جامعة جنوب الوادي قسم النبات والميكروبيولوجي كلية العلوم

Lectures and practical lessons

Cytogenetics

2nd year biology students

Prepared by
Dr/ Abdelraheem Qenawy

First semester
2024 / 2025

المحاضرة الأولي

Content:

Introduction:

1- Historical review

a – Chromosome b- Gene

c – Relation between the two (Theory of chromosomes)

 # Fell Helm raw 1983; Suggested that chromosomes are regular bodies have the ability to divide accurately during cell division

 # Johannse ; Suggested gene instead of factor

Theodor Boveri 1902 Supported theory and Walter Sutton of chromosomes by strong practical evidence

 # Morgan *et all*

which indicate the relation shipe between gene and chromosome

Oswald Avery was an American scientist who studied bacteriology, immunology, and molecular biology. He lived from 1877-1955 and was responsible for identifying DNA as the substance to carry genetic information.

The Hershey–Chase experiments were a series of experiments conducted in 1952[1] by Alfred Hershey and Martha Chase that helped to confirm that DNA is genetic material.

2- Mendel and principles of heredity

3- Main divisions of genetics

a-Mendelian genetics (General genetics)

b- Cytogenetics

c- Gene-engineering

4- Branches of genetics

Genetic engineering •
Mutagenetics •
biochemical genetics •
Radiation genetics •
Molecular genetics •
Cytogenetics •
Quantitative genetics •
Developmental genetics •
Population genetics •
Human genetics •
Microbial genetics •

Physiological genetics •

Technical of genetic engineering

1- Cleavage of DNA •
2- Gel-electrophoresis •
3- DNA sequencing •
4- Nucleic acid hybridization •
5- DNA cloning •
6- DNA engineering

2- Discussion

- لست بحاجة الي أن أذكرك بأن علم الوراثة قديم جدا وحديث جدا في نفس الوقت فالأنسان القديم عرف الوراثة في الكائنات الحية بفطرته الذكية من خلال مشاهداته لما يحدث من تعاقب الأجيال دون أن تندثر صفات الأباء المؤسسين بينوع الواحد فلا القمح عل مر الأجيال أنتج برسيما ولا الفيل أنتج ذئبا هذا يعني أن المعلومات الخاصة بمواصفات كل نوع محفوظة في مكان أمن داخل الخلية الحية وأنها قادرة علي التعبير عن نفسها بدقة شديدة بل متناهية الدقة متي أتيح لها ذلك لم تخطئ أبدا الخطأ الذي يقلب النوع رأسا علي عقب رغم مرور ملايين السنين علي تعاقب هذه المعلومات من جيل الي جيل.
- الشيئ الذي كانت تفتقده البشرية خلال قرون عديده وأزمان بعيده أن يوضع لهذا السر الكامن في الخلايا علم يوضحة ويشرحه علم تكون له أسس وقواعد شارحة ومفسرة بل ومتنبأة احيانا بما سيكشف عنه الغيب . لكن لأن هذا السر الوراثي كامن وخفي كان لابد من وجود كاشف يكشف ما في أعماق الخلية من أسرار لذلك نهض علم الوراثة نهضة كبيرة بعد اكتشاف الميكروسكوب وظهور علم الخلية فبعد أن كان علم الوراثة معتمدا فقط علي مشاهدة الظاهر من الصفات بالعين المجردة وهوما سمي فيما بعد الوراثة العامة أو الوراثة المندلية اعتمادا علي أن مندل كان من الرواد الكبار في علم الوراثة الذي اعتمدت دراساته علي الظاهر من الصفات لأنه لم يكن بأمكانه دراسة المورثات من الداخل فلم يكن علم الخلية قد ظهر بعد ومع ظهور هذا العلم وتقدمه ظهرت الوراثة الخلوية وهي الفرع الثاني الكبير بعد الوراثة العامة من أفرع علم الوراثة وساعد علي ذلك وضوح العلاقة بين الجينات حاملة الصفات والكروموسومات حاملة الجينات. ولكن وبالرغم من معرفة **DNA**وتحليله كيميائيا ومعرفة مكوناته من سكر وفوسفات وقواعد نيتروجينية وعلاقة هذه القواعد بعضها ببعض ظل السر الوراثي المكتوب

داخل الحمض النووي أمرا محيرا للغاية بل مستحيلا أن يري أويقرأ . وفي قلب هذه الحيرة المظلمة أشرقت شمس المعرفة من جديد حاملة البشري بعهد جديد في علم الوراثة وهوالوراثة البيولوجية الدي تمخض عنه والذي تمخض عنه أخطر واعظم وأدق علوم البيولوجي وهوالهندسة الوراثية الذي سوف نتناوله في حينه ان شاء الله بالتفصيل هذه ياعزيزي علي الأجمال رحلتك ان شاء الله مع علم الوراثة تمنياتي للجميع بالتوفيق.

دعني أذكرك بما درسته من قبل في الوراثة المندلية ونحن الأن بحاجة اليه لفهم الوراثة العامة وتطبيقها في حل مسائل وراثية في العملي

- أقدم أفرع علم الوراثة الثلاثة الكبار وأوسعها انتشارا (الوراثة العامة – الوراثة الخلوية – الوراثة الجزيئية)
- 1- اسسها مندل ثم كورينر- دي فريز وتشيرماك
- 2- تشمل:
- ا- الوراثة المندلية ب- تداخل الجينات
- ج- وراثة العشائر د- الصفات الكمية

عوامل نجاح تدارب مندل

- 1- اختيار النبات المناسب بعناية ودقة
  2- تصميم تجاربه
- 3- الاعتماد علي التحليل العددي والاحصائي للنتائج التي حصل عليها
- 4- دراسة الصفات المختلفة كل علي حدة بدلا من دراسة الفرد كاملا بكل صفاته مجتمعة

مميزات البسلة

- 1- لها صفات ثابته يمكن التعرف عليها بسهولة

- 2- ذاتية التلقيح ويمكن اجراء التلقيح الصناعي بسهولة
- 3- النسل الناتج تام الخصوبة
- 4- سهولة زراعتها وقصر فترة نموها

الصفات السبع للبسلة التي تم دراستها

- حصل علي 34 صنف انتخب 22 لتجاربه ودرس سبع صفات متفارقة وهي:
- 1- طول الساق 2- وضع الأزهار علي المحور الزهري
- 3- لون غلاف البذرة 4- شكل البذور
- 5- لون فلقتي البذرة 6- شكل القرون
- 7- لون القرون

قوانين مندل

- **القانون الأول** (قانون الانعزال) Law of segregation وينص علي أنه:
- " اذا اختلف فردان في زوج من صفاتهما المتفارقة فانهما ينتجان عند تزاوجهما جيلا به صفة أحد الأبوين فقط وهي الصفة السائدة بينما تعاود الصفتان الظهور مرة أخري في أفراد الجيل الثاني بنسبة عددية ثابته 3 : 1 (سائد الي متنحي)."
- مثال:
- نبات طويل (سائد) مع نبات قصير (متنحي)
- الجيل الأول : نباتات طويلة ,,,,,,,,,,,,,,,,اذا تركت تتزاوج ذاتيا تعطي الجيل الثاني كالاتي:
- نباتات طويلة 3
- نباتات قصيرة 1

P1: TT x tt .

G1: T t .

F1: Tt .

.

P2 Tt x Tt .

G2: T t .

F2 TT Tt Tt tt .

القانون الثاني (التوزيع الحر أو المستقل)
Law of independent assortment

- " اذا اختلف فردان في زوجين من الصفات المتفارقة تظهر صفتي كل زوج بنسبة 3 : 1 في الجيل الثاني وتوزع صفتي كل زوج توزيعا مستقلا عن صفتي الزوج الأخر"
- مثال: نبات طويل أحمر الأزهار مع قصير أبيض الأزهار

P1: TTRR x ttrr

G2: TR tr

F1: TtRr

طويل أحمر الأزهار

- يترك ليتلقح ذاتيا

- الطرز المظهرية في الجيل الثاني
-
- طويل أحمر 9
- .....أبيض 3
- قصير أحمر 3
- قصير أبيض 1
- النسبة:
- 9 : 3 : 3 : 1
- الطويل : القصير (3 : 1)
- الأحمر : الأبيض (3 : 1) اذا كل صفة تتوزع مستقلة

التطبيقات العملية لقوانين مندل

- 1- اختبار التراكيب الوراثية
- 2- استنباط السلالات الجديدة
- 3- الوراثة المندلية في الانسان

المحاضرة الثانية

Effects of chemical agents on chromosome

The effects of these chemicals grouped into:

A- Metaphase arrest

Chemicals such as : colchicine – gammaxane – chloral hyrdrate – acenaphthene – actidione

Mode of action:

- **They cause metaphase arrest by inhibiting the operation of the spindle mechanism.**

B- Polyploidy

- **The importance of polyploidy due to the increase in gene dosage resulting from multiplication of chromosome sets bring about giganticism in all characters in general.**
- **Chemicals: colchicine**

C- Chromosome fragmentation

- **The importance:**
- **1- Fragmentation followed by translocation of some fragments may bring about a new patterning of chromosome segments resulting in heritable phenotypic difference.**

- **2- The chromosome breaking property of chemicals has an important bearing on the chemotherapy of cancer.**

Mode of action vary as follow:

- **1- Some of them affect sulphydryl groups of protein**
- **2- Others act through their influence on hydrogen bonds of nucleic acids.**
- **3- Some agents may affect the oxidation –reduction system within the nucleus.**
- **4- May be due to some specific reaction with RNA.**

Chemicals as;

- **Dyes (methyl blue – orcein) – coumarin - plant pigments – vegetable oils (fats and essential oils) – drugs (antibiotics – DDTinsecticides – sulphur compounds) –bacterial products – alkaloids (caffeine – vincristine) – vitamins – growth regulators – phenols – mustard**

D- Somatic reduction

- **Mode of action:**
- **Prevent replication of centromere in metaphase during mitotic division leading to daughter cells each one contain the half number of chromosomes.**
- **Chemicals:**
- **Sodium nucleate**

المحاضرة الثالثة

Chromosomal variations

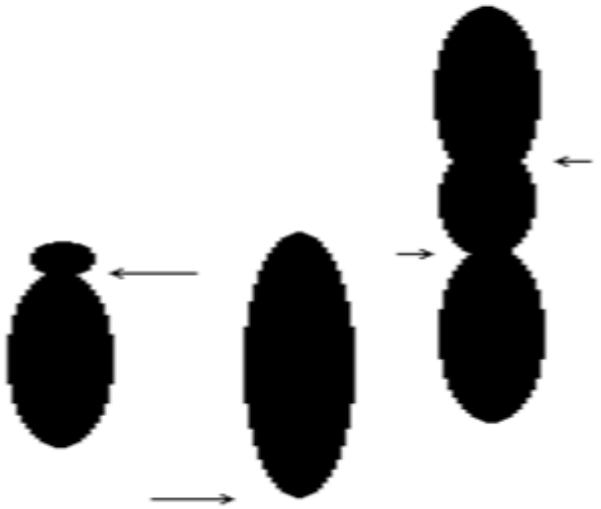
هذا التنوع الذي سوف نتناوله في هذه الباب يشمل أشكال الكروموسومات داخل الخلية في النوع الواحد وبين الخلايا في الأنواع المختلفة وأن الكروموسوم لا يأخذ شكلا واحدا يشبه فيه كل إخوانه بل هناك تباين واسع في الأشكال وفي الأطوال وفي الأحجام وفي الأعداد من نوع لأخر بل حتي في استجابة هذه الكروموسومات للصبغات المختلفة حين يتم صباغتها بهذه الصبغات هذه التنوع الواسع كما سنري جعل الكروموسومات مادة خصبة لدراسة التنوع بين الكائنات المختلفة وقياس درجة التقارب أو التباعد بينها وبين بعضها بل والتعرف علي العينات المجهولة القرابة ونسبتها الي الأجناس والعائلات التي تمثلها فعلا فهيا بنا نعرف أكثر

Types of variation:

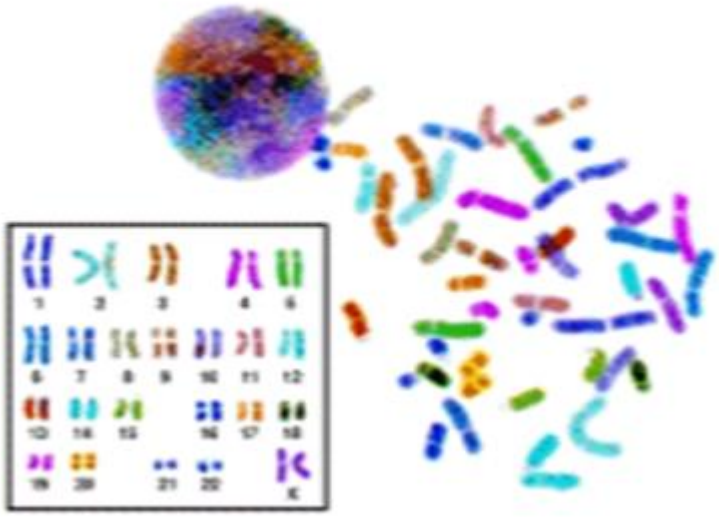
1- Variation in chromosome morphology

- **This variation based on the position of cenromere as follow:**
- **أي أن وضع السنترومير علي الكروموسوم يجعل الكروموسومات جميعها من ناحية الشكل تقع تحت واحد من هذه الأشكال الأربعة التي لا خامس لها وهي : الطرفي وتحت الطرفي والوسطي وتحت الوسطي**
- **a- Telocentric b- Metacentric**
  c- Submetacentric d- Acrocentric

Centromere positions استنتج بنفسك من الصورة في هذه الشريحة والتي بعدها



karyotype



2-Variation based on number

- **The numbers were recorded ranged between: 2*n*=4 in *Haplopappus gracilis* (Asteraceae)**
- **to 2*n*=260 in *Poa litorosa* (Poaceae)**
- **علينا أن ننتبه الي أن أعداد الكروموسومات ثابتة في خلايا النوع الواحد ومتغيرة بين الأنواع وبالتالي كل كائن حي له عدد واحد ثابت يميزه وبالتالي يمكن التعرف عليه في عينه مجهولة من خلال هذا العدد ...الانسان مثلا 46 كروموسوم في كل خلية من خلاياه الجسدية فأينما وجد الأنسان وجد هذا العدد في خلاياه فاذا وجد زيادة أو نقص في حالات معينة تحت الفحص فهذه حالات شاذه بعضها يؤدي الي أمراض وراثية خطيرة تعالوا لنري.....**

Abnormal cases

- Although the number of chromosomes in the species is constant, there is abnormal cases:
- 1- Aneuploidy (+ Chromosome 2n+1 or – chromosome 2n-1) : in this case the normal number of chromosomes in living cell have chromosome excess or less chromosome as follow:

2n +1 or 2n – 1

Examples:

- a- Trisomy 21 (Daween Syndrome) , the excess chromosome follow pair number21 and this case discovered by Daween so named Daween syndrome
- b- Trisomy 18 (Edward Syndrome) , the excess chromosome follow pair number 18 and this case discovered by Edware so named Edward syndrome
- c-Trisomy 13 (Patou Syndrome) , the excess chromosome follow pair number 13 and this case discovered by Patou so named Patou syndrome
-
- The last cases the increasing in somatic chromosomes , in the following cases the increasing or decreasing in sex chromosomes

d- A female with only one X
2n+x instead 2n+xx)
Turner syndrome(called

e- An additional (X) chromosome n a male… resulting in 47 chromosomes…..has Klinefelter Syndrome (2n + XXY)

-

2- Polyploidy: This means chromsomes in the nucleus increase (+ group or more 3n , 4n ,)

3- Monoploidy: In the opposite side chromosomes decrease (- one group n)

3- Variation based on long

- **Chromosome inside cell or between different living organisms vary in their longs and these longs ranged from 0.25 to 30 m (in the different species) .**

4- Variation based on size

The differences between chromosomes in size related to the content of chromatin material inside each chromosome . Chromatin material content vary ; so chromosomal sizes vary too inside one cell and between cells of different living organisms.

5- Variation based on staining (Chromosome bands)

- **The metaphase chromosomes are digested and stained. Heterochromatin regions (gene-poor) (gene-rich) stained more than Euchromatin regions and bands appeared, these bands form number , position and size constant on the one chromosome in same species and vary between different chromosomes and different species , this in case one dye use for compare.**
- **Chromosome bands resulting from stain system type this true.**

Examples:

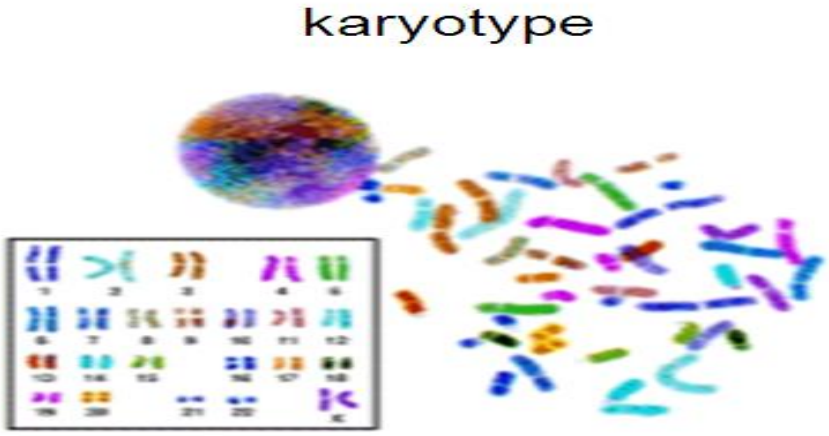
1- G- banding (Giemsa)
2- C- banding (constitutive heterochromatine or centromere)
3- Q- banding (quinacrine)
4- T-banding (Telomeric)
5- R- banding (reverse giemsa)

المحاضرة الرابعة

Applications of chromosomal variations

1- Karyotype

What is karyotype?



The karyotype is the morphological aspect of the chromosome complement as seen at mitotic metaphase.

The principal morphological criteria and their significance

1- Differences in absolute size of the chromosome which probably reflect different amounts of gene duplication between related species or genera

2- Differences in position of the centromere which reflect changes structural changes of chromosomes

3- Differences in relative chromosome size

4- Differences in basic number

5- Differences in the number and position of satellites

6- Differences in degree and distribution of heterochromatic regions

Applications of karyotype

1- Cytological classification (genotaxnomy)

2- Discover syndromes

3- Definition the type of tumor

4- Definition the type of embryo

5- Discover translocation , inversion and deletion of chromosome

6- Duplication of bands

7- Presence or absence of genetic rearrangement •

Chromosomal variation during mitotic cycle

Mitotic cycle divided into:

1- Metabolic stage

In this stage chemical changes tackle place . The principal chemical event is replication of DNA. On the opposite side , DNA of the centromere replicates during metaphase not in this stage. •

2- Mitotic stages

During these successive stages , chromosomes undergo physical changes. The physical changes consist partly of the coiling and folding of the chromonemata (chromonemata is the coiled filament in which the genes are located, which extends the entire length of a chromosome while chromonema is the coiled central thread of a chromatid along which chromomere lies.) , which produces the shorting of the chromosomes from mid-prophase to metaphase , and of their uncoiling and unfolding during telophase. •

…….discuss in detail ?

المحاضرة الخامسة

Linkage and Crossing over

1- Results of Bateson- Punnets on sweet pea crosses (Leads to Linkage)

2- Results of Morgan leads to crossing over which explained linkage

In mendelian genetics this experiment was done as follow:

- نبات طويل أحمر الأزهار مع قصير أبيض الأزهار
- P1: TTRR Homozygous dominant x ttrr
- H. recessive
- G2: TR tr
- F1: TtRr
- طويل Heterozygous dominant
- أحمر الأزهار (أباء)
- يترك ليتلقح ذاتيا

P2: TtRr x TtRr •

G2: TR , Tr , tR, tr •

F2: •

	TR	Tr	tR	tr
TR	TTRR	TTRr	TtRR	TtRr
Tr	TTRr	TTrr	TtRr	Ttrr
tR	TtRR	TtRr	ttRR	ttRr
tr	TtRr	Ttrr	ttRr	ttrr

Phenotypes in F2:

1-Tall and red = 9

2- Tall and white =3

3- Short and red = 3

4- Short and white = 1

This is the rule in mendelian genetics . The strange action that Batsun and Punnet did the same thing but the results were different , what they did?

This experiment was done:

1- Flower color (Purple – red) •

2-Pollen grains shape (long – round) •

F1:

P1: rrpp (Red round) x RRPP (Purple long) •

G: (rp) (RP) •

F1: RrPp (Purple long)

The phenotypes in F2 as follow:

1- Purple long : 11.1 (Expected ratio 9) •
2- Purple round : 0.83 (Expected ratio 3) •
3- Red long : 1.03 (Expected ratio 3) •
4- Red round : 3.08 (Expected ratio 1) •

How and why 9:3:3:1 ratio was not obtained ?

Bateson and Punnet were unable to explain. •

When they did test cross , results were as follow:

1- Purple long : 43.7 •
2- Purple round : 6.3 •
3- Red long : 6.3 •
4- Red round : 43.7 •

While mendele did the same experiment by using other characters ; the results were as follow: •

•

TtRr x ttrr •
G: TR – Tr – tR- tr tr •

F: TtRr - Ttrr - ttRr - ttrr •

•

(phenotypes) 1 : 1 : 1 :1 •

The difference between two experiments large what happened?

1- Genes for flower color and pollen grain shape •

were linked together on the same chromosome ,

this linkage recognized as incomplete linkage. When genes are closely linked on the same chromosome they are transmitted together from parent to progeny this is called complete linkage.

2- Law of independent assortment was not rule in all crosses. •

3- Linkage may be complete (rare) or incomplete •

This is introduced after discovering crossing over by Morgan and others.

Genotype determination

1- Back cross

2- Self cross

Lecturer 6

Theory of chromosomes

Chromosomal Theory of Inheritance was proposed long before there was any direct evidence that traits were carried on chromosomes

1-During meiosis, homologous chromosome pairs migrate as discrete structures that are independent of other chromosome pairs. -1

2-The sorting of chromosomes from each homologous pair into pre-gametes appears to be random. -2

3- Each parent synthesizes gametes that contain only half of their chromosomal complement.

4-Even though male and female gametes (sperm and egg) differ in size and morphology they have the same number of chromosomes, suggesting equal genetic contributions from each parent.

5- The gametic chromosomes combine during fertilization to produce offspring with the same chromosome number as their parents

Through of time ongl

search, evedences practical to pprocha scientests

follow as orythe isth lead to:

1- Protenor
In this insect the number of

Chromosomes in male less than

Female one chromosome as follow:

- **Female have 14 Chromosome (7 + 7)**
- **Male have 13 Chromosome (7 + 6)**

characters eneticg of because ivingl in malef of

seual in speciallye male than more always chracters

rganismso

os in cessex hromosomec het propably from

exuals tabou esp nsibler seng carry femal

characters

2 – *Tennerio moliton (Netti & Steves)*

- **Male and female have the same number of chromosomes.**
- **Chromosomes divided into two types**
- **A- Somatic chromosomes (the same in male and female)**
- **B- Sex chromosomes (2 in male and 2 in female) one of male chromosome similar to the two of female and the other different ……Steves put the following formula .**

Male (XY) Female (XX)

Because of of somatic character in

female and male similar also chromosomes

Similar while characters correlate with

Sex differ , sex chromosomes differ

3- *Drosophila melanogaster*, a fruit fly (Thomas Hunt Morgan)

- **A – Normal fly : Male and female each have 8 chromosomes**

…… 4 pairs , 3 pairs somatic (similar) and 1 pair sex

chromosomes different (rode shape).

Sex chromosomes in male: one of the two rode

with end has simple curve so it differ from the others in male

and female.

Parents 3AA+XX x 3AA+XY •
Gametes (3A+X) (3A+Y) (3A+X) •
•
F1 3AA+XX (Female) 3AA+XY (Male) •

3-b: Wild fly

dlwi the ni hatt oundf Morgan fly,

This ectsni

characterized by red eye and the color red dominant. •
When married tack placed results •
were as follow:

Characterized by red eye when made his •
experiments gains this results:

P XrXr (Wild red eyes) x XY White eyes •

G (Xr) (X) (Y) •

F1 XrX (Wild eyes) XrY (Wild eyes) •
This not happened (50% male : 50% female) if gene •
carried on somatic chromosome
•

F2

P2 XrX x XrY •

G2 (Xr) (X) (Xr) (Y) •

F2 XrXr XrY XXr XY •
3 red eye : 1 white eye •
This is too (25% white male only) •

Reverse

P1 XX (white eye) x XrY (red eye)

G1 (X) (Xr) (Y)

F1 : XrX XY

50% females (red eye) 50% males (white eye)

This is too (50% white eye, male only not female)

F2

P2 XrX x XY

G2 (Xr) (X) (X) (Y)

F2: XrX Xr Y XX XY

This is too (50% white eye, males and females)

that detroper yehT

- **gene responsible for this character carried on X chromosome**
- **this means that traits carrying chromosomes**

Lecturer 7

Mutation and Mutagenesis

Mutation is a sudden heritable change in the genetic material of an organism. Mutations are created

Mainly by external factors , including chemical and physical agents, called mutagens.

Types of mutations

Mutations can be grouped into :

1- Spontaneous mutations (normal)

Mutations can occur spontaneously due to errors in DNA replication , repair and recombination.

a- Somatic mutations

b-Germinal mutations

1- Chromosomal mutations (result of aberrations)

2- Gene mutations

 # Macromutations (major gene)

 # Micromutations (minor gene)

 # Molecular or point mutation

2- Induced mutations (by human)

3- Cytoplasmic mutations (Classed as plasmon and plastidom)
4- Conditional mutations (related to environment)
a- If show phenotypic expression (restrictive conditions)
b- if do not show phenotypic expression (permissive conditions)

Classes of conditional mutations

1- Auxotrophic (restrictive condition)
2- Temprature-senstive
can grow under high and low (temprature)
3- Suppressor-sensetive (need another genetic factor called suppresor

Mutagenic agents

1- Physical mutagens
a- High-energy ionizing radiations (x-rays , gamma , beta and alpha particles , neutrons)
b- Low-energy non-ionizing (ultraviolet light)
2- Chemical mutagens

Chemical mutagens

1- Base analogues and related compounds

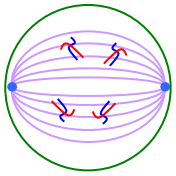
- **(similar to one of the four DNA bases in which can be incorporated into a DNA molecule during replication)**
- **2- Antibiotics (chromosome breaking)**
- **3- Alkylating agents (alkyl group)**
- **4- Azide (carcinogenic**

--الانقسام الخلوي Cell division

* الانقسام الميتوزي (غير مباشر) Mitosis or mitotic division :
- مكان الحدوث : الخلايا الجسمية Somatic cells.
- الهدف : النمو "بزيادة الخلايا" .

- تعتبر المرحلة الوسطية (البينية) Interphase للانقسام الميتوزي لب هذه العمليات .
- لأنه يتم خلالها مضاعفة المادة النووية DNA لنفسها .

Three stages : Gape 1 phase --------Synthetic phase--------Gape 2 phase-

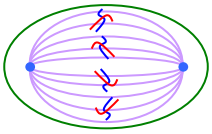


- مراحل الانقسام الميتوزي :

أ – المرحلة التمهيدية Prophase:

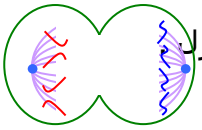
1- انقسام الجسم المركزي إلى نجيمين مع تكوين خيوط المغزل .
2- اختفاء النوية و الغشاء النووي .
3- تفكك الشبكة النووية إلى كروموسومات (2n) و كل كروموسوم عبارة عن كروماتيدين يلتقيان بالسنترومير .

ب – المرحلة الاستوائية Metaphase :



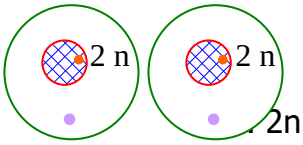
- تنتظم الكروموسومات في وسط الخلية مكونة صفيحة استوائية واحدة .

ج – المرحلة الانفصالية Anaphase:



1- ينقسم السنترومير و يبتعد كل كروماتيد عن قرينه بفعل تقلص خيوط المغزل.
2- تخصر السيتوبلازم .

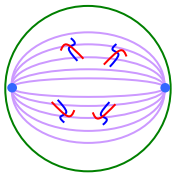
د – المرحلة النهائية Telophase :



1- يتحول كل نجيم إلى جسم مركزي مع اختفاء خيوط المغزل .
2- تظهر النوية و الغشاء النووي .
3- تتشكل الشبكة النووية .
4- تتكون خليتان بهما نفس العدد الكروموسومي للخلية المنقسمة 2n .

- الانقسام الميوزي (الاختزالي) Meiosis or meiotic division:

- مكان الحدوث : الخلايا الجنسية Germ cells.
- الهدف : تكوين و زيادة عدد الأمشاج .

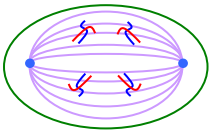


- مراحل الانقسام الميوزي I :

أ – المرحلة التمهيدية Prophase 1:

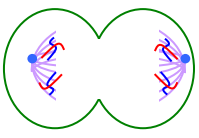
1- انقسام الجسم المركزي إلى نجيمين مع تكوين خيوط المغزل .
2- اختفاء النوية و الغشاء النووي .
3- تفكك الشبكة النووية إلى كروموسومات (2n) و كل كروموسوم عبارة عن كروماتيدين يلتقيان بالسنترومير . ويتم ذلك خلال خمسة مراحل وهي:

ا- اللبتوتين leptotene
ب- الزيجوتين Zygotene تترتب الكروموسومات في أزواج ويلتصق كل كروموسومين متشابهين ببعضهما ويكونان Bivalent وتعرف هذه الظاهرة بالاقتران Synapsis
ج- الباكتين Pachytene
د- الدبلوتين Diplotene وفيه تتكون الوصلات بين الكروماتيدات غير الشقيقة التي تسمي Chiasma ويحدث ما يسمي بالعبور Crossing over وتتوقف نسبة التراكيب الجديدة في النسل علي نسبة العبور وعدد الكيازما التي يمكن عدها
س- التشتتي Diakinesis



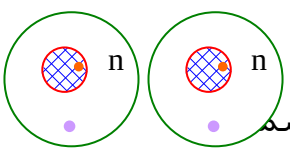
ب – المرحلة الاستوائية Metaphase 1 :

- انتظام الكروموسومات في وسط الخلية لتكوين صفيحتين استوائيتين .



ج – المرحلة الانفصالية Anaphase 1 :

1- ينفصل كل كروموسوم عن قرينه بفعل خيوط المغزل .
2- تختصر السيتوبلازم .

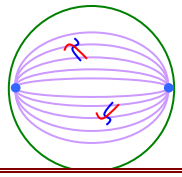


د – المرحلة النهائية Telophase 1 :

1- يتحول كل نجيم إلى جسم مركزي مع اختفاء خيوط المغزل .
2- تظهر النوية و الغشاء النووي .
3- تتشكل الشبكة النووية .
4- يتكون خليتان بكل منهما نصف العدد الكروموسومي للخلية المنقسمة

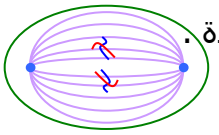
- مراحل الانقسام الميوزي II :

أ – المرحلة التمهيدية Phrophase2:



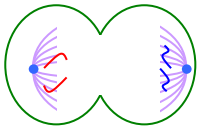
4- انقسام الجسم المركزي إلى نجيمين مع تكوين خيوط المغزل .
5- اختفاء النوية و الغشاء النووي .
6- تفكك الشبكة النووية إلى كروموسومين و كل كروموسوم عبارة عن كروماتيدين يلتقيان بالسنترومير .

ب – المرحلة الاستوائية Metaphase 2 :



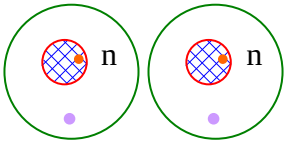
- تنتظم الكروموسومات في وسط الخلية لتكوين صفيحة استوائية واحدة .

ج – المرحلة الانفصالية Anaphase 2 :



3- ينقسم السنترومير و ابتعاد كل كروماتيد عن قرينه .
4- تخصر السيتوبلازم .

د – المرحلة النهائية Telophase 2 :



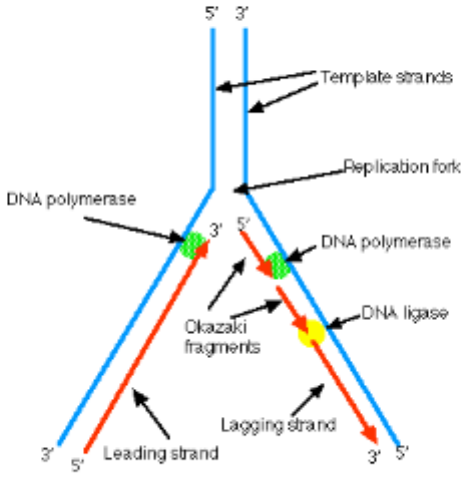
5- يتحول كل نجيم إلى جسم مركزي مع اختفاء خيوط المغزل .
6- تظهر النوية و الغشاء النووي .
7- تتشكل الشبكة النووية .
8- تتكون خليتان بهما نصف العدد الكروموسومي للخلية المنقسمة .

المحاضرة التاسعة

DNA replication

(Semi-conservative method)

Molecular Aspects of the DNA Semi-conservative Replication



Introdution

In molecular biology, DNA replication is the biological process of producing two identical replicas of DNA from one original DNA molecule.[1] DNA replication occurs in all living organisms acting as the most essential part for biological inheritance. This is essential for cell division during growth and repair of damaged tissues, while it also ensures that each of the new cells receives its own copy of the DNA.[2] The cell possesses the distinctive property of division, which makes replication of DNA essential.

DNA is made up of a double helix of two complementary strands. The double helix describes the appearance of a double-stranded DNA which is thus composed of two linear strands that run opposite to each other and twist together to form.[3] During replication, these strands are separated. Each strand of the original DNA molecule then serves as a template for the production of its counterpart, a process referred to as semiconservative replication. As a result of semi-conservative replication, the new helix will be composed of an original DNA strand as well as a newly synthesized strand.[4] Cellular proofreading and error-checking mechanisms ensure near perfect fidelity for DNA replication.[5][6]

In a cell, DNA replication begins at specific locations, or origins of replication, in the genome[7] which contains the genetic material of an organism.[8] Unwinding of DNA at the origin and synthesis of new strands, accommodated by an enzyme known as helicase, results in replication forks growing bi-directionally from the origin. A number of proteins are associated with the replication fork to help in the initiation and continuation of DNA synthesis. Most prominently, DNA polymerase synthesizes the new strands by adding nucleotides that complement each (template) strand. DNA replication occurs during the S-stage of interphase.

DNA replication (DNA amplification) can also be performed *in vitro* (artificially, outside a cell). DNA polymerases isolated from cells and artificial DNA primers can be used to start DNA synthesis at known sequences in a template DNA molecule. Polymerase chain reaction (PCR), ligase chain reaction (LCR), and transcription-mediated amplification (TMA) are examples. In March 2021, researchers reported evidence suggesting that a preliminary form of transfer RNA, a necessary component of translation, the biological synthesis of new proteins in accordance with the genetic code, could

have been a replicator molecule itself in the very early development of life, or abiogenesis.[9][10]

Contents

DNA structure[edit]

DNA exists as a double-stranded structure, with both strands coiled together to form the characteristic double-helix. Each single strand of DNA is a chain of four types of nucleotides. Nucleotides in DNA contain a deoxyribose sugar, a phosphate, and a nucleobase. The four types of nucleotide correspond to the four nucleobases adenine, cytosine, guanine, and thymine,

commonly abbreviated as A, C, G and T. Adenine and guanine are purine bases, while cytosine and thymine are pyrimidines. These nucleotides form phosphodiester bonds, creating the phosphate-deoxyribose backbone of the DNA double helix with the nucleobases pointing inward (i.e., toward the opposing strand). Nucleobases are matched between strands through hydrogen bonds to form base pairs. Adenine pairs with thymine (two hydrogen bonds), and guanine pairs with cytosine (three hydrogen bonds).

DNA strands have a directionality, and the different ends of a single strand are called the "3′ (three-prime) end" and the "5′ (five-prime) end". By convention, if the base sequence of a single strand of DNA is given, the left end of the sequence is the 5′ end, while the right end of the sequence is the 3′ end. The strands of the double helix are anti-parallel with one being 5′ to 3′, and the opposite strand 3′ to 5′. These terms refer to the carbon atom in deoxyribose to which the next phosphate in the chain attaches. Directionality has consequences in DNA synthesis, because DNA polymerase can synthesize DNA in only one direction by adding nucleotides to the 3′ end of a DNA strand.

The pairing of complementary bases in DNA (through hydrogen bonding) means that the information contained within each strand is redundant. Phosphodiester (intra-strand) bonds are stronger than hydrogen (inter-strand) bonds. The actual job of the phosphodiester bonds is where in DNA polymers connect the 5' carbon atom of one nucleotide to the 3' carbon atom of another nucleotide, while the hydrogen bonds stabilize DNA double helices across the helix axis but not in the direction of the axis 1.[11] This allows the strands to be separated from one another. The nucleotides on a single strand can therefore be used to reconstruct nucleotides on a newly synthesized partner strand.[12]

DNA polymerase[edit]

Main article: DNA polymerase

DNA polymerases adds nucleotides to the 3′ end of a strand of DNA.[13] If a mismatch is accidentally incorporated, the polymerase is inhibited from further extension. Proofreading removes the mismatched nucleotide and extension continues.

DNA polymerases are a family of enzymes that carry out all forms of DNA replication.[14] DNA polymerases in general cannot initiate synthesis of new strands, but can only extend an existing DNA or RNA strand paired with a template strand. To begin synthesis, a short fragment of RNA, called a primer, must be created and paired with the template DNA strand.

DNA polymerase adds a new strand of DNA by extending the 3′ end of an existing nucleotide chain, adding new nucleotides matched to the template strand one at a time via the creation of phosphodiester bonds. The energy

for this process of DNA polymerization comes from hydrolysis of the high-energy phosphate (phosphoanhydride) bonds between the three phosphates attached to each unincorporated base. Free bases with their attached phosphate groups are called nucleotides; in particular, bases with three attached phosphate groups are called nucleoside triphosphates. When a nucleotide is being added to a growing DNA strand, the formation of a phosphodiester bond between the proximal phosphate of the nucleotide to the growing chain is accompanied by hydrolysis of a high-energy phosphate bond with release of the two distal phosphate groups as a pyrophosphate. Enzymatic hydrolysis of the resulting pyrophosphate into inorganic phosphate consumes a second high-energy phosphate bond and renders the reaction effectively irreversible.[Note 1]

In general, DNA polymerases are highly accurate, with an intrinsic error rate of less than one mistake for every 10^7 nucleotides added.[15] In addition, some DNA polymerases also have proofreading ability; they can remove nucleotides from the end of a growing strand in order to correct mismatched bases. Finally, post-replication mismatch repair mechanisms monitor the DNA for errors, being capable of distinguishing mismatches in the newly synthesized DNA strand from the original strand sequence. Together, these three discrimination steps enable replication fidelity of less than one mistake for every 10^9 nucleotides added.[15]

The rate of DNA replication in a living cell was first measured as the rate of phage T4 DNA elongation in phage-infected *E. coli*.[16] During the period of exponential DNA increase at 37 °C, the rate was 749 nucleotides per second. The mutation rate per base pair per replication during phage T4 DNA synthesis is 1.7 per 10^8.[17]

Replication process[edit]

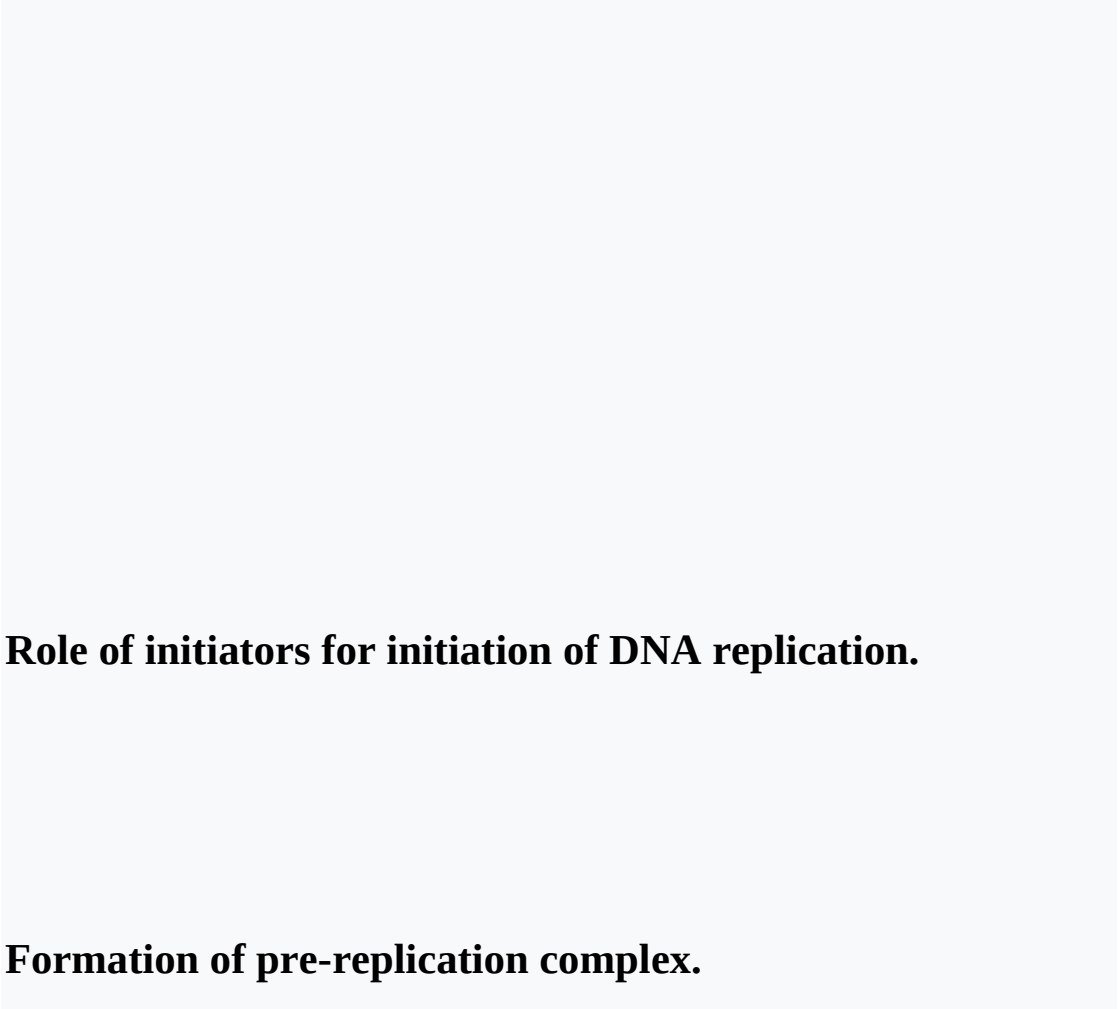
*Main articles: **Prokaryotic DNA replication** and **Eukaryotic DNA replication***

Overview of the steps in DNA replication

Steps in DNA synthesis

DNA replication, like all biological polymerization processes, proceeds in three enzymatically catalyzed and coordinated steps: initiation, elongation and termination.

Initiation[edit]



Role of initiators for initiation of DNA replication.

Formation of pre-replication complex.

For a cell to divide, it must first replicate its DNA.[18] DNA replication is an all-or-none process; once replication begins, it proceeds to completion. Once replication is complete, it does not occur again in the same cell cycle. This is made possible by the division of initiation of the pre-replication complex.

Pre-replication complex[edit]

Main article: Pre-replication complex

In late mitosis and early G1 phase, a large complex of initiator proteins assembles into the pre-replication complex at particular points in the DNA, known as "origins".[7] In *E. coli* the primary initiator protein is DnaA; in yeast, this is

the origin recognition complex.[19] Sequences used by initiator proteins tend to be "AT-rich" (rich in adenine and thymine bases), because A-T base pairs have two hydrogen bonds (rather than the three formed in a C-G pair) and thus are easier to strand-separate.[20] In eukaryotes, the origin recognition complex catalyzes the assembly of initiator proteins into the pre-replication complex. Cdc6 and Cdt1 then associate with the bound origin recognition complex at the origin in order to form a larger complex necessary to load the Mcm complex onto the DNA. The Mcm complex is the helicase that will unravel the DNA helix at the replication origins and replication forks in eukaryotes. The Mcm complex is recruited at late G1 phase and loaded by the ORC-Cdc6-Cdt1 complex onto the DNA via ATP-dependent protein remodeling. The loading of the Mcm complex onto the origin DNA marks the completion of pre-replication complex formation.[21]

If environmental conditions are right in late G1 phase, the G1 and G1/S cyclin-Cdk complexes are activated, which stimulate expression of genes that encode components of the DNA synthetic machinery. G1/S-Cdk activation also promotes the expression and activation of S-Cdk complexes, which may play a role in activating replication origins depending on species and cell type. Control of these Cdks vary depending cell type and stage of development. This regulation is best understood in budding yeast, where the S cyclins Clb5 and Clb6 are primarily responsible for DNA replication.[22] Clb5,6-Cdk1 complexes directly trigger the activation of replication origins and are therefore required throughout S phase to directly activate each origin.[21]

In a similar manner, Cdc7 is also required through S phase to activate replication origins. Cdc7 is not active throughout the cell cycle, and its activation is strictly timed to avoid premature initiation of DNA replication. In late G1, Cdc7 activity rises abruptly as a result of association with the regulatory subunit Dbf4, which binds Cdc7 directly and promotes its protein kinase activity. Cdc7 has been found to

be a rate-limiting regulator of origin activity. Together, the G1/S-Cdks and/or S-Cdks and Cdc7 collaborate to directly activate the replication origins, leading to initiation of DNA synthesis.[21]

Preinitiation complex[edit]

In early S phase, S-Cdk and Cdc7 activation lead to the assembly of the preinitiation complex, a massive protein complex formed at the origin. Formation of the preinitiation complex displaces Cdc6 and Cdt1 from the origin replication complex, inactivating and disassembling the pre-replication complex. Loading the preinitiation complex onto the origin activates the Mcm helicase, causing unwinding of the DNA helix. The preinitiation complex also loads α-primase and other DNA polymerases onto the DNA.[21]

After α-primase synthesizes the first primers, the primer-template junctions interact with the clamp loader, which loads the sliding clamp onto the DNA to begin DNA synthesis. The components of the preinitiation complex remain associated with replication forks as they move out from the origin.[21]

Elongation[edit]

DNA polymerase has 5′–3′ activity. All known DNA replication systems require a free 3′ hydroxyl group before synthesis can be initiated (note: the DNA template is read in 3′ to 5′ direction whereas a new strand is synthesized in the 5′ to 3′ direction—this is often confused). Four distinct mechanisms for DNA synthesis are recognized:

1. **All cellular life forms and many DNA viruses, phages and plasmids use a primase to synthesize a short RNA primer with a free 3′ OH group which is subsequently elongated by a DNA polymerase.**
2. **The retroelements (including retroviruses) employ a transfer RNA that primes DNA replication by**

providing a free 3′ OH that is used for elongation by the reverse transcriptase.

3. In the adenoviruses and the φ29 family of bacteriophages, the 3′ OH group is provided by the side chain of an amino acid of the genome attached protein (the terminal protein) to which nucleotides are added by the DNA polymerase to form a new strand.
4. In the single stranded DNA viruses—a group that includes the circoviruses, the geminiviruses, the parvoviruses and others—and also the many phages and plasmids that use the rolling circle replication (RCR) mechanism, the RCR endonuclease creates a nick in the genome strand (single stranded viruses) or one of the DNA strands (plasmids). The 5′ end of the nicked strand is transferred to a tyrosine residue on the nuclease and the free 3′ OH group is then used by the DNA polymerase to synthesize the new strand.

The first is the best known of these mechanisms and is used by the cellular organisms. In this mechanism, once the two strands are separated, primase adds RNA primers to the template strands. The leading strand receives one RNA primer while the lagging strand receives several. The leading strand is continuously extended from the primer by a DNA polymerase with high processivity, while the lagging strand is extended discontinuously from each primer forming Okazaki fragments. RNase removes the primer RNA fragments, and a low processivity DNA polymerase distinct from the replicative polymerase enters to fill the gaps. When this is complete, a single nick on the leading strand and several nicks on the lagging strand can be found. Ligase works to fill these nicks in, thus completing the newly replicated DNA molecule.

The primase used in this process differs significantly between bacteria and archaea/eukaryotes. Bacteria use a primase belonging to the DnaG protein superfamily which contains a catalytic domain of the TOPRIM fold type.[23] The

TOPRIM fold contains an α/β core with four conserved strands in a Rossmann-like topology. This structure is also found in the catalytic domains of topoisomerase Ia, topoisomerase II, the OLD-family nucleases and DNA repair proteins related to the RecR protein.

The primase used by archaea and eukaryotes, in contrast, contains a highly derived version of the RNA recognition motif (RRM). This primase is structurally similar to many viral RNA-dependent RNA polymerases, reverse transcriptases, cyclic nucleotide generating cyclases and DNA polymerases of the A/B/Y families that are involved in DNA replication and repair. In eukaryotic replication, the primase forms a complex with Pol α.[24]

Multiple DNA polymerases take on different roles in the DNA replication process. In *E. coli*, DNA Pol III is the polymerase enzyme primarily responsible for DNA replication. It assembles into a replication complex at the replication fork that exhibits extremely high processivity, remaining intact for the entire replication cycle. In contrast, DNA Pol I is the enzyme responsible for replacing RNA primers with DNA. DNA Pol I has a 5′ to 3′ exonuclease activity in addition to its polymerase activity, and uses its exonuclease activity to degrade the RNA primers ahead of it as it extends the DNA strand behind it, in a process called nick translation. Pol I is much less processive than Pol III because its primary function in DNA replication is to create many short DNA regions rather than a few very long regions.

In eukaryotes, the low-processivity enzyme, Pol α, helps to initiate replication because it forms a complex with primase.[25] In eukaryotes, leading strand synthesis is thought to be conducted by Pol ε; however, this view has recently been challenged, suggesting a role for Pol δ.[26] Primer removal is completed Pol δ[27] while repair of DNA during replication is completed by Pol ε.

As DNA synthesis continues, the original DNA strands continue to unwind on each side of the bubble, forming a replication fork with two prongs. In bacteria, which have a single origin of replication on their circular chromosome, this process creates a "theta structure" (resembling the Greek letter theta: θ). In contrast, eukaryotes have longer linear chromosomes and initiate replication at multiple origins within these.[28]

Replication fork[edit]

Scheme of the replication fork.
a: template, b: leading strand, c: lagging strand, d: replication fork, e: primer, f: Okazaki fragments

Many enzymes are involved in the DNA replication fork.

The replication fork is a structure that forms within the long helical DNA during DNA replication. It is created by helicases, which break the hydrogen bonds holding the two DNA strands together in the helix. The resulting structure has two branching "prongs", each one made up of a single strand of DNA. These two strands serve as the template for the leading and lagging strands, which will be created as DNA polymerase matches complementary nucleotides to the templates; the templates may be properly referred to as the leading strand template and the lagging strand template.

DNA is read by DNA polymerase in the 3′ to 5′ direction, meaning the new strand is synthesized in the 5' to 3' direction. Since the leading and lagging strand templates are oriented in opposite directions at the replication fork, a major issue is how to achieve synthesis of new lagging strand DNA, whose direction of synthesis is opposite to the direction of the growing replication fork.

Leading strand[edit]

The leading strand is the strand of new DNA which is synthesized in the same direction as the growing replication fork. This sort of DNA replication is continuous.

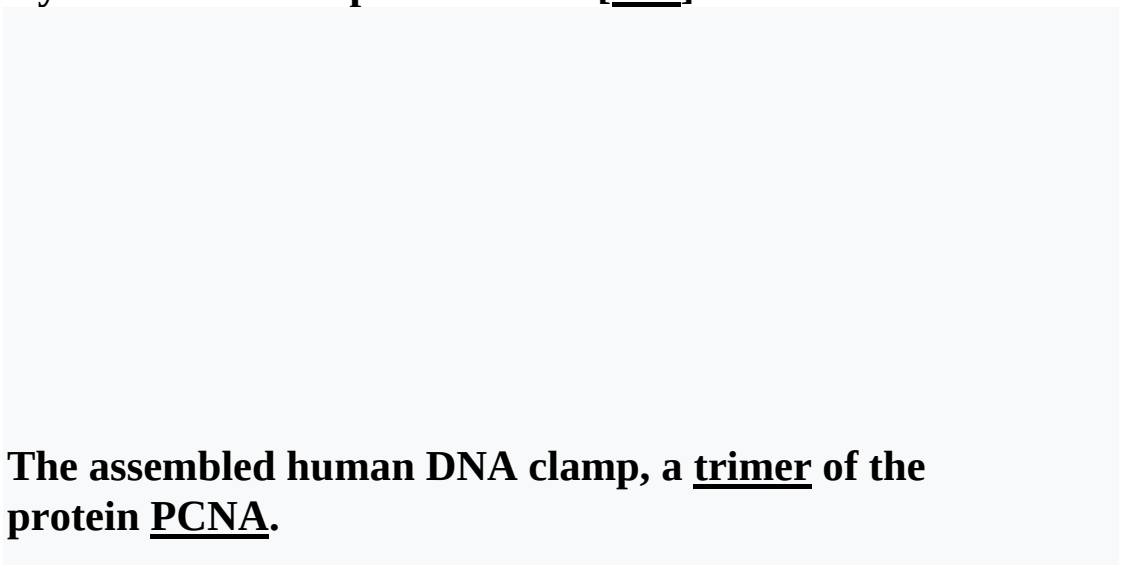
Lagging strand[edit]

The lagging strand is the strand of new DNA whose direction of synthesis is opposite to the direction of the growing replication fork. Because of its orientation, replication of the lagging strand is more complicated as compared to that of the leading strand. As a consequence, the DNA polymerase on this strand is seen to "lag behind" the other strand.

The lagging strand is synthesized in short, separated segments. On the lagging strand *template*, a primase "reads" the template DNA and initiates synthesis of a short complementary RNA primer. A DNA polymerase extends the primed segments, forming Okazaki fragments. The RNA

primers are then removed and replaced with DNA, and the fragments of DNA are joined by DNA ligase.

Dynamics at the replication fork[edit]



The assembled human DNA clamp, a trimer of the protein PCNA.

In all cases the helicase is composed of six polypeptides that wrap around only one strand of the DNA being replicated. The two polymerases are bound to the helicase heximer. In eukaryotes the helicase wraps around the leading strand, and in prokaryotes it wraps around the lagging strand.[29]

As helicase unwinds DNA at the replication fork, the DNA ahead is forced to rotate. This process results in a build-up of twists in the DNA ahead.[30] This build-up forms a torsional resistance that would eventually halt the progress of the replication fork. Topoisomerases are enzymes that temporarily break the strands of DNA, relieving the tension caused by unwinding the two strands of the DNA helix; topoisomerases (including DNA gyrase) achieve this by adding negative supercoils to the DNA helix.[31]

Bare single-stranded DNA tends to fold back on itself forming secondary structures; these structures can interfere with the movement of DNA polymerase. To prevent this, single-strand binding proteins bind to the DNA until a second strand is synthesized, preventing secondary structure formation.[32]

Double-stranded DNA is coiled around histones that play an important role in regulating gene expression so the

replicated DNA must be coiled around histones at the same places as the original DNA. To ensure this, histone chaperones disassemble the chromatin before it is replicated and replace the histones in the correct place. Some steps in this reassembly are somewhat speculative.[33]

Clamp proteins form a sliding clamp around DNA, helping the DNA polymerase maintain contact with its template, thereby assisting with processivity. The inner face of the clamp enables DNA to be threaded through it. Once the polymerase reaches the end of the template or detects double-stranded DNA, the sliding clamp undergoes a conformational change that releases the DNA polymerase. Clamp-loading proteins are used to initially load the clamp, recognizing the junction between template and RNA primers.[6]:274-5

DNA replication proteins[edit]

At the replication fork, many replication enzymes assemble on the DNA into a complex molecular machine called the replisome. The following is a list of major DNA replication enzymes that participate in the replisome:[34]

Enzyme	Function in DNA replication
DNA helicase	Also known as helix destabilizing enzyme. Helicase separates the two strands of DNA at the Replication Fork behind the topoisomerase.
DNA polymerase	The enzyme responsible for catalyzing the addition of nucleotide substrates to DNA in the 5′ to 3′ direction during DNA replication. Also performs proof-reading and error correction. There exist many different types of DNA Polymerase, each of which perform different functions in

	different types of cells.
DNA clamp	**A protein which prevents elongating DNA polymerases from dissociating from the DNA parent strand.**
Single-strand DNA-binding protein	**Bind to ssDNA and prevent the DNA double helix from re-annealing after DNA helicase unwinds it, thus maintaining the strand separation, and facilitating the synthesis of the new strand.**
Topoisomerase	**Relaxes the DNA from its super-coiled nature.**
DNA gyrase	**Relieves strain of unwinding by DNA helicase; this is a specific type of topoisomerase**
DNA ligase	**Re-anneals the semi-conservative strands and joins Okazaki Fragments of the lagging strand.**
Primase	**Provides a starting point of RNA (or DNA) for DNA polymerase to begin synthesis of the new DNA strand.**
Telomerase	**Lengthens telomeric DNA by adding repetitive nucleotide sequences to the ends of eukaryotic chromosomes. This allows germ cells and stem cells to avoid**

	the Hayflick limit on cell division.

E. coli Replisome. Notably, the DNA on lagging strand forms a loop. The exact structure of replisome is not well understood.

Replication machineries consist of factors involved in DNA replication and appearing on template ssDNAs. Replication machineries include primosotors are replication enzymes; DNA polymerase, DNA helicases, DNA clamps and DNA topoisomerases, and replication proteins; e.g. single-stranded DNA binding proteins (SSB). In the replication machineries these components coordinate. In most of the bacteria, all of the factors involved in DNA replication are located on replication forks and the complexes stay on the forks during DNA replication. These replication machineries are called replisomes or DNA replicase systems. These terms are generic terms for proteins located on replication forks. In eukaryotic and some bacterial cells the replisomes are not formed.

Since replication machineries do not move relatively to template DNAs such as factories, they are called a replication factory.[36] In an alternative figure, DNA factories are similar to projectors and DNAs are like as cinematic films passing constantly into the projectors. In the replication factory model, after both DNA helicases for

leading strands and lagging strands are loaded on the template DNAs, the helicases run along the DNAs into each other. The helicases remain associated for the remainder of replication process. Peter Meister et al. observed directly replication sites in budding yeast by monitoring green fluorescent protein (GFP)-tagged DNA polymerases α. They detected DNA replication of pairs of the tagged loci spaced apart symmetrically from a replication origin and found that the distance between the pairs decreased markedly by time.[37] This finding suggests that the mechanism of DNA replication goes with DNA factories. That is, couples of replication factories are loaded on replication origins and the factories associated with each other. Also, template DNAs move into the factories, which bring extrusion of the template ssDNAs and new DNAs. Meister's finding is the first direct evidence of replication factory model. Subsequent research has shown that DNA helicases form dimers in many eukaryotic cells and bacterial replication machineries stay in single intranuclear location during DNA synthesis.[36]

The replication factories perform disentanglement of sister chromatids. The disentanglement is essential for distributing the chromatids into daughter cells after DNA replication. Because sister chromatids after DNA replication hold each other by Cohesin rings, there is the only chance for the disentanglement in DNA replication. Fixing of replication machineries as replication factories can improve the success rate of DNA replication. If replication forks move freely in chromosomes, catenation of nuclei is aggravated and impedes mitotic segregation.

Termination[edit]

Eukaryotes initiate DNA replication at multiple points in the chromosome, so replication forks meet and terminate at many points in the chromosome. Because eukaryotes have linear chromosomes, DNA replication is unable to reach the very end of the chromosomes. Due to this problem, DNA is

lost in each replication cycle from the end of the chromosome. Telomeres are regions of repetitive DNA close to the ends and help prevent loss of genes due to this shortening. Shortening of the telomeres is a normal process in somatic cells. This shortens the telomeres of the daughter DNA chromosome. As a result, cells can only divide a certain number of times before the DNA loss prevents further division. (This is known as the Hayflick limit.) Within the germ cell line, which passes DNA to the next generation, telomerase extends the repetitive sequences of the telomere region to prevent degradation. Telomerase can become mistakenly active in somatic cells, sometimes leading to cancer formation. Increased telomerase activity is one of the hallmarks of cancer.

Termination requires that the progress of the DNA replication fork must stop or be blocked. Termination at a specific locus, when it occurs, involves the interaction between two components: (1) a termination site sequence in the DNA, and (2) a protein which binds to this sequence to physically stop DNA replication. In various bacterial species, this is named the DNA replication terminus site-binding protein, or Ter protein.

Because bacteria have circular chromosomes, termination of replication occurs when the two replication forks meet each other on the opposite end of the parental chromosome. *E. coli* regulates this process through the use of termination sequences that, when bound by the Tus protein, enable only one direction of replication fork to pass through. As a result, the replication forks are constrained to always meet within the termination region of the chromosome.[38]

Regulation[edit]

Main articles: Cell division and Cell cycle

The cell cycle of eukaryotic cells.

Eukaryotes[edit]

Within eukaryotes, DNA replication is controlled within the context of the cell cycle. As the cell grows and divides, it progresses through stages in the cell cycle; DNA replication takes place during the S phase (synthesis phase). The progress of the eukaryotic cell through the cycle is controlled by cell cycle checkpoints. Progression through checkpoints is controlled through complex interactions between various proteins, including cyclins and cyclin-dependent kinases.[39] Unlike bacteria, eukaryotic DNA replicates in the confines of the nucleus.[40]

The G1/S checkpoint (or restriction checkpoint) regulates whether eukaryotic cells enter the process of DNA replication and subsequent division. Cells that do not proceed through this checkpoint remain in the G0 stage and do not replicate their DNA.

After passing through the G1/S checkpoint, DNA must be replicated only once in each cell cycle. When the Mcm complex moves away from the origin, the pre-replication complex is dismantled. Because a new Mcm complex cannot be loaded at an origin until the pre-replication subunits are reactivated, one origin of replication can not be used twice in the same cell cycle.

Activation of S-Cdks in early S phase promotes the destruction or inhibition of individual pre-replication complex components, preventing immediate reassembly. S and M-Cdks continue to block pre-replication complex assembly even after S phase is complete, ensuring that

assembly cannot occur again until all Cdk activity is reduced in late mitosis.

In budding yeast, inhibition of assembly is caused by Cdk-dependent phosphorylation of pre-replication complex components. At the onset of S phase, phosphorylation of Cdc6 by Cdk1 causes the binding of Cdc6 to the SCF ubiquitin protein ligase, which causes proteolytic destruction of Cdc6. Cdk-dependent phosphorylation of Mcm proteins promotes their export out of the nucleus along with Cdt1 during S phase, preventing the loading of new Mcm complexes at origins during a single cell cycle. Cdk phosphorylation of the origin replication complex also inhibits pre-replication complex assembly. The individual presence of any of these three mechanisms is sufficient to inhibit pre-replication complex assembly. However, mutations of all three proteins in the same cell does trigger reinitiation at many origins of replication within one cell cycle.

In animal cells, the protein geminin is a key inhibitor of pre-replication complex assembly. Geminin binds Cdt1, preventing its binding to the origin recognition complex. In G1, levels of geminin are kept low by the APC, which ubiquitinates geminin to target it for degradation. When geminin is destroyed, Cdt1 is released, allowing it to function in pre-replication complex assembly. At the end of G1, the APC is inactivated, allowing geminin to accumulate and bind Cdt1.

Replication of chloroplast and mitochondrial genomes occurs independently of the cell cycle, through the process of D-loop replication.

Replication focus[edit]

In vertebrate cells, replication sites concentrate into positions called replication foci.[37] Replication sites can be detected by immunostaining daughter strands and replication enzymes and monitoring GFP-tagged replication factors. By these methods it is found that replication foci of

varying size and positions appear in S phase of cell division and their number per nucleus is far smaller than the number of genomic replication forks.

P. Heun *et al*., (2001) tracked GFP-tagged replication foci in budding yeast cells and revealed that replication origins move constantly in G1 and S phase and the dynamics decreased significantly in S phase.[37] Traditionally, replication sites were fixed on spatial structure of chromosomes by nuclear matrix or lamins. The Heun's results denied the traditional concepts, budding yeasts do not have lamins, and support that replication origins self-assemble and form replication foci.

By firing of replication origins, controlled spatially and temporally, the formation of replication foci is regulated. D. A. Jackson et al.(1998) revealed that neighboring origins fire simultaneously in mammalian cells. Spatial juxtaposition of replication sites brings clustering of replication forks. The clustering do rescue of stalled replication forks and favors normal progress of replication forks. Progress of replication forks is inhibited by many factors; collision with proteins or with complexes binding strongly on DNA, deficiency of dNTPs, nicks on template DNAs and so on. If replication forks stall and the remaining sequences from the stalled forks are not replicated, the daughter strands have nick obtained un-replicated sites. The un-replicated sites on one parent's strand hold the other strand together but not daughter strands. Therefore, the resulting sister chromatids cannot separate from each other and cannot divide into 2 daughter cells. When neighboring origins fire and a fork from one origin is stalled, fork from other origin access on an opposite direction of the stalled fork and duplicate the un-replicated sites. As other mechanism of the rescue there is application of dormant replication origins that excess origins do not fire in normal DNA replication.

Dam methylates adenine of GATC sites after replication.

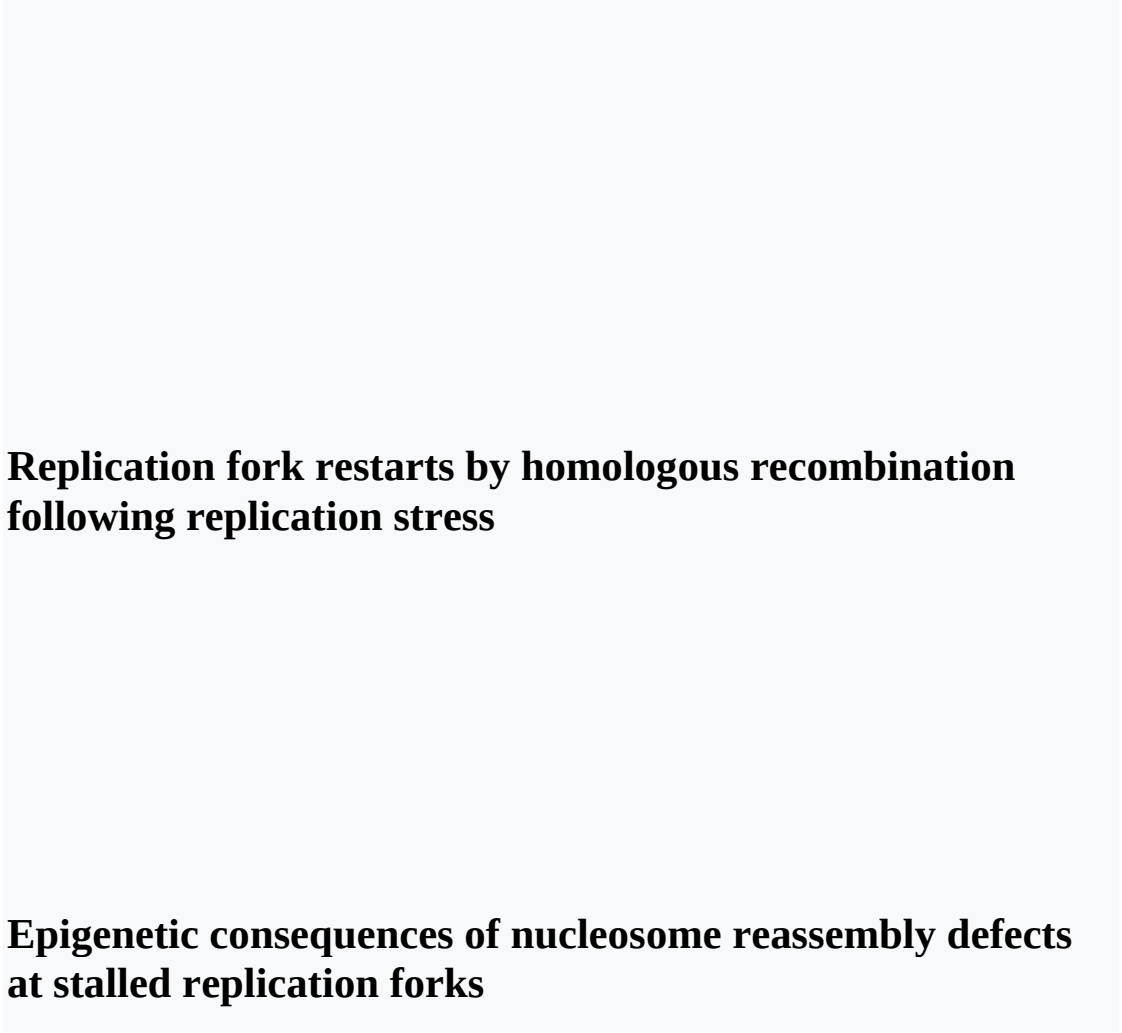
Most bacteria do not go through a well-defined cell cycle but instead continuously copy their DNA; during rapid growth, this can result in the concurrent occurrence of multiple rounds of replication. In *E. coli*, the best-characterized bacteria, DNA replication is regulated through several mechanisms, including: the hemimethylation and sequestering of the origin sequence, the ratio of adenosine triphosphate (ATP) to adenosine diphosphate (ADP), and the levels of protein DnaA. All these control the binding of initiator proteins to the origin sequences.

Because *E. coli* methylates GATC DNA sequences, DNA synthesis results in hemimethylated sequences. This hemimethylated DNA is recognized by the protein SeqA, which binds and sequesters the origin sequence; in addition, DnaA (required for initiation of replication) binds less well to hemimethylated DNA. As a result, newly replicated origins are prevented from immediately initiating another round of DNA replication.

ATP builds up when the cell is in a rich medium, triggering DNA replication once the cell has reached a specific size. ATP competes with ADP to bind to DnaA, and the DnaA-

ATP complex is able to initiate replication. A certain number of DnaA proteins are also required for DNA replication — each time the origin is copied, the number of binding sites for DnaA doubles, requiring the synthesis of more DnaA to enable another initiation of replication.

In fast-growing bacteria, such as *E. coli*, chromosome replication takes more time than dividing the cell. The bacteria solve this by initiating a new round of replication before the previous one has been terminated. The new round of replication will form the chromosome of the cell that is born two generations after the dividing cell. This mechanism creates overlapping replication cycles.



Replication fork restarts by homologous recombination following replication stress

Epigenetic consequences of nucleosome reassembly defects at stalled replication forks

There are many events that contribute to replication stress, including:

- **Misincorporation of ribonucleotides**

- **Unusual DNA structures**
- **Conflicts between replication and transcription**
- **Insufficiency of essential replication factors**
- **Common fragile sites**
- **Overexpression or constitutive activation of oncogenes**
- **Chromatin inaccessibility**

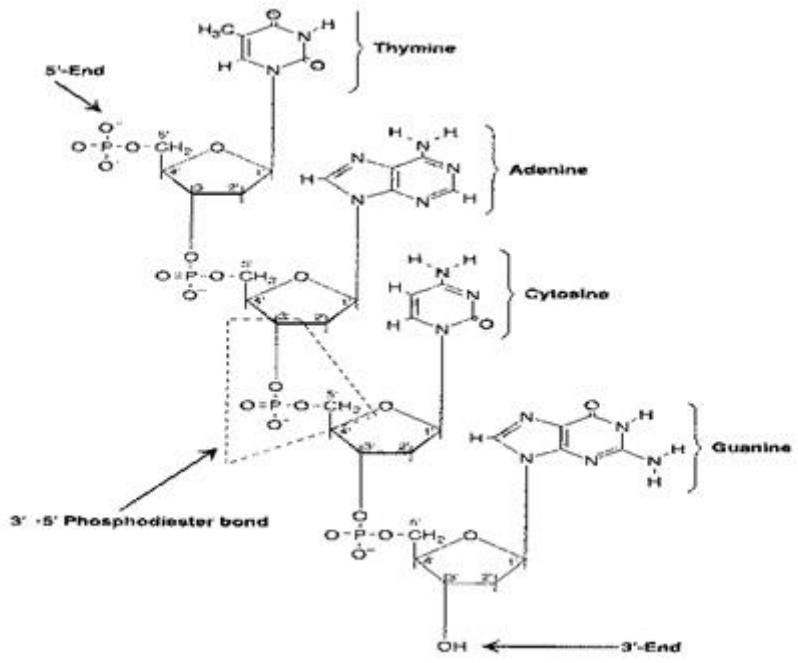
Researchers commonly replicate DNA *in vitro* using the polymerase chain reaction (PCR). PCR uses a pair of primers to span a target region in template DNA, and then polymerizes partner strands in each direction from these primers using a thermostable DNA polymerase. Repeating this process through multiple cycles amplifies the targeted DNA region. At the start of each cycle, the mixture of template and primers is heated, separating the newly synthesized molecule and template. Then, as the mixture cools, both of these become templates for annealing of new primers, and the polymerase extends from these. As a result, the number of copies of the target region doubles each round, increasing exponentially.

المحاصرة العاشرة

DNA

سوف نتناول بالدراسة الموضوعات التالية وهي متعلقة بالحامض النووي

- ١- تركيبه الجزيئ وانواع الروابط الموجوده به
- ٢- الأدلة العملية كونه مادة الوراثة فعلا
- ٣- صوره المختلفة والفروق بينها بعدما كان يعتقد أن له صورة واحدة فقط
- ٤- بعض الصفات التي يتميز بها دون غيره من المركبات العضوية والتي توفر له الثبات حتي عند درجة الغليان
- ٥- عملية تضاعف الحامض النووي تلقائيا داخل الخلية وكيف تتم ولولاها لما نتج نسل ولبقي الأباء بدون أبناء ولفشلت عملية الوراثة برمتها



Molecular structure

A- Nucleotide (DNA composed of repeated units of nucleotides connected together by phosphodiester bonds)

Nucleotide = sugar +base + phosphate group

B- Nucleoside= nucleotide without phosphate

C- Bonds (hydrogen bonds ,dihydrogen bonds and tri – glycosidic bonds – phosphodiester bonds)

D- Rule of Charagaff (1949 – 1953) found that the content of A=T and also C=G

DNA

1- Molecular structure
2- Genetic matter
3- Different forms of DNA
4- Abnormal characters of DNA
   a- Stability of Tautameric form
   b- Denaturation and Annealing
5- Replication
6- Cutting
7- Restriction fragments
8- Vectors
9- Electrophoresis

Molecular structure

- A- Nucleotide (DNA composed of repeated units of nucleotides connected together by phosphodiester bonds)
- Nucleotide = sugar +base + phosphate group
- B- Nucleoside= nucleotide without phosphate
- C- Bonds (hydrogen bonds ,dihydrogen bonds and tri – glycosidic bonds – phosphodiester bonds)
- D- Rule of Charagaff (1949 – 1953) found that the content of A=T and also C=G

Different forms of DNA

1- A
2- B
3- Z

Comparison

Z	B	A	Character
Very rare	distributed	rare	Range of distribution
Oblong and narrow	Long and narrow	Short and wide	General form
3.8A	3.32A	2.3A	High between base pairs
Left	Wright	Wright	Direction of helix
12	10	١١	No. of bases/ helix cycle

Comparison

Z	B	A	Character
60	35.9	33.6	Turn (mean) /base pair

Follow

Z	B	A	Character
-9	-1.2	+19.0	Disposition degree of bases for axis
Large groove	Through base pairs	Large groove	Site of helix axis
Splay on helix surface	Wide and mid Deeping	Very narrow and very deep	Percentages of large groove
Very narrow very deep	Narrow and mid Deeping	Very wide and shallow	Percentages of small groove

علينا أن نعلم ان موضوع DNA لم يعد قاصرا علي علماء البيولوجي وانما دخلت في حساباته الرياضيات والفيزياء بثقلهما مما نتج عنه هذا الكم الهائل من القياسات في ما يسمي حسابات DNA التي كانت تبدوا مستحيلة لعلماء البيولوجي كما رأيتم بوضوح في جدول المقارنات

Different characters of DNA

Stability of Tautameric form

In nitrogen bases ; hydrogen connected with nitrogen or oxygen mainly in the form of NH2 (Amino form) not NH (Imino form) except in rare cases and this case called Tautamerically stable. Also oxygen connected with carbon six in Thymine and Guanine is in the form C= O (keto form) not Enol form (COH).

Denaturation and Annealing

DNA separate by heat(100 c) and coil by decreasing heat

G+C / A+T….. this percentage is directly proportional to heat responsible for separate double helix in laboratory

هذه بعض الخواص التي يتمتع بها الحامض النووي
دون غيره من المركبات الحيوية كما ذكرنا من قبل

المحاضرة الحادية عشرة

Restriction Enzymes

INTRODUCTION

Restriction endonucleases popularly referred to as restriction enzymes, are ubiquitously present in prokaryotes. The function of restriction endonucleases is mainly protection against foreign genetic material especially against bacteriophage DNA. The other functions attributed to these enzymes are recombination and transposition. Restriction endonucleases make up the restriction-modification (R-M) systems comprised of endonuclease and methytransferase activities. The endonuclease recognizes and cleaves foreign DNA on the defined recognition sites. The methyltransferase modifies the recognition sites in the host DNA and protects it against the activity of endonucleases. The sequences in foreign DNA are generally not methylated and are subjected to restriction digestion. Each restriction enzyme recognizes a specific sequence of 4–8 nucleotides in DNA and cleaves at these sites. Endonucleases isolated by different organisms with identical recognition sites are termed isoschizomers.

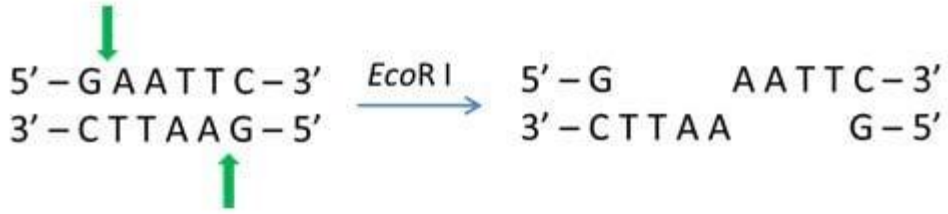
NOMENCLATURE

Different bacterial species synthesize endonucleases depending on the infecting viral DNA. The guidelines for naming restriction enzymes are based on the original suggestion by Smith and Nathans.[1] The enzyme names begin with an italicized three-letter acronym; the first letter of the acronym is the first letter of the genus of bacteria from which the enzyme was isolated, the next two letters are the two letters of the species. These are followed by extra letters or numbers to indicate the serotype or strain, a space, then a Roman numeral to indicate the chronology of identification. For example, the first endonuclease isolated from *Escherichia coli*, strain RY13 is named as *Eco*R I. *Hin*d III is the third endonuclease of four isolated from *Haemophilus influenza*, serotype d.

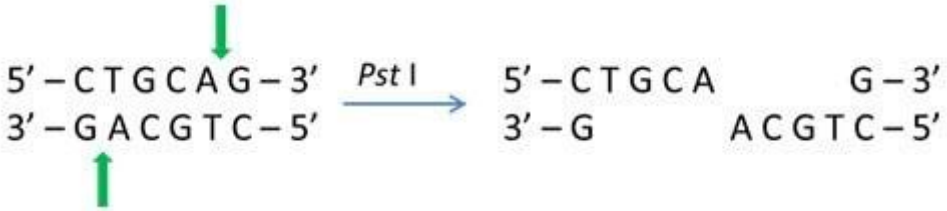
PRODUCTS OF RESTRICTION DIGESTION

The digestion product of DNA strands may result in a fragment with either blunt ends or cohesive (sticky) ends at both the ends.

Cohesive 5' ends generated by *Eco*R I (Catalog Number R6265):



Cohesive 3' ends generated by *Pst* I (Catalog Number R7023):



Blunt ends generated by *Hae* III (Catalog Number R5628):

METHYLATION OF DNA STRANDS

Methylation of DNA is a major factor that affects the specificity of restriction endonucleases. DNA isolated from strains of bacteria expressing methylases such as Dam or Dcm may be resistant to cleavage by endonucleases as the recognition sites are methylated. Plasmid DNA isolated from *E. coli* expressing Dam is methylated at the GATC sites making it resistant to cleavage by *Mbo* I. Certain methylation-specific endonucleases also degrade methylated DNA without affecting the methylated host DNA. *Bam*H I isolated from *Bacillus amyloliquefaciens* H (Catalog Number R0260) cleaves methylated GGATCC sites of plasmid from *E. coli* expressing Dam. *Dpn* I (Catalog Number R8381) isolated from *Diplococcus pneumonia* targets methylated T7 DNA.

APPLICATIONS OF RESTRICTION ENZYMES

Restriction endonucleases are widely used in molecular biology research for the following applications:

Genetic Engineering: The most popular application of restriction endonucleases is as a tool for genetic engineering. The endonuclease activity enables manipulation of the genome as well as introduction of sequences of interest in the host organism. This results in the production of the desired gene product by the host. This concept has wide range of applications in biotechnology in the production of antibiotics, antibodies, enzymes, and several secondary metabolites.

DNA mapping: DNA mapping using restriction enzymes (also known as restriction mapping) is a method to obtain structural information of the DNA fragment. In this technique the DNA is digested with a series of restriction enzymes to produce DNA fragments of various sizes. The resultant fragments are separated by agarose gel electrophoresis and the distance between the restriction enzyme sites can be estimated. This can be used to determine the structure of an unknown DNA fragment.

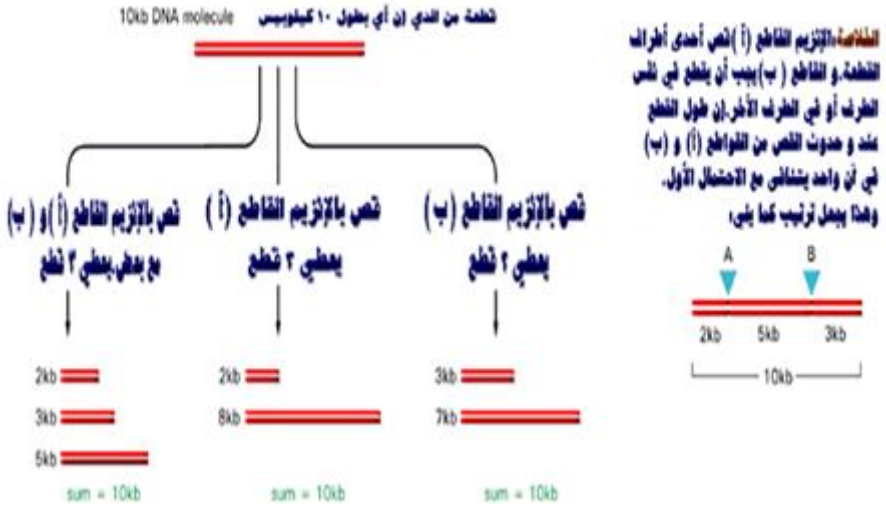
Gene Sequencing: A large DNA molecule is digested using restriction enzymes and the resulting fragments are processed through DNA sequencer to obtain the nucleotide sequence.

The other applications of restriction endonucleases include gene expression and mutation studies and examination of population polymorphisms.

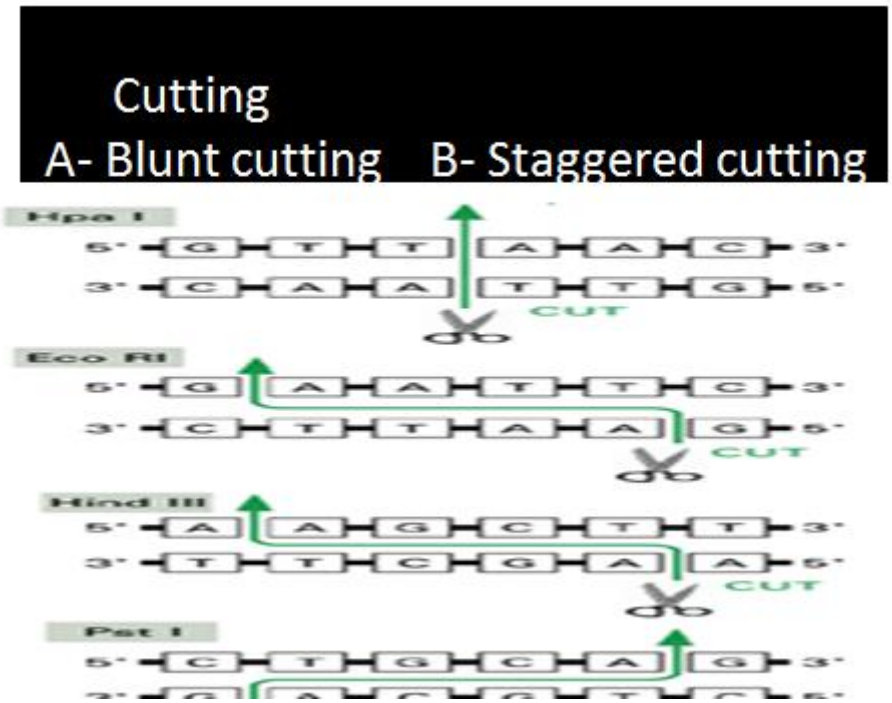
هذه Cutting enzymes (Restriction nucleases) produce هي البداية restriction fragments

If mutation take placed , Restriction Fragment Length Polymorphism produced (RFLP)

Restriction fragments



Maps of restriction fragments



Methylation

وجود انزيمات القطع(at A or C) add methyl group
المتعرج واستخدامها أدي الي سهولة انتاج قطع من DNAتقبل
اللصق بسهولة مما يسهل عملية الحصول علي قطع هجينة من
DNA.....أما عملية وجود مجموعة مثيل في القاعدة
النيتروجينية فهو نظام بيولوجي يوفر الحماية للخلايا التي
تحتوي انزيمات القطع حيث يبقي الحامض النووي الخاص بها
محصنا ضد القطع فاذا نزعت مجموعة المثيل قطع

Recombinant DNA = two fragments of two different DNA

Hpa I = GTTAAC

Eco RI = GAATTC

Mechanism of cutting depends on certain sequence for each enzyme as above

Nomenclature

Hpa 1 = Hemophilus parainfluenzae

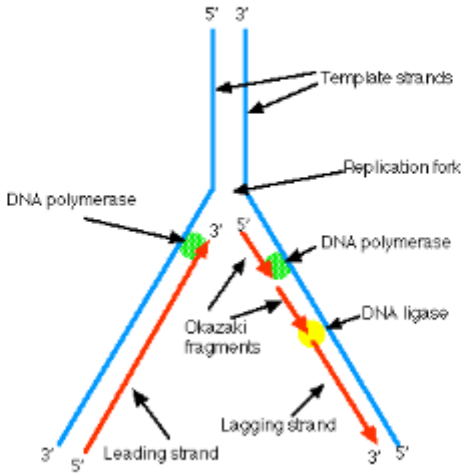
Eco R1 = Escherichia coli

Replication

A- Conservative method

B- Semi-conservative method

Molecular Aspects of the DNA Semi-conservative Replication



المحاضرة الثانية عشرة

Genetic matter

- 1- Genetic transformation (Avery 1944)
- A- S-type (Smooth)
- B- R-type (Rough)
- 2- Genetic transduction (Hershy and Chase 1952)

For many years, genetic information was thought to be contained in cell protein. Continuing the research done by Frederick Griffith, Avery worked with Colin MacLeod and Maclyn McCarty on the mystery of inheritance. He had received emeritus status from the Rockefeller Institute in 1943, but continued working for five years, though by that time he was in his late sixties. Techniques were available to remove various organic compounds from bacteria, and if the remaining organic compounds were still able to cause R strain bacteria to transform then the substances removed could not be the carrier of genes. S-bacteria first had the large cellular structures removed. Then they were treated

with protease enzymes, which removed the proteins from the cells before the remainder was placed with R strain bacteria. The R strain bacteria transformed, meaning that proteins did not carry the genes causing disease. Then the remnants of the S strain bacteria were treated with a deoxyribonuclease enzyme which removed the DNA. After this treatment, the R strain bacteria no longer transformed. This showed that DNA was the substance that transformed R strain into S strain bacteria and indicated that it was the carrier of genes in cells.

Hershey and Chase needed to be able to examine different parts of the phages they were studying separately, so they needed to distinguish the phage subsections. Viruses were known to be composed of a protein shell and DNA, so they chose to uniquely label each with a different elemental isotope. This allowed each to be observed and analyzed separately. Since phosphorus is contained in DNA but not amino acids, radioactive phosphorus-32 was used to label the DNA contained in the T2 phage. Radioactive sulfur-35 was used to label the protein sections of the T2 phage, because sulfur is contained in protein but not DNA.

Hershey and Chase inserted the radioactive elements in the bacteriophages by adding the isotopes to separate media within which bacteria were allowed to grow for 4 hours before bacteriophage introduction. When the bacteriophages infected the bacteria, the progeny contained the radioactive isotopes in their structures. This procedure was performed once for the sulfur-labeled phages and once for phosphorus-labeled phages. The labeled progeny were then allowed to infect unlabeled bacteria. The phage coats remained on the outside of the bacteria, while genetic material entered. Disruption of phage from the bacteria by agitation in a blender followed by centrifugation allowed for the

separation of the phage coats from the bacteria. These bacteria were lysed to release phage progeny. The progeny of the phages that were labeled with radioactive phosphorus remained labeled, whereas the progeny of the phages labeled with radioactive sulfur were unlabeled. Thus, the Hershey–Chase experiment helped to confirm that DNA, not protein, is the genetic material. Hershey and Chase showed that the introduction of deoxyribonuclease (referred to as DNase), an enzyme that breaks down DNA, into a solution containing the labeled bacteriophages did not introduce any ^{32}P into the solution. This demonstrated that the phage is resistant to the enzyme while intact. Additionally, they were able to plasmolyze the bacteriophages so that they went into osmotic shock, which effectively created a solution containing most of the ^{32}P and a heavier solution containing structures called "ghosts" that contained the ^{35}S and the protein coat of the virus. It was found that these "ghosts" could adsorb to bacteria that were susceptible to T2, although they contained no DNA and were simply the remains of the original viral capsule. They concluded that the protein protected the DNA from DNase, but that once the two were separated and the phage was inactivated, the DNase could hydrolyze the phage DNA.

Hershey and Chase were also able to prove that the DNA from the phage is inserted into the bacteria shortly after the virus attaches to its host. Using a high-speed blender they were able to force the bacteriophages from the bacterial cells after adsorption. The lack of ^{32}P-labeled DNA remaining in the solution after the bacteriophages had been allowed to adsorb to the bacteria showed that the phage DNA was transferred into the bacterial cell. The presence of almost all the radioactive ^{35}S in the solution showed that the protein coat that protects the DNA before adsorption stayed outside the cell.[1]

Hershey and Chase concluded that DNA, not protein, was the genetic material. They determined that a protective protein coat was formed around the bacteriophage, but that

the internal DNA is what conferred its ability to produce progeny inside a bacterium. They showed that, in growth, protein has no function, while DNA has some function. They determined this from the amount of radioactive material remaining outside of the cell. Only 20% of the ^{32}P remained outside the cell, demonstrating that it was incorporated with DNA in the cell's genetic material. All of the ^{35}S in the protein coats remained outside the cell, showing it was not incorporated into the cell, and that protein is not the genetic material.

Hershey and Chase's experiment concluded that little sulfur-containing material entered the bacterial cell. However no specific conclusions can be made regarding whether material that is sulfur-free enters the bacterial cell after phage adsorption. Further research was necessary to conclude that it was solely bacteriophages' DNA that entered the cell and not a combination of protein and DNA where the protein did not contain any sulfur.

المحاضرة الثالثة عشرة

Vectors

- 1- Plasmids (Yeast – E. coli)
- a- DNA b- RNA
- 2- Viral Vectors (Phages e.g. lambda phage)
- 3- Cosmides (phage + plasmide)
- Plasmide + Cos sites = Cosmide (Cos = cohesive sequence)

4- Artificial Chromosome vectors

- 1- Yeast Artificial Chromosomes / YAC (more than 500 kb)
- 2- Bacterial Artificial Chromosomes / BAC (150 kb)

Types of plasmids

- **YIP**, yeast integrating plasmid = selectable marker +cloning sites
- **YRP**, yeast replicating plasmid = YIP + ARS origin of replication
- **YEP**, yeast expression plasmid = YIP + 2 micron origin of replication
- **YCP**, yeast centromere plasmid = YRP + centromere sequence

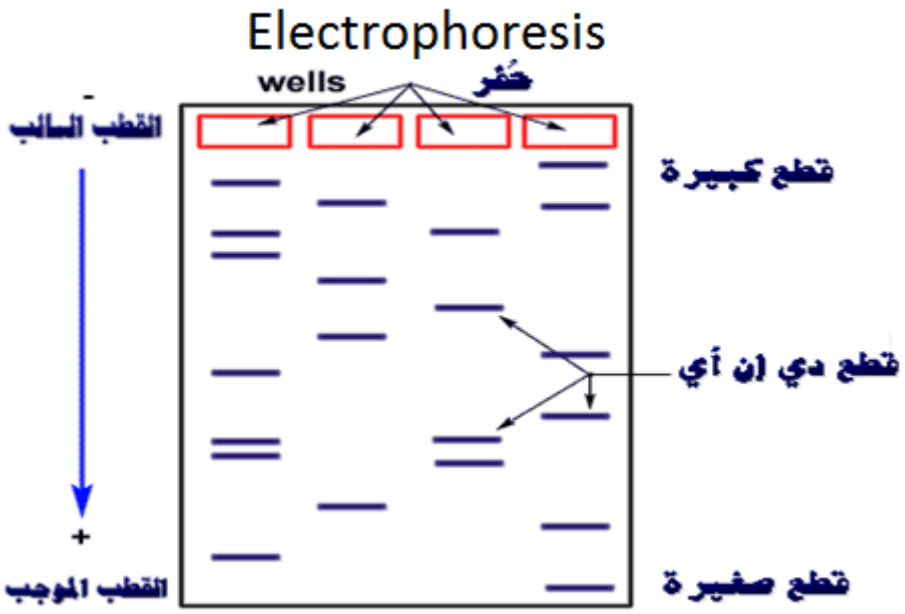
Electrophoresis

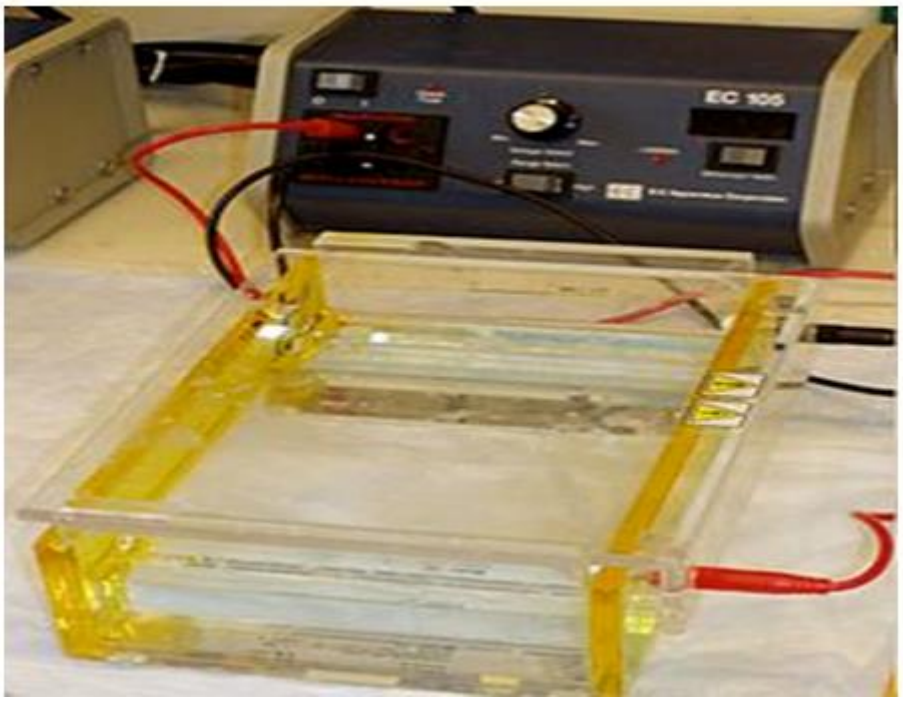
A- Gel

B- Polyacrylamide

عملية الفصل الكهربي لقطع الحامض النووي ما كانت لتتم لولم تكن جزيئات الحامض مشحونة بشحنة كهربية

- مادة الجل أثبتت فعاليتها العالية كمادة يتم عليها الفصل لكن بمرور الوقت بدأت القطع الصغيرة من الحامض النووي تجد صعوبة في حركتها علي لوح الجل واتضح أن السبب هوكبر المسافة بين جزيئات الجل مما يؤدي الي سقوط هذه القطع في هذه الفراغات مما يعيق حركتها لذلك في هذه الحالات يتم استبدال مادة الجل بمادة البولي اكريلميد ذات المسافات الجزيئية الصغيرة . مع الوقت بدأت القطع الكبيرة نسبيا تتأرجح أثناء حركتها علي الجل مما يجعلها تخرج عن المسار المستقيم لحركة القطع وأمكن اعادتها الي الوضع الطبيعي باستخدام ومضات من تيار متردد عند اللزوم بجانب التيار المستمر المستخدم أصلا


Electrophoresis
wells
حُفَر
-
القطب السالب
+
القطب الموجب
قطع كبيرة
قطع دي إن أي
قطع صغيرة


EC 105

المحاضة الثالثة عشرة

Vectors

In molecular cloning, a vector is any particle (e.g., plasmids, cosmids, Lambda phages) used as a vehicle to artificially carry a foreign nucleic sequence – usually DNA – into another cell, where it can be replicated and/or expressed.[1] A vector containing foreign DNA is termed recombinant DNA. The four major types of vectors are plasmids, viral vectors, cosmids, and artificial chromosomes. Of these, the most commonly used vectors are plasmids.[2] Common to all engineered vectors have an origin of replication, a multicloning site, and a selectable marker.

The vector itself generally carries a DNA sequence that consists of an insert (in this case the transgene) and a larger sequence that serves as the "backbone" of the vector. The purpose of a vector which transfers genetic information to another cell is typically to isolate, multiply, or express the insert in the target cell. All vectors may be used for cloning and are therefore cloning vectors, but there are also vectors designed specially for cloning, while others may be designed specifically for other purposes, such as transcription and protein expression. Vectors designed specifically for the expression of the transgene in the target cell are called expression vectors, and generally have a promoter sequence that drives expression of the transgene. Simpler vectors called transcription vectors are only capable of being transcribed but not translated: they can be replicated in a target cell but not expressed, unlike expression vectors. Transcription vectors are used to amplify their insert.

The manipulation of DNA is normally conducted on *E. coli* vectors, which contain elements necessary for their maintenance in *E. coli*. However, vectors may also have elements that allow them to be maintained in another organism such as yeast, plant or mammalian

cells, and these vectors are called shuttle vectors. Such vectors have bacterial or viral elements which may be transferred to the non-bacterial host organism, however other vectors termed intragenic vectors have also been developed to avoid the transfer of any genetic material from an alien species.[3]

Insertion of a vector into the target cell is usually called transformation for bacterial cells,[4] transfection for eukaryotic cells, although insertion of a viral vector is often called transduction.

Plasmids

Plasmids are double-stranded extra chromosomal and generally circular DNA sequences that are capable of replication using the host cell's replication machinery.[7] Plasmid vectors minimalistically consist of an origin of replication that allows for semi-independent replication of the plasmid in the host. Plasmids are found widely in many bacteria, for example in *Escherichia coli*, but may also be found in a few eukaryotes, for example in yeast such as *Saccharomyces cerevisiae*.[8] Bacterial plasmids may be conjugative/transmissible and non-conjugative:

- **conjugative - mediate DNA transfer through conjugation and therefore spread rapidly among the bacterial cells of a population; e.g., F plasmid, many R and some col plasmids.**
- **nonconjugative - do not mediate DNA through conjugation, e.g., many R and col plasmids.**

The pBR322 plasmid is one of the first plasmids widely used as a cloning vector.

Plasmids with specially-constructed features are commonly used in laboratory for cloning purposes. These plasmid are generally non-conjugative but may have many more features, notably a "multiple cloning site" where multiple restriction enzyme cleavage sites allow for the insertion of a transgene insert. The bacteria containing the plasmids can generate millions of copies of the vector within the bacteria in hours, and the amplified vectors can be extracted from the bacteria for further manipulation. Plasmids may be used specifically as transcription vectors and such plasmids may lack crucial sequences for protein expression. Plasmids used for protein expression, called expression vectors, would include elements for translation of protein, such as a ribosome binding site, start and stop codons.

Viral vectors

Viral vectors are generally genetically engineered viruses carrying modified viral DNA or RNA that has been rendered noninfectious, but still contain viral promoters and also the transgene, thus allowing for translation of the transgene through a viral promoter. However, because viral vectors frequently are lacking infectious sequences, they require helper viruses or packaging lines for large-scale transfection. Viral vectors are often designed for permanent incorporation of the insert into the host genome, and thus leave distinct genetic markers in the host genome after incorporating the transgene. For example, retroviruses leave a

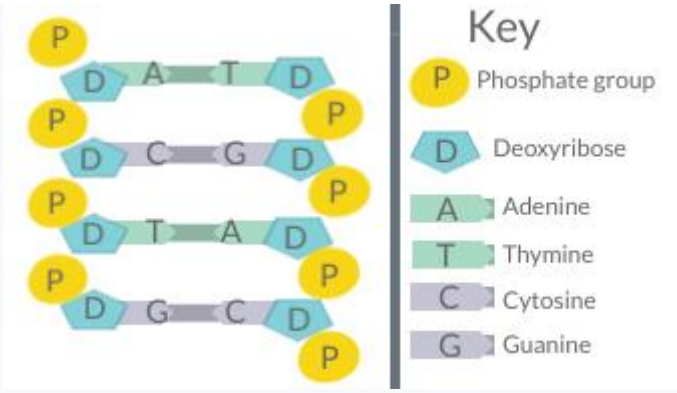
characteristic retroviral integration pattern after insertion that is detectable and indicates that the viral vector has incorporated into the host genome.

Artificial chromosomes

Artificial chromosomes are manufactured chromosomes in the context of yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), or human artificial chromosomes (HACs). An artificial chromosome can carry a much larger DNA fragment than other vectors.[9] YACs and BACs can carry a DNA fragment up to 300,000 nucleotides long. Three structural necessities of an artificial chromosome include an origin of replication, a centromere, and telomeric end sequences.

المحاضرة الرابعة عشرة

Chargaff's rules



A diagram of DNA base pairing, demonstrating the basis for Chargaff's rules.

Chargaff's rules state that DNA from any species of any organism should have a 1:1 stoichiometric ratio of purine and pyrimidine bases (i.e., A+G=T+C) and, more specifically, that the amount of guanine should be equal to cytosine and the amount of adenine should be equal to thymine. This pattern is found in both strands of the DNA. They were discovered by Austrian-born chemist Erwin Chargaff, in the late 1940s.

Definitions

First parity rule

The first rule holds that a double-stranded DNA molecule, *globally* has percentage base pair equality: A% = T% and G% = C%. The rigorous validation of the rule constitutes the basis of Watson-Crick pairs in the DNA double helix model.

Second parity rule

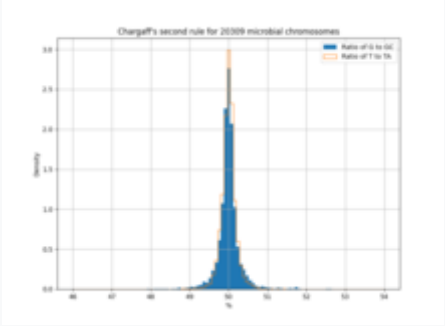
The second rule holds that both A%≈ T% and G% ≈ C% are valid for each of the two DNA strands.[3] This describes only a global feature of the base composition in a single DNA strand.[4]

Research

The second parity rule was discovered in 1968.[3] It states that, in single-stranded DNA, the number of adenine units is *approximately* equal to that of thymine (%A ≈ %T), and the number of cytosine units is *approximately* equal to that of guanine (%C ≈ %G).

The first empirical generalization of Chargaff's second parity rule, called the Symmetry Principle, was proposed by Vinayakumar V. Prabhu [5] in 1993. This principle states that for any given oligonucleotide, its frequency is approximately equal to the frequency of its complementary reverse oligonucleotide. A theoretical generalization[6] was mathematically derived by Michel E. B. Yamagishi and Roberto H. Herai in 2011.

In 2006, it was shown that this rule applies to four[2] of the five types of double stranded genomes; specifically it applies to the eukaryotic chromosomes, the bacterial chromosomes, the double stranded DNA viral genomes, and the archaeal chromosomes.[8] It does not apply to organellar genomes (mitochondria and plastids) smaller than ~20-30 kbp, nor does it apply to single stranded DNA (viral) genomes or any type of RNA genome. The basis for this rule is still under investigation, although genome size may play a role.



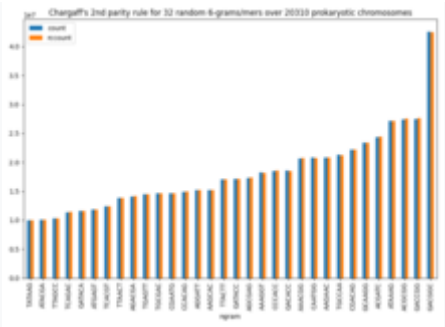
Histogram showing how 20309 chromosomes adhere to Chargaff's second parity rule

The rule itself has consequences. In most bacterial genomes (which are generally 80-90% coding) genes are arranged in such a fashion that approximately 50% of the coding sequence lies on either strand. Wacław Szybalski, in the 1960s, showed that in bacteriophage coding sequences purines (A and G)

exceed pyrimidines (C and T).[9] This rule has since been confirmed in other organisms and should probably be now termed "Szybalski's rule". While Szybalski's rule generally holds, exceptions are known to exist.[10][11][12] The biological basis for Szybalski's rule, like Chargaff's, is not yet known.

The combined effect of Chargaff's second rule and Szybalski's rule can be seen in bacterial genomes where the coding sequences are not equally distributed. The genetic code has 64 codons of which 3 function as termination codons: there are only 20 amino acids normally present in proteins. (There are two uncommon amino acids—selenocysteine and pyrrolysine—found in a limited number of proteins and encoded by the stop codons—TGA and TAG respectively.) The mismatch between the number of codons and amino acids allows several codons to code for a single amino acid - such codons normally differ only at the third codon base position.

Multivariate statistical analysis of codon use within genomes with unequal quantities of coding sequences on the two strands has shown that codon use in the third position depends on the strand on which the gene is located. This seems likely to be the result of Szybalski's and Chargaff's rules. Because of the asymmetry in pyrimidine and purine use in coding sequences, the strand with the greater coding content will tend to have the greater number of purine bases (Szybalski's rule). Because the number of purine bases will, to a very good approximation, equal the number of their complementary pyrimidines within the same strand and, because the coding sequences occupy 80-90% of the strand, there appears to be (1) a selective pressure on the third base to minimize the number of purine bases in the strand with the greater coding content; and (2) that this pressure is proportional to the mismatch in the length of the coding sequences between the two strands.



Chargaff's 2nd parity rule for prokaryotic 6-mers

The origin of the deviation from Chargaff's rule in the organelles has been suggested to be a consequence of the mechanism of replication.[13] During replication the DNA strands separate. In single stranded DNA, cytosine spontaneously slowly deaminates to adenosine (a C to A transversion). The longer the strands are separated the greater the quantity of deamination. For reasons that are not yet clear the strands tend to exist longer in single form in mitochondria than in chromosomal DNA. This process tends to yield one strand that is enriched in guanine (G) and thymine (T) with its complement enriched in cytosine (C) and adenosine (A), and this process may have given rise to the deviations found in the mitochondria.[*citation needed*][*dubious* – *discuss*]

Chargaff's second rule appears to be the consequence of a more complex parity rule: within a single strand of DNA any oligonucleotide (k-mer or n-gram; length ≤ 10) is present in equal numbers to its reverse complementary nucleotide. Because of the computational requirements this has not been verified in all genomes for all oligonucleotides. It has been verified for triplet oligonucleotides for a large data set.[14] Albrecht-Buehler has suggested that this rule is the consequence of genomes evolving by a process of inversion and transposition.[14] This process does not appear to have acted on the mitochondrial genomes. Chargaff's second parity rule appears to be extended from the nucleotide-level to populations of codon triplets, in the case of whole single-stranded Human genome DNA.[15] A kind of "codon-level second Chargaff's parity rule" is proposed as follows:

Intra-strand relation among percentages of codon populations			
First codon	Second codon	Relation proposed	Details
Twx (1st base position is T)	yzA (3rd base position is A)	% Twx % yzA	Twx and yzA are mirror codons, e.g. TCG and CGA
Cwx (1st base position is C)	yzG (3rd base position is G)	% Cwx % yzG	Cwx and yzG are mirror codons, e.g. CTA and TAG
wTx (2nd base position is T)	yAz (2nd base position is A)	% wTx % yAz	wTx and yAz are mirror codons, e.g. CTG and CAG
wCx (2nd base position is C)	yGz (2nd base position is G)	% wCx % yGz	wCx and yGz are mirror codons, e.g. TCT and AGA
wxT (3rd base position is T)	Ayz (1st base position is A)	% wxT % Ayz	wxT and Ayz are mirror codons, e.g. CTT and AAG
wxC (3rd	Gyz (1st	% wxC % Gyz	wxC and Gyz are

Intra-strand relation among percentages of codon populations			
First codon	Second codon	Relation proposed	Details
base position is C)	base position is G)		mirror codons, e.g. GGC and GCC

Examples — computing whole human genome using the first codons reading frame provides:

36530115 TTT and 36381293 AAA (ratio % = 1.00409).
2087242 TCG and 2085226 CGA (ratio % = 1.00096), etc...

In 2020, it is suggested that the physical properties of the dsDNA (double stranded DNA) and the tendency to maximum entropy of all the physical systems are the cause of Chargaff's second parity rule.[16] The symmetries and patterns present in the dsDNA sequences can emerge from the physical peculiarities of the dsDNA molecule and the maximum entropy principle alone, rather than from biological or environmental evolutionary pressure.

Percentages of bases in DNA

The following table is a representative sample of Erwin Chargaff's 1952 data, listing the base composition of DNA from various organisms and support both of Chargaff's rules.[17] An organism such as φX174 with significant variation from A/T and G/C equal to one, is indicative of single stranded DNA.

Organism	Taxon	% A	% G	% C	% T	A / T	G / C	% GC	% AT

Organism	Taxon	%A	%G	%C	%T	A/T	G/C	%GC	%AT
Maize	*Zea*	26.8	22.8	23.2	27.2	0.99	0.98	46.1	54.0
Octopus	*Octopus*	33.2	17.6	17.6	31.6	1.05	1.00	35.2	64.8
Chicken	*Gallus*	28.0	22.0	21.6	28.4	0.99	1.02	43.7	56.4
Rat	*Rattus*	28.6	21.4	20.5	28.4	1.01	1.00	42.9	57.0
Human	*Homo*	29.3	20.7	20.0	30.0	0.98	1.04	40.7	59.3
Grasshopper	Orthoptera	29.3	20.5	20.7	29.3	1.00	0.99	41.2	58.6
Sea urchin	Echinoidea	32.8	17.7	17.3	32.1	1.02	1.02	35.0	64.9
Wheat	*Triticum*	27.3	22.7	22.8	27.1	1.01	1.00	45.5	54.4
Yeast	*Saccharomyces*	31.3	18.7	17.1	32.9	0.95	1.09	35.8	64.4

Organism	Taxon	%A	%G	%C	%T	A/T	G/C	%GC	%AT
E. coli	*Escherichia*	24.7	26.0	25.7	23.6	1.05	1.01	51.7	48.3
φX174	*PhiX174*	24.0	23.3	21.5	31.2	0.77	1.08	44.8	55.2

المحاضرة الخامسة عشرة

Population genetics

Population genetics is a subfield of genetics that deals with genetic differences within and between populations, and is a part of evolutionary biology. Studies in this branch of biology examine such phenomena as adaptation, speciation, and population structure.

Population genetics was a vital ingredient in the emergence of the modern evolutionary synthesis. Its primary founders were Sewall Wright, J. B. S. Haldane and Ronald Fisher, who also laid the foundations for the related discipline of quantitative genetics. Traditionally a highly mathematical discipline, modern population genetics encompasses theoretical, laboratory, and field work. Population genetic models are used both for statistical inference from DNA sequence data and for proof/disproof of concept.

What sets population genetics apart from newer, more phenotypic approaches to modelling evolution, such as evolutionary game theory and adaptive dynamics, is its emphasis on such genetic phenomena as dominance, epistasis, the degree to which genetic recombination breaks linkage disequilibrium, and the random phenomena of mutation and genetic drift. This makes it appropriate for comparison to population genomics data.

History

Population genetics began as a reconciliation of Mendelian inheritance and biostatistics models. Natural selection will only cause evolution if there is enough genetic variation in a population. Before the discovery of Mendelian genetics, one

common hypothesis was blending inheritance. But with blending inheritance, genetic variance would be rapidly lost, making evolution by natural or sexual selection implausible. The Hardy–Weinberg principle provides the solution to how variation is maintained in a population with Mendelian inheritance. According to this principle, the frequencies of alleles (variations in a gene) will remain constant in the absence of selection, mutation, migration and genetic drift.[3]

The typical white-bodied form of the peppered moth.

Industrial melanism: the black-bodied form of the peppered moth appeared in polluted areas.

The next key step was the work of the British biologist and statistician Ronald Fisher. In a series of papers starting in 1918 and culminating in his 1930 book *The Genetical Theory of Natural Selection*, Fisher showed that the continuous variation measured by the biometricians could be produced by the combined action of many discrete genes, and that natural selection could change allele frequencies in a population, resulting in evolution. In a series of papers beginning in 1924, another British geneticist, J. B. S. Haldane, worked out the mathematics of allele frequency change at a single gene locus under a broad range of conditions. Haldane also applied statistical analysis to real-world examples of natural selection, such as peppered moth evolution and industrial melanism, and showed that selection coefficients could be larger than Fisher assumed, leading to more rapid adaptive evolution as a camouflage strategy following increased pollution.[4][5]

The American biologist Sewall Wright, who had a background in animal breeding experiments, focused on combinations of interacting genes, and the effects of inbreeding on small, relatively isolated populations that exhibited genetic drift. In 1932 Wright introduced the concept of an adaptive landscape and argued that genetic drift and inbreeding could drive a small, isolated sub-population away from an adaptive peak, allowing natural selection to drive it towards different adaptive peaks

The work of Fisher, Haldane and Wright founded the discipline of population genetics. This integrated natural selection with Mendelian genetics, which was the critical first step in developing a unified theory of how evolution worked.[4][5] John Maynard Smith was Haldane's pupil, whilst W. D. Hamilton was influenced by the writings of Fisher. The American George R. Price worked with both Hamilton and Maynard Smith. American Richard Lewontin and Japanese Motoo Kimura were influenced by Wright and Haldane.

Modern synthesis

The mathematics of population genetics were originally developed as the beginning of the modern synthesis. Authors such as Beatty[6] have asserted that population genetics defines the core of the modern synthesis. For the first few decades of the 20th century, most field naturalists continued to believe that Lamarckism and orthogenesis provided the best explanation for the complexity they observed in the living world. During the modern synthesis, these ideas were purged, and only evolutionary causes that could be expressed in the mathematical framework of population genetics were retained. Consensus was reached as to which evolutionary factors might influence evolution, but not as to the relative importance of the various factors.

Theodosius Dobzhansky, a postdoctoral worker in T. H. Morgan's lab, had been influenced by the work on genetic

diversity by Russian geneticists such as Sergei Chetverikov. He helped to bridge the divide between the foundations of microevolution developed by the population geneticists and the patterns of macroevolution observed by field biologists, with his 1937 book *Genetics and the Origin of Species*. Dobzhansky examined the genetic diversity of wild populations and showed that, contrary to the assumptions of the population geneticists, these populations had large amounts of genetic diversity, with marked differences between sub-populations. The book also took the highly mathematical work of the population geneticists and put it into a more accessible form. Many more biologists were influenced by population genetics via Dobzhansky than were able to read the highly mathematical works in the original.[9]

In Great Britain E. B. Ford, the pioneer of ecological genetics, continued throughout the 1930s and 1940s to empirically demonstrate the power of selection due to ecological factors including the ability to maintain genetic diversity through genetic polymorphisms such as human blood types. Ford's work, in collaboration with Fisher, contributed to a shift in emphasis during the modern synthesis towards natural selection as the dominant force.

Neutral theory and origin-fixation dynamics

The original, modern synthesis view of population genetics assumes that mutations provide ample raw material, and focuses only on the change in frequency of alleles within populations. The main processes influencing allele frequencies are natural selection, genetic drift, gene flow and recurrent mutation. Fisher and Wright had some fundamental disagreements about the relative roles of selection and drift. The availability of molecular data on all genetic differences led to the neutral theory of molecular evolution. In this view, many mutations are deleterious and so never observed, and most of the remainder are neutral, i.e. are not under selection. With the fate of each neutral mutation left to chance (genetic drift), the direction of

evolutionary change is driven by which mutations occur, and so cannot be captured by models of change in the frequency of (existing) alleles alone.

The origin-fixation view of population genetics generalizes this approach beyond strictly neutral mutations, and sees the rate at which a particular change happens as the product of the mutation rate and the fixation probability.

Selection

Natural selection, which includes sexual selection, is the fact that some traits make it more likely for an organism to survive and reproduce. Population genetics describes natural selection by defining fitness as a propensity or probability of survival and reproduction in a particular environment. The fitness is normally given by the symbol w=1-s where s is the selection coefficient. Natural selection acts on phenotypes, so population genetic models assume relatively simple relationships to predict the phenotype and hence fitness from the allele at one or a small number of loci. In this way, natural selection converts differences in the fitness of individuals with different phenotypes into changes in allele frequency in a population over successive generations.

Before the advent of population genetics, many biologists doubted that small differences in fitness were sufficient to make a large difference to evolution. Population geneticists addressed this concern in part by comparing selection to genetic drift. Selection can overcome genetic drift when s is greater than 1 divided by the effective population size. When this criterion is met, the probability that a new advantageous mutant becomes fixed is approximately equal to 2s. The time until fixation of such an allele depends little on genetic drift, and is approximately proportional to log(sN)/s.

Dominance means that the phenotypic and/or fitness effect of one allele at a locus depends on which allele is present in

the second copy for that locus. Consider three genotypes at one locus, with the following fitness values

Genotype:	**A_1A_1**	**A_1A_2**	**A_2A_2**
Relative fitness:	**1**	**1-hs**	**1-s**

Population Genetics Glossary

- **species – a group of closely related organisms that are capable of interbreeding and producing fertile offspring**
- **population – the total number of individuals of a particular species in a given area**
- **interbreed – the ability to breed with another species (homo sapiens with homo neanderthal)**
- **gene pool – the collective genetic information contained within a population of sexually reproducing organisms**
- **allele frequency – the frequency or proportion of a particular allele of a gene within a population**
- **hardy–weinberg principle – an equation linking allele frequency to genotype frequency in a population $p^2+2pq+q^2=1$**

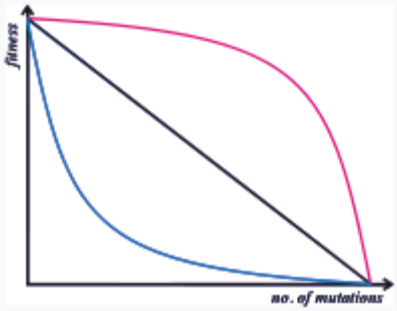
- **v**
- **t**
- **e**

s is the selection coefficient and h is the dominance coefficient. The value of h yields the following information:

h=0	**A_1 dominant, A_2 recessive**
h=1	**A_2 dominant, A_1 recessive**

0<h<1	**incomplete dominance**
h<0	**overdominance**
h>1	**Underdominance**

Epistasis[edit]



The logarithm of fitness as a function of the number of deleterious mutations. Synergistic epistasis is represented by the red line - each subsequent deleterious mutation has a larger proportionate effect on the organism's fitness. Antagonistic epistasis is in blue. The black line shows the non-epistatic case, where fitness is the product of the contributions from each of its loci.

Epistasis means that the phenotypic and/or fitness effect of an allele at one locus depends on which alleles are present at other loci. Selection does not act on a single locus, but on a phenotype that arises through development from a complete genotype.[20] However, many population genetics models of sexual species are "single locus" models, where the fitness of an individual is calculated as the product of the contributions from each of its loci—effectively assuming no epistasis.

In fact, the genotype to fitness landscape is more complex. Population genetics must either model this complexity in detail, or capture it by some simpler average rule. Empirically, beneficial mutations tend to have a smaller fitness benefit when added to a genetic background that

already has high fitness: this is known as diminishing returns epistasis.[21] When deleterious mutations also have a smaller fitness effect on high fitness backgrounds, this is known as "synergistic epistasis". However, the effect of deleterious mutations tends on average to be very close to multiplicative, or can even show the opposite pattern, known as "antagonistic epistasis".

Synergistic epistasis is central to some theories of the purging of mutation load and to the evolution of sexual reproduction.

Mutation

Drosophila melanogaster

Mutation is the ultimate source of genetic variation in the form of new alleles. In addition, mutation may influence the direction of evolution when there is mutation bias, i.e. different probabilities for different mutations to occur. For example, recurrent mutation that tends to be in the opposite direction to selection can lead to mutation–selection balance. At the molecular level, if mutation from G to A happens more often than mutation from A to G, then genotypes with A will tend to evolve. Different insertion vs. deletion mutation biases in different taxa can lead to the evolution of different genome sizes. Developmental or mutational biases have also been observed in morphological evolution. For example, according to the phenotype-first theory of evolution, mutations can eventually cause the genetic assimilation of traits that were previously induced by the environment.

Mutation bias effects are superimposed on other processes. If selection would favor either one out of two mutations, but there is no extra advantage to having both, then the mutation that occurs the most frequently is the one that is most likely to become fixed in a population.

Mutation can have no effect, alter the product of a gene, or prevent the gene from functioning. Studies in the fly *Drosophila melanogaster* suggest that if a mutation changes a protein produced by a gene, this will probably be harmful, with about 70 percent of these mutations having damaging effects, and the remainder being either neutral or weakly beneficial. Most loss of function mutations are selected against. But when selection is weak, mutation bias towards loss of function can affect evolution. For example, pigments are no longer useful when animals live in the darkness of caves, and tend to be lost. This kind of loss of function can occur because of mutation bias, and/or because the function had a cost, and once the benefit of the function disappeared, natural selection leads to the loss. Loss of sporulation ability in a bacterium during laboratory evolution appears to have been caused by mutation bias, rather than natural selection against the cost of maintaining sporulation ability.[36] When there is no selection for loss of function, the speed at which loss evolves depends more on the mutation rate than it does on the effective population size, indicating that it is driven more by mutation bias than by genetic drift.

Mutations can involve large sections of DNA becoming duplicated, usually through genetic recombination. This leads to copy-number variation within a population. Duplications are a major source of raw material for evolving new genes. Other types of mutation occasionally create new genes from previously noncoding DNA.

Genetic drift

***Genetic drift* is a change in allele frequencies caused by random sampling.- That is, the alleles in the offspring are**

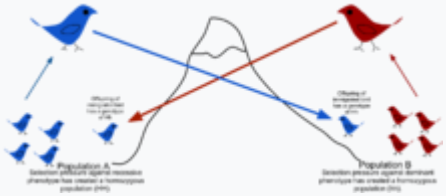
a random sample of those in the parents. Genetic drift may cause gene variants to disappear completely, and thereby reduce genetic variability. In contrast to natural selection, which makes gene variants more common or less common depending on their reproductive success,the changes due to genetic drift are not driven by environmental or adaptive pressures, and are equally likely to make an allele more common as less common.

The effect of genetic drift is larger for alleles present in few copies than when an allele is present in many copies. The population genetics of genetic drift are described using either branching processes or a diffusion equation describing changes in allele frequency. These approaches are usually applied to the Wright-Fisher and Moran models of population genetics. Assuming genetic drift is the only evolutionary force acting on an allele, after t generations in many replicated populations, starting with allele frequencies of p and q, the variance in allele frequency across those populations is

Ronald Fisher held the view that genetic drift plays at the most a minor role in evolution, and this remained the dominant view for several decades. No population genetics perspective have ever given genetic drift a central role by itself, but some have made genetic drift important in combination with another non-selective force. The shifting balance theory of Sewall Wright held that the combination of population structure and genetic drift was important. Motoo Kimura's neutral theory of molecular evolution claims that most genetic differences within and between populations are caused by the combination of neutral mutations and genetic drift.[47]

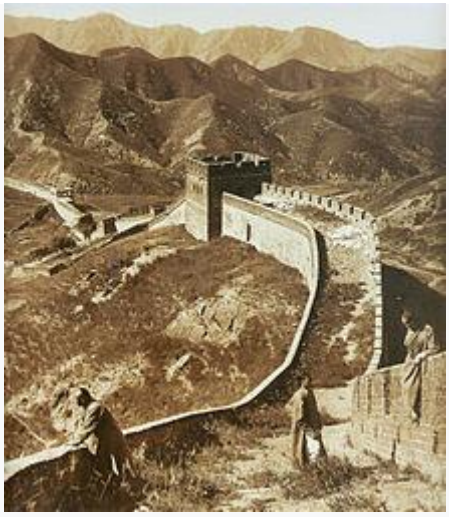
The role of genetic drift by means of sampling error in evolution has been criticized by John H Gillespieand Will Provine, who argue that selection on linked sites is a more important stochastic force, doing the work traditionally ascribed to genetic drift by means of sampling error. The

mathematical properties of genetic draft are different from those of genetic drift. The direction of the random change in allele frequency is autocorrelated across generations.



Gene flow is the transfer of alleles from one population to another population through immigration of individuals. In this example, one of the birds from population A immigrates to population B, which has fewer of the dominant alleles, and through mating incorporates its alleles into the other population.

Because of physical barriers to migration, along with the limited tendency for individuals to move or spread (vagility), and tendency to remain or come back to natal place (philopatry), natural populations rarely all interbreed as may be assumed in theoretical random models (panmixy). There is usually a geographic range within which individuals are more closely related to one another than those randomly selected from the general population. This is described as the extent to which a population is genetically structured.



The Great Wall of China is an obstacle to gene flow of some terrestrial species.

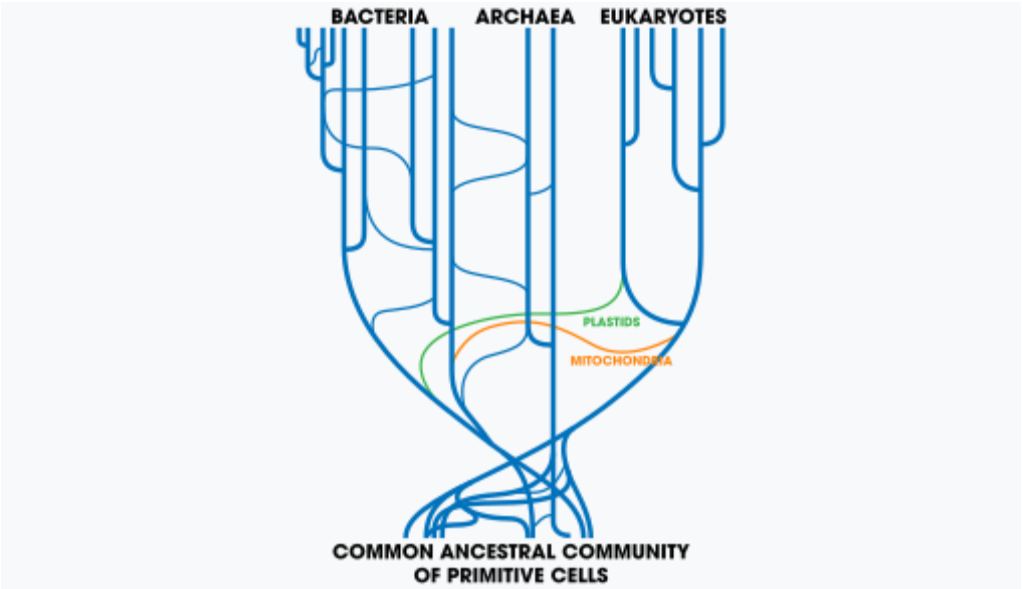
Genetic structuring can be caused by migration due to historical climate change, species range expansion or current availability of habitat. Gene flow is hindered by mountain ranges, oceans and deserts or even man-made structures such as the Great Wall of China, which has hindered the flow of plant genes.

Gene flow is the exchange of genes between populations or species, breaking down the structure. Examples of gene flow within a species include the migration and then breeding of organisms, or the exchange of pollen. Gene transfer between species includes the formation of hybrid organisms and horizontal gene transfer. Population genetic models can be used to identify which populations show significant genetic isolation from one another, and to reconstruct their history.

Subjecting a population to isolation leads to inbreeding depression. Migration into a population can introduce new genetic variants, potentially contributing to evolutionary rescue. If a significant proportion of individuals or gametes migrate, it can also change allele frequencies, e.g. giving rise to migration load.

In the presence of gene flow, other barriers to hybridization between two diverging populations of an outcrossing species are required for the populations to become new species.

Horizontal gene transfer



Current tree of life showing vertical and horizontal gene transfers.

Horizontal gene transfer

Horizontal gene transfer is the transfer of genetic material from one organism to another organism that is not its offspring; this is most common among prokaryotes. In medicine, this contributes to the spread of antibiotic resistance, as when one bacteria acquires resistance genes it can rapidly transfer them to other species. Horizontal transfer of genes from bacteria to eukaryotes such as the yeast *Saccharomyces cerevisiae* and the adzuki bean beetle *Callosobruchus chinensis* may also have occurred. An example of larger-scale transfers are the eukaryotic bdelloid rotifers, which appear to have received a range of genes from bacteria, fungi, and plants. Viruses can also carry DNA between organisms, allowing transfer of genes even across biological domains. Large-scale gene transfer has also occurred between the ancestors of eukaryotic cells and prokaryotes, during the acquisition of chloroplasts and mitochondria.

Linkage

If all genes are in linkage equilibrium, the effect of an allele at one locus can be averaged across the gene pool at

other loci. In reality, one allele is frequently found in linkage disequilibrium with genes at other loci, especially with genes located nearby on the same chromosome. Recombination breaks up this linkage disequilibrium too slowly to avoid genetic hitchhiking, where an allele at one locus rises to high frequency because it is linked to an allele under selection at a nearby locus. Linkage also slows down the rate of adaptation, even in sexual populations. The effect of linkage disequilibrium in slowing down the rate of adaptive evolution arises from a combination of the Hill–Robertson effect (delays in bringing beneficial mutations together) and background selection (delays in separating beneficial mutations from deleterious hitchhikers).

Linkage is a problem for population genetic models that treat one gene locus at a time. It can, however, be exploited as a method for detecting the action of natural selection via selective sweeps.

In the extreme case of an asexual population, linkage is complete, and population genetic equations can be derived and solved in terms of a travelling wave of genotype frequencies along a simple fitness landscape. Most microbes, such as bacteria, are asexual. The population genetics of their adaptation have two contrasting regimes. When the product of the beneficial mutation rate and population size is small, asexual populations follow a "successional regime" of origin-fixation dynamics, with adaptation rate strongly dependent on this product. When the product is much larger, asexual populations follow a "concurrent mutations" regime with adaptation rate less dependent on the product, characterized by clonal interference and the appearance of a new beneficial mutation before the last one has fixed.

Applications

Explaining levels of genetic variation

Neutral theory predicts that the level of nucleotide diversity in a population will be proportional to the product of the population size and the neutral mutation rate. The fact that levels of genetic diversity vary much less than population sizes do is known as the "paradox of variation". While high levels of genetic diversity were one of the original arguments in favor of neutral theory, the paradox of variation has been one of the strongest arguments against neutral theory.

It is clear that levels of genetic diversity vary greatly within a species as a function of local recombination rate, due to both genetic hitchhiking and background selection. Most current solutions to the paradox of variation invoke some level of selection at linked sites. For example, one analysis suggests that larger populations have more selective sweeps, which remove more neutral genetic diversity. A negative correlation between mutation rate and population size may also contribute.

Life history affects genetic diversity more than population history does, e.g. r-strategists have more genetic diversity.

Detecting selection

Population genetics models are used to infer which genes are undergoing selection. One common approach is to look for regions of high linkage disequilibrium and low genetic variance along the chromosome, to detect recent selective sweeps.

A second common approach is the McDonald–Kreitman test which compares the amount of variation within a species (polymorphism) to the divergence between species (substitutions) at two types of sites; one assumed to be neutral. Typically, synonymous sites are assumed to be neutral. Genes undergoing positive selection have an excess of divergent sites relative to polymorphic sites. The test can also be used to obtain a genome-wide estimate of the proportion of substitutions that are fixed by positive selection, α.[73][74] According to the neutral theory of

molecular evolution, this number should be near zero. High numbers have therefore been interpreted as a genome-wide falsification of neutral theory.

Demographic inference

The simplest test for population structure in a sexually reproducing, diploid species, is to see whether genotype frequencies follow Hardy-Weinberg proportions as a function of allele frequencies. For example, in the simplest case of a single locus with two alleles denoted A and a at frequencies p and q, random mating predicts freq(AA) = p^2 for the AA homozygotes, freq(aa) = q^2 for the aa homozygotes, and freq(Aa) = $2pq$ for the heterozygotes. In the absence of population structure, Hardy-Weinberg proportions are reached within 1-2 generations of random mating. More typically, there is an excess of homozygotes, indicative of population structure. The extent of this excess can be quantified as the inbreeding coefficient, F.

Individuals can be clustered into K subpopulations. The degree of population structure can then be calculated using F_{ST}, which is a measure of the proportion of genetic variance that can be explained by population structure. Genetic population structure can then be related to geographic structure, and genetic admixture can be detected.

Coalescent theory relates genetic diversity in a sample to demographic history of the population from which it was taken. It normally assumes neutrality, and so sequences from more neutrally-evolving portions of genomes are therefore selected for such analyses. It can be used to infer the relationships between species (phylogenetics), as well as the population structure, demographic history (e.g. population bottlenecks, population growth), biological dispersal, source–sink dynamics[78] and introgression within a species.

Another approach to demographic inference relies on the allele frequency spectrum.

Evolution of genetic systems

By assuming that there are loci that control the genetic system itself, population genetic models are created to describe the evolution of dominance and other forms of robustness, the evolution of sexual reproduction and recombination rates, the evolution of mutation rates, the evolution of evolutionary capacitors, the evolution of costly signalling traits, the evolution of ageing, and the evolution of co-operation. For example, most mutations are deleterious, so the optimal mutation rate for a species may be a trade-off between the damage from a high deleterious mutation rate and the metabolic costs of maintaining systems to reduce the mutation rate, such as DNA repair enzymes. One important aspect of such models is that selection is only strong enough to purge deleterious mutations and hence overpower mutational bias towards degradation if the selection coefficient s is greater than the inverse of the effective population size. This is known as the drift barrier and is related to the nearly neutral theory of molecular evolution. Drift barrier theory predicts that species with large effective population sizes will have highly streamlined, efficient genetic systems, while those with small population sizes will have bloated and complex genomes containing for example introns and transposable elements. However, somewhat paradoxically, species with large population sizes might be so tolerant to the consequences of certain types of errors that they evolve higher error rates, e.g. in transcription and translation, than small populations.

Practical evidences that DNA is the genetic matter

Firstly chromosome consists of protein and DNA , protein miss the following:

1-The ability of replication accurately and constancy

2- ……mutation by very low ratio allow by genetic variations transport to progeny

3- Not contain on all genetic knowledge responsible for bioactivity

Genetic transformation – التحول الوراثي

- **1944 حصل Avery علي سلالتين من البكتريا المسببة للالتهاب الرئوي في الانسان أحدهما smooth وتكون كبسولة S-type والثانية Rough لا تكون كبسولة R-type**
- **التجربة الأولي: نزع بروتين السلالة S ونزع RNA وأضاف DNA الي مزرعة من R-type فحصل علي خلايا من نوع السلالة S-type**

التجربة الثانية : تخلص من DNA بـDNase وأبقي البروتين و RNA وأضاف الي مزرعة R-type فلم يحصل علي S-type

Genetic transduction

الاستنقال الوراثي (النقل الفاجي)

- **تمكن Hershy and Chase 1952من عدوي E. Coli بالفاج T2 بعد تعليم بروتينه بالكبريت المشع 32 وال DNA الخاص به بالفوسفور المشع 35 وقد وجدا أن الفاجات الناشئة كلها تحتوي**

علي DNA المشع ولا يحتوي علي البروتين المشع مما يؤكد أن ال DNA وليس البروتين هو المسئول عن التوريث

The third evidence the content of DNA constant in the cell and metabolism (catabolism- anabolism) not take place as protein