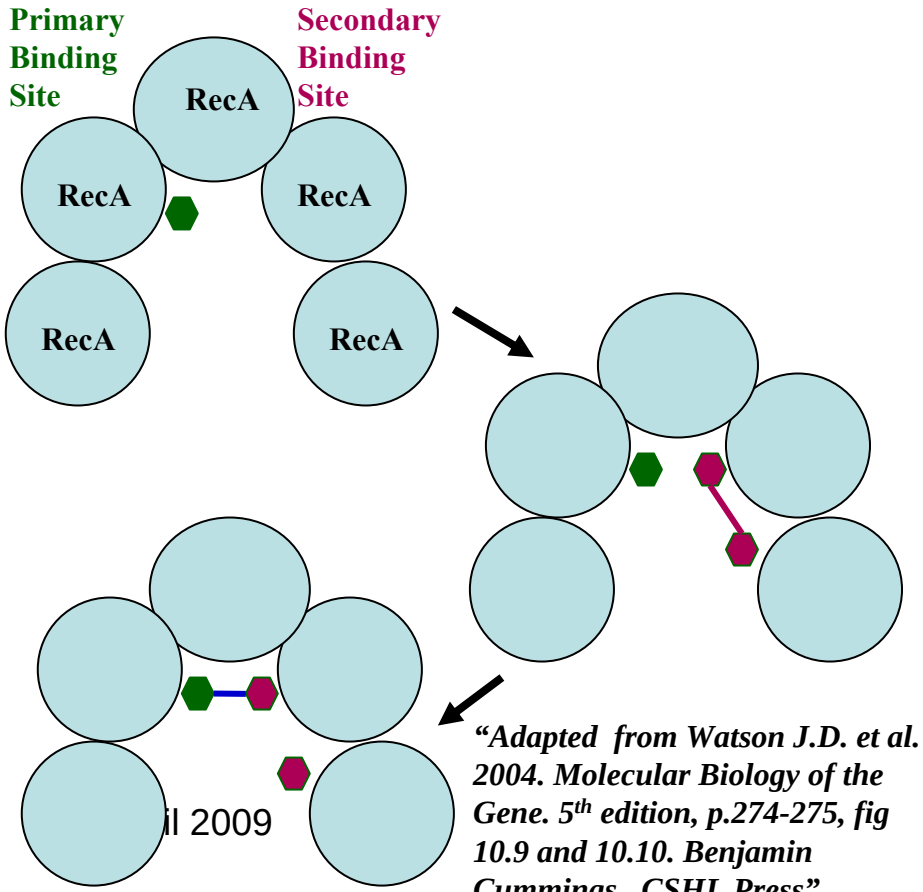
RecBCD complex is composed of three proteins RecB, C, D and has two enzymatic activities, helicase and nuclease. This complex operates as follows :

- **RecBCD complex binds to a site of DNA double strand break, moves along the DNA, progressively unwinds and cleaves the two strands.**
- **When it encounters a chi (χ) site (5’-GCTGGTGG-3’), its nuclease activity changes so that it preferentially cleaves the strand having the 5’-3’ polarity, resulting in a 3’ overhang single strand**

The 3’ single-stranded region is required for the binding of strand-exchange protein RecA to promote recombination

PROKARYOTIC HOMOLOGOUS RECOMBINATION FACTORS





"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.274-275, fig 10.9 and 10.10. Benjamin Cummings., CSHL Press"

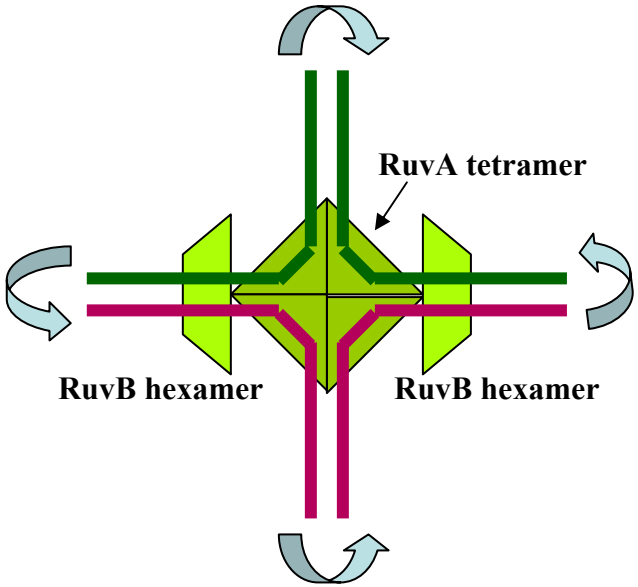
Many RecA subunits (~100) assembles on the 3'single-stranded DNA to form a protein-DNA filament. This assembly grows in the 5'→3' direction → only 3'overhang DNA strand could be the substrate for RecA binding.

Functions of RecA : (1) Search for sequences complementarity, (2) base pairing between homologous DNA and strand exchange

☙ To search for sequence matches, RecA keeps the single strand in its primary binding site and screens many DNA molecules through their transient binding to the secondary binding site.

☙ When the region of complementarity is found, RecA mediates strand exchange through breaking old set of base pairs (–) and formation of new one (–) → switch of base pairing between the three DNA strands

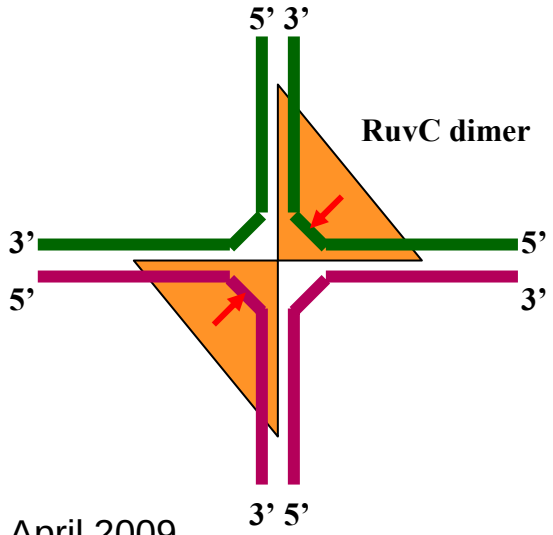
PROKARYOTIC HOMOLOGOUS RECOMBINATION FACTORS



RuvAB complex mediates branch migration :

- **RuvA tetramer recognizes and binds to Holliday junction, recruits two RuvB hexamers**
- **RuvB hexamer is an ATPase which provides energy to drive strand exchange leading to branch migration**

"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.277, fig 10.12. Benjamin Cummings., CSHL Press"



RuvC resolves the Holliday junction by cleaving (↓) the two homologous strands belonging to the two DNA duplexes.

The cleaved ends are then joined by DNA ligase.

Depending on the strand pairs cut, resolution of Holliday junction can give rise to cross over or non cross over products.

"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.278, fig 10.13. Benjamin Cummings., CSHL Press"

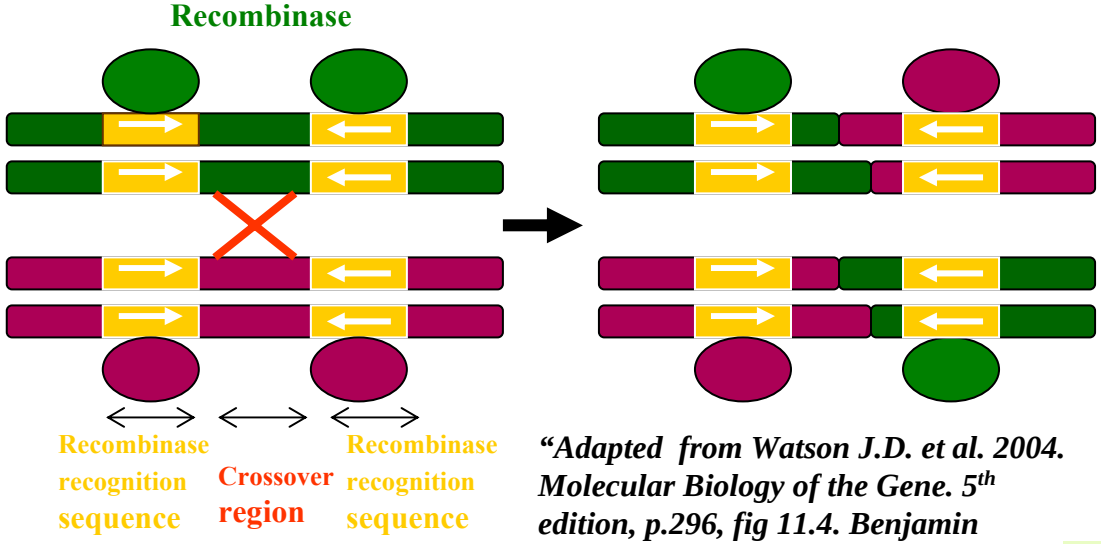
RECOMBINATION FACTORS IN PROKARYOTES AND EUKARYOTES

Recombination process in prokaryotes (*E. coli*) and eukaryotes are very similar including specific recombination proteins having similar functions acting in the same order

RECOMBINATION STEP	*E. COLI*	EUKARYOTE
Generating DSB	None	Spo 11, HO
Generating single strands	RecBCD	MRX (Rad50/58/60)
Assembly of strand exchange proteins	RecBCD	Rad52, Rad59
Pairing homologous DNAs and strand invasion	RecA	Rad51, Dcm1
Holliday junction recognition and branch migration	RuvAB	Unknown
Resolution of Holliday junction	RuvC	Mus81, … ?

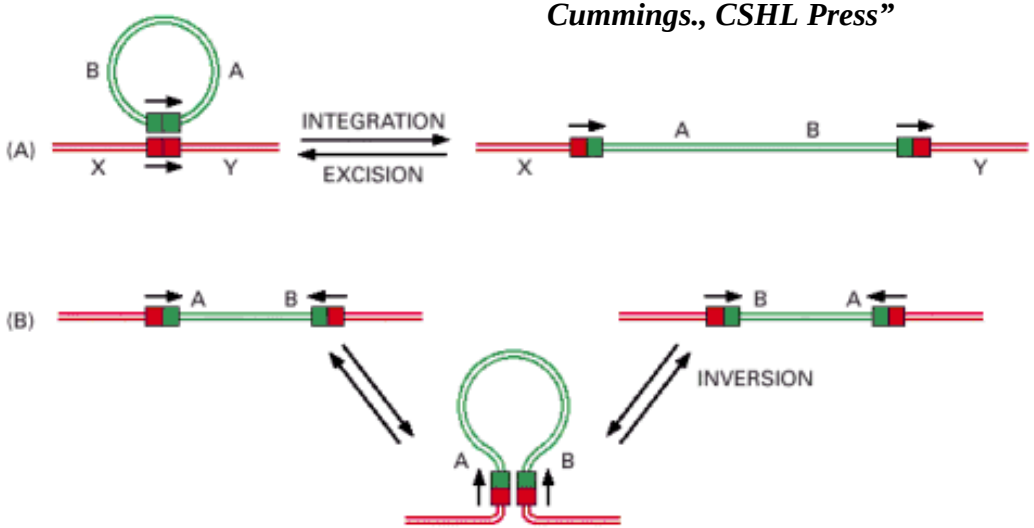
SITE-SPECIFIC RECOMBINATION

Site-specific recombination (SSR) does not involve long similar DNA sequences as in homologous recombination. It occurs between short DNA sequences known as recombination sites.



"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.296, fig 11.4. Benjamin Cummings., CSHL Press"

SSR occurs when the protein recombinase recognizes and binds to recombination sites on DNA. Recombination site includes two types of sequences : two recombinase recognition sequences bound by recombinase flanking a crossover sequence where DNA cleavage and rejoining occur.



"Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC."

SSR can make DNA rearrangements in three ways :

- **Integration (insertion) : when crossover occurs between recombination sites of two different molecules**
- **Excision (deletion) : when recombination sites are direct repeats (→ →)**
- **Inversion : when recombination sites are inverted repeats (→ ←)**

SSR RECOMBINASES

SSR recombinases operate in two steps :

- **When bound to recombination sites, recombinases induce single strand break and form a covalent bond with the cleavage site → formation of a protein-DNA intermediate.**
- **The cleaved DNA end on the other strand interacts with the protein-DNA intermediate, hydrolyzes the covalent bond. The released primary DNA end is then sealed with the second cleavage site using energy stocked in the protein-DNA intermediate → generation of recombinant molecules.**

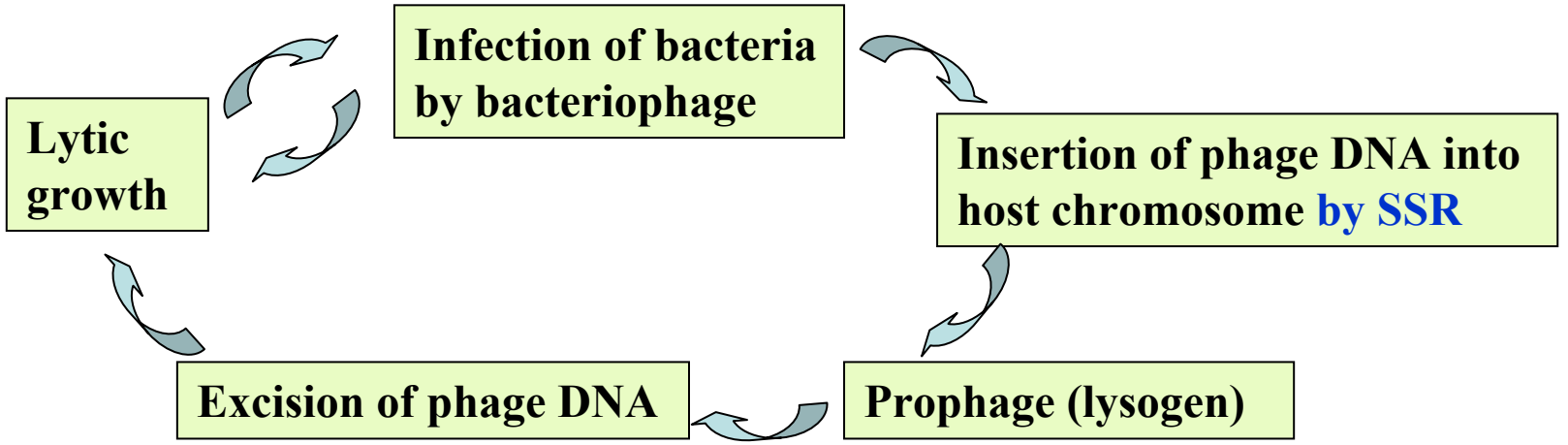
Recombinase + DNA end → Protein-DNA intermediate → DNA end + DNA end → Recombinant molecule

SSR recombinases are classified into two families : serine recombinases and tyrosine recombinases :

- **Serine recombinases : The active group which interacts with DNA cleavage site is a serine. Four subunits of the enzyme work simultaneously to induce break in the four strands of the two exchanged DNA molecules → Strand swap → Reseal → Recombinant molecule.**
- **Tyrosine recombinases : The active group is a tyrosine. Two subunits of the enzyme cleave and rejoin two DNA strands first followed by the action of the other subunits.**

BIOLOGICAL MEANINGS OF SITE-SPECIFIC RECOMBINATION

Phage insertion into host chromosome is an example of site-specific recombination



Modification of gene expression through site-specific recombination, e.g Hin recombination in *Salmonella.*

ꟷ Flagellin is a surface protein of the bacteria, thus is a target of host immune system.

ꟷ Genes encoding flagellin exist in two alternative forms, H1 and H2, which have inverted orientation. When H1 is in "on" orientation, it expresses whereas H2 is repressed ("off"), and *vice versa*.

ꟷ Hin recombinase induces inversion of these genes allowing the expression of different flagellins in some individuals of the bacterial population.

→ These individuals can escape host immune system and proliferate

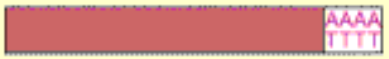
DNA TRANSPOSITION

Barbara McClintock (1902-1992) elucidated genetic bases of transposable elements through her studies in maize. She won a Nobel Prize in 1983, 35 years after her first publication on Transposition.

- **Transposable elements (transposons) are mobile fragments which can move from one site to another site in the genome.**
- **Transposons exist in the genome of all species. Their percentage varies – very high in some species such as human or maiz, .., low in *Drosophila* and yeast, ..**
- **Insertion of transposons into new sites can disrupt gene products (insertion into coding region) or change gene expression (insertion into regulatory sequences)**
- **Transposition is the most common cause of DNA mutations in many organisms.**
- **Transposition is used to disrupt gene functions in Genetic engineering.**

Based on structure and mechanism of action, transposable elements are classified into different groups

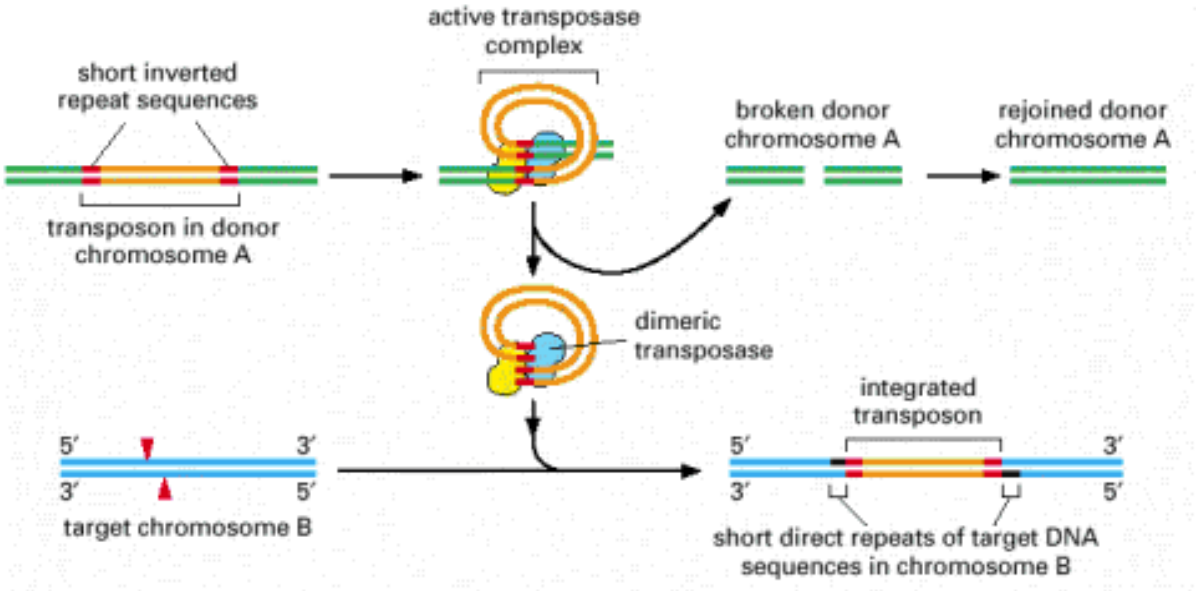
STRUCTURE AND MECHANISM OF ACTION OF THE THREE FAMILIES OF TRANSPOSABLE ELEMENTS

CLASS DESCRIPTION AND STRUCTURE	GENES IN COMPLETE ELEMENT	MODE OF MOVEMENT	EXAMPLES
DNA-only transposons			
			
short inverted repeats at each end	encodes transposase	moves as DNA, either excising or following a replicative pathway	P element (*Drosophila*) Ac-Ds (maize) Tn3 and IS1 (*E.coli*) Tam3 (snapdragon)
Retroviral-like retrotransposons			
			
directly repeated long terminal repeats (LTRs) at ends	encodes reverse transcriptase and resembles retrovirus	moves via an RNA intermediate produced by promoter in LTR	Copia (*Drosophila*) Ty1 (yeast) THE-1 (human) Bs1 (maize)
Nonretroviral retrotransposons			
			
Poly A at 3′ end of RNA transcript; 5′ end is often truncated	encodes reverse transcriptase	moves via an RNA intermediate that is often produced from a neighboring promotor	F element (*Drosophila*) L1 (human) Cin4 (maize)

These elements range in length from 1000 to about 12,000 nucleotide pairs; each family contains many members, only a few of which are listed here. In addition to transposable elements, there are selected viruses that can move in and out of host cell chromosomes; these viruses are related to the first two classes of transposons.

TRANSPOSITION CAN OCCUR IN TWO WAYS : NON-REPLICATIVE (CUT –AND PASTE) AND REPLICATIVE

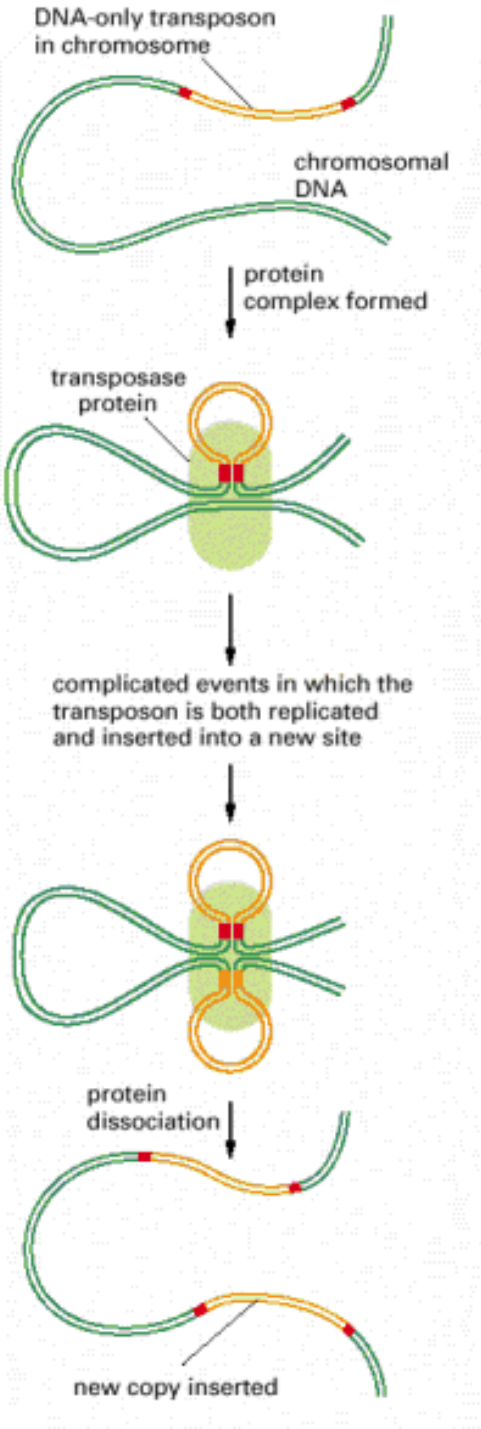
CUT-AND-PASTE (NON-REPLICATIVE) TRANSPOSITION



Excision of the transposon from its initial location on the donor chromosome and integration of this transposon into new DNA site on the target chromosome

"Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC."

- **Subunits of transposase, the enzyme responsible for transposition, bind to the inverted repeats at each end of the transposon and bring them together to form a transpososome**
- **Transposase subunits cleave the junction between the transposon and host DNA**
- **The ends of the excised transposon are joined to cleaved ends of the new site resulting in its insertion in the new site.**
- **DSB at the "old" site is repaired by homologous recombination or NEJ ; gaps at the "new" site are filled by repair polymerases and sealed by DNA ligase → duplication of short repeats of the target DNA**



REPLICATIVE TRANSPOSITION

The replicative transposition follows the following steps :

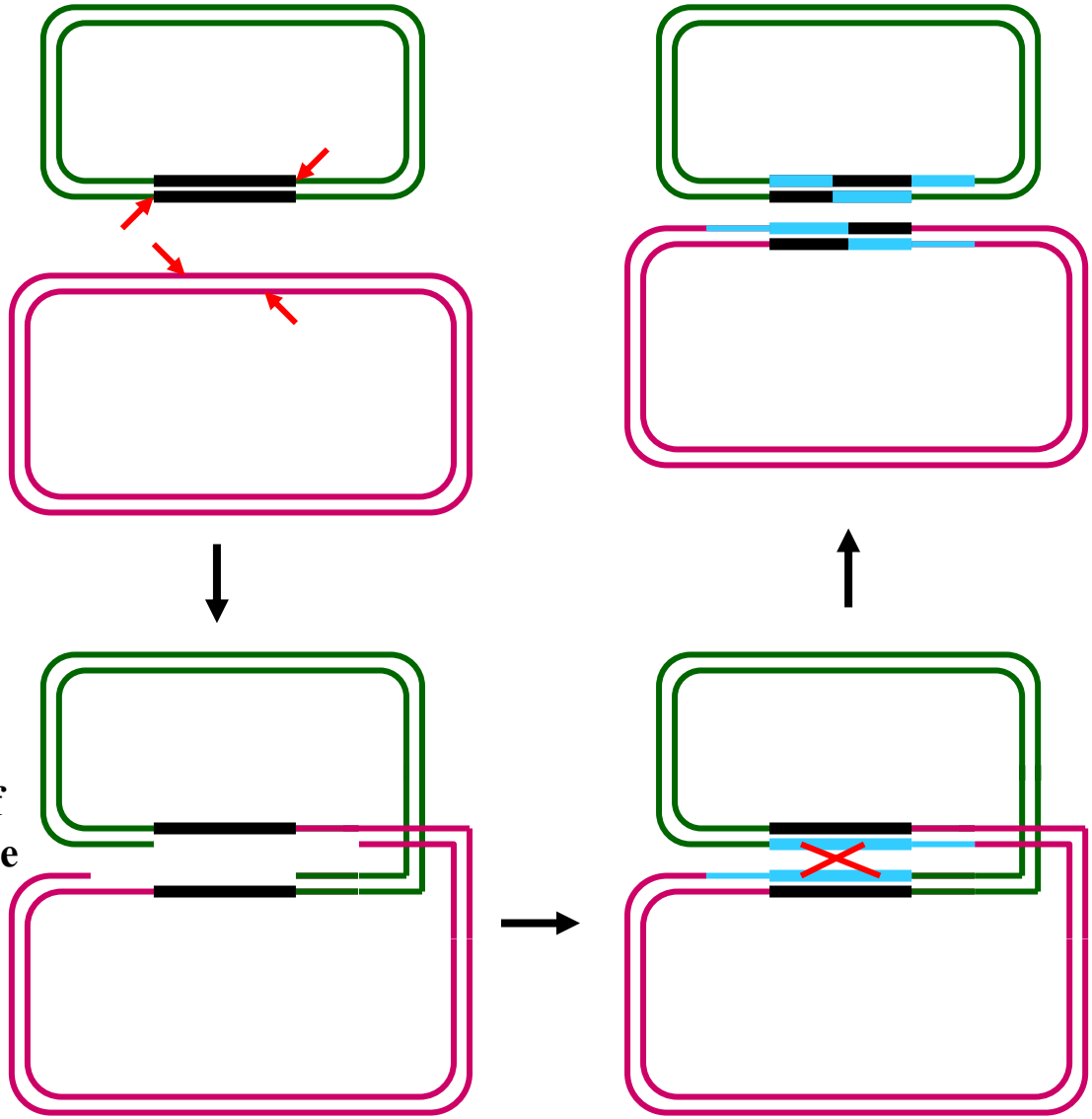
- **Assembly of transposase subunits on inverted repeats at each end of the transposon and formation of the transpososome**
- **Transposase subunits cleave at the end of the transposon to release the 3'OH group at each end. The transposon is not excised from its initial location. This is the main difference between replicative and non-replicative transposition**
- **The 3'OH ends of the transposon are joined to 5'ends of the "new" cleaved site whereas the 5'P ends of the transposon remain linked to the "old" site**
- **Replication generates two copies of the transposon, one remains at the initial location, the other is inserted in a new location**

REPLICATIVE TRANSPOSITION PROCESS

Transposon is cleaved releasing two 3'OH ends

The target DNA is cleaved releasing two 5'P ends

The 3'OH ends of the transposon are ligated to the 5'P ends of target DNA

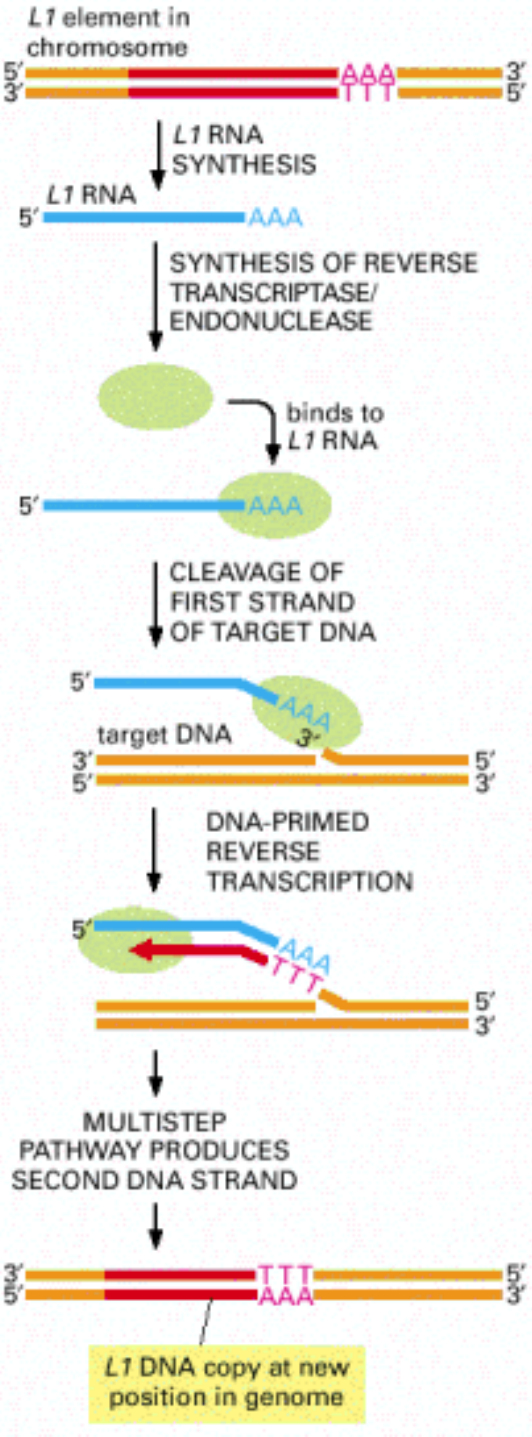


Replicative transposition results in two copies of the initial transposon

The transposon and adjacent sequences are replicated. The resulting structure is called a cointegrate. Recombination occurs between the two transposon copies

"Adapted from Watson J.D. et al. 2004. Molecular Biology of the Gene. 5th edition, p.319, fig 11.22. Benjamin Cummings., CSHL Press"

TRANSPOSITION OF POLY-A RETROTRANSPOSON

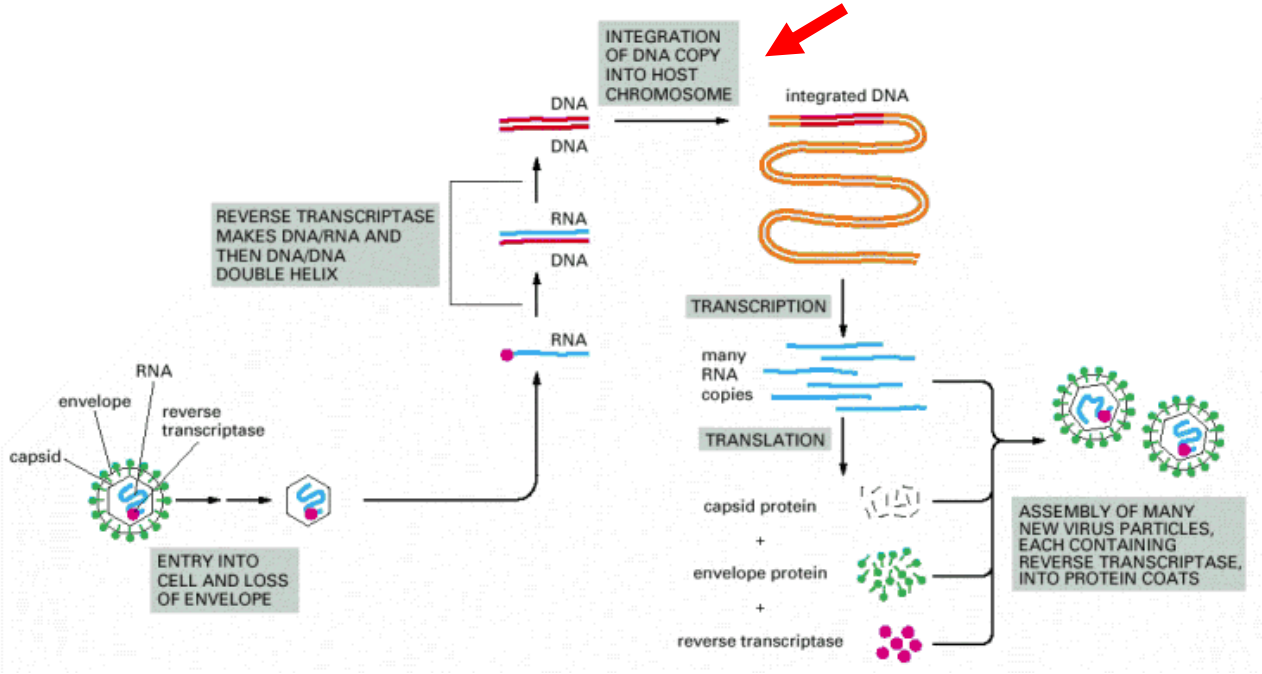


The transposition of a poly-A retrotransposon includes many steps :

- **Transcription of the retrotransposon by cellular RNA polymerase**
- **The resulting mRNA is translated into proteins which bind to its 3'end**
- **The protein-mRNA complex binds to a T-rich site in the target DNA, induces a nick in this site and forms a RNA:DNA hybrid**
- **The 3'end of the cleaved target site is used as a primer to synthesize a cDNA on the RNA template**
- **The initial mRNA is degraded and the second DNA strand is synthesized followed by DNA joining and ligation.**

→ A new copy of the poly-A retrotransposon is inserted in a new location

TRANSPOSITION OF RETROVIRAL-LIKE RETROTRANSPOSON



The life cycle of a retrovirus. The integration (↙) of the retrovirus into host genome is catalyzed by a viral enzyme, the integrase

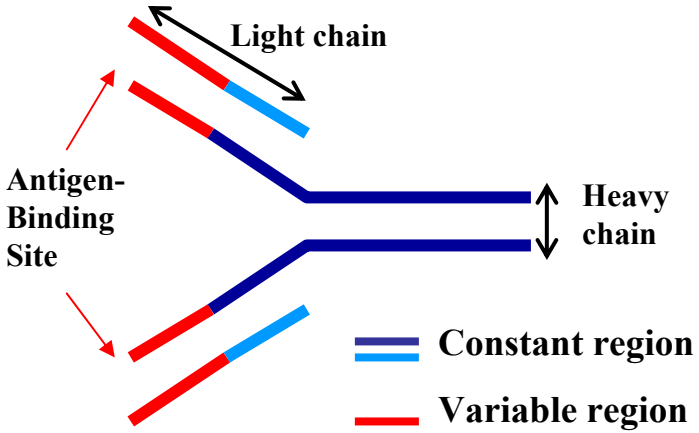
"Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC."

Retroviral-like retrotransposon and retrovirus transpose as follows :

- **Retrotransposon (or prophage) is transcribed from a promoter located within one of the LTR**
- **The resulting RNA is then reverse transcribed into a double-stranded cDNA**
- **Integrase removes some nucleotides from the 3'end of the double-stranded cDNA**
- **The processed retrotransposon is then inserted to a new site on the target DNA**

AN EXAMPLE OF TRANSPOSITION : V(D)J RECOMBINATION

The immune system of vertebrates has to recognize a vast spectrum of "foreign" molecules encountered by the organism during its life time, and to destroy them. The specific recognition of these "foreign" molecules is made by T-cell receptors produced by T cells and antibodies produced by B cells. Take the example of an antibody



An antibody is composed of two light chains and two heavy chains *(figure)*. Each chain includes variable and constant regions. Variable regions are involved in the recognition of various antigens.

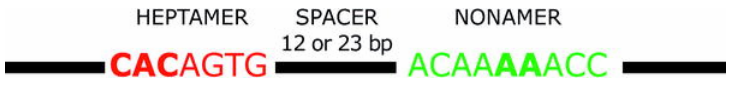
The diversity of these variable regions is created by differently combining V (variable), D (diveristy) and J (joining) genes in a process called V(D)J recombination (table below)

Element	Immunoglobulin		α:β Receptor	
	H	κ + λ	β	α
V segmats	65	70	52	70
D segments	27	—	2	—
J segments	6	5κ 4λ	13	61
Number of V region combinations	3.4×10^6	3.4×10^6	5.8×10^6	5.8×10^6
Junctional diversity	3×10^7	3×10^7	2×10^{11}	2×10^{11}
Total diversity	10^{14}	10^{14}	10^{18}	10^{18}

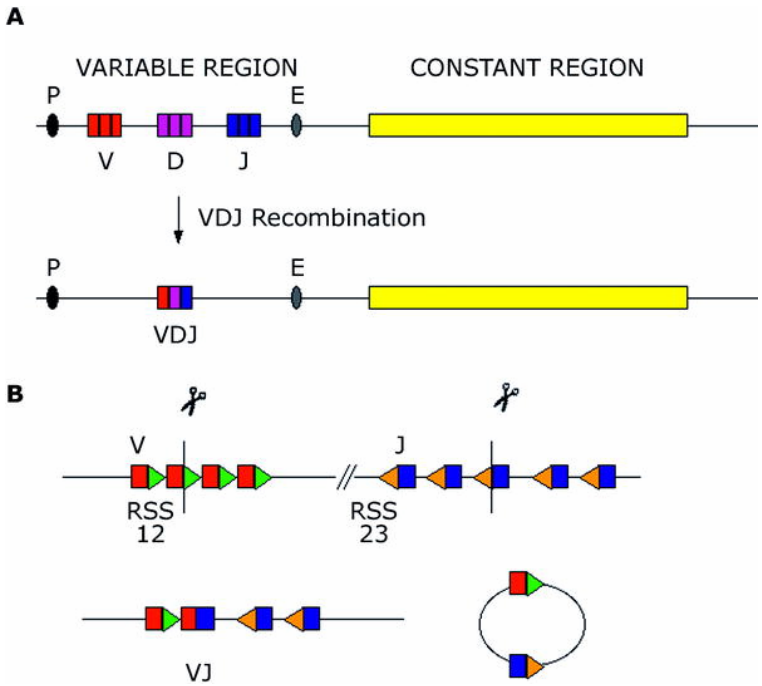
Number of V, D, and J segments contributes to combinatorial diversity. Further changes are introduced by junctional diversity, to give the total number of BCR and TCR repertoires. DOI: 10.1371/journal.pbio.00000016.t001

Market E., Papavasiliou F.N. 2003. V(D)J recombination and the evolution of the adaptive immune system. PLoS Biol 1(1)

V(D)J RECOMBINATION (continued)



RSS sequence composed of 7-mer and 9-mer conserved motifs which are bound by recombinases



Market E., Papavasiliou F.N. 2003. V(D)J recombination and the evolution of the adaptive immune system. PLoS Biol 1(1) e16
doi:101371/journal.pbio.0000016

V(D)J recombination is a "cut-and-paste" reaction including many steps :

- **Two recombinases, RAG1 and RAG2, recognize recombination signal sequences (RSS) flanking V, D, J genes and bring them together.**
- **RAG complex makes a single-strand cleavage at the junction between a RSS and V, D, or J gene flanked by this RSS.**
- **In the case of a light chain, the released 3'OH free ends of one of the V genes will then attack the 5'P ends of one of the J gene at random to create a defined VJ combination associated with the C gene**

In a heavy chain, the recombination reaction occurs in two steps : (1) random recombination between one D gene and one J gene, (2) subsequent recombination between a V gene and the preformed DJ combination.

V(D)J recombination is similar to Cut-and-Paste recombination, involving recombination sites (RSS) and transposases (RAG1, RAG2) acting in the same way

SUMMARY

DNA variations are the consequences of three major phenomena : DNA Mutation, DNA Recombination and Transposition.

DNA MUTATION :

- **Mutations occurred during or just after replication can be repaired by DNA polymerases proofreading acivities or Mismatch repair system**
- **Other DNA damages, caused by water, chemical or physical agents, can be repaired by one of the four categories of repair systems : (1) Direct reversal, (2) Excision repair systems including Base Excision and Nucleotide excision repair, (3) Recombinational repair systems including Double Strand Break (DSB) and Non Homologous End Joining (NHEJ) repair, (4) Translesion repair.**

HOMOLOGOUS RECOMBINATION (HR)

- **HR is the exchange of fragments between two homologous DNA molecules**
- **HR induces genetic variability, controls the DSB repair system, regulates gene expression**
- **HR can be explained by the Holliday model**
- **Recombinational machinery in *E. coli* includes : RecBCD, RecA, RuvAB, RuvC which are responsible of the generation of 3' overhang single strand, strand exchange between homologous DNA, branch migration and resolution of the Holliday junction. Homologs of these recombination proteins are found in eukaryotes.**

SUMMARY (continued)

SITE-SPECIFIC RECOMBINATION (SSR)

- SSR involves the exchange of short sequences called recombination sites
- SSR can cause insertion, deletion or inversion of DNA sequences between recombination sites
- SSR is used for phage insertion into host genome or to modify gene expression

TRANSPOSITION

- Transposable elements are composed of transposon, retroviral-like retrotransposon and poly-A retrotransposon
- Transposition can disrupt gene function or change gene expression by insertion into gene coding or regulatory sequences
- Transposable elements can operate by non-replicative (cut-and-paste) or replicative process
- The retroviral-like and poly-A retrotransposon tranpose through an RNA intermediates

An example of transposition concerns V(D)J recombination which generates a huge diversity of the immune factors (antibodies and receptors)