M2 "BIM" 2009 Cours : Génétique, Génomes et Evolution

22/09 - semaine 1:	Introduction à la génétique et à la génomique	3h00 cours Gilles Fischer
29/09 - semaine 2:	Les différents niveaux d'information codés dans les génomes	3h00 cours Claude Thermes
06/10 - semaine 3:	Comparaison de séquences et notion d'homologie	3h00 cours Ingrid Lafontaine
13/10 - semaine 4:	Anatomie et annotations des génomes	2h00 cours Gilles Fischer + 2h00 TP Cécile Neuvéglise
20/10 - semaine 5:	Biais de codons	3h00 cours Alessandra Carbone
27/10 - semaine 6:	Les réarrangements chromosomiques et la notion de synténie	2h00 cours Gilles Fischer + 2h00 TP Guénola Drillon
03/11 - semaine 7:	L'évolution des génomes	3h00 cours Bernard Dujon

Introduction à la génétique et à la génomique

- Une introduction à la Biologie
- Le fonctionnement des génomes
  - ADN
  - ARN
  - Dogme de la biologie moléculaire
  - Code génétique
- La division cellulaire
  - Réplication
  - Mitose
  - Méiose
- Quelques éléments de génétique
  - Phénotype-Génotype
  - Mutations

LA BIOLOGIE

L'étude de la vie (Bios-logia)

"Tout ce qui est généralement commun aux végétaux et aux animaux comme toutes les facultés qui sont propres à chacun de ces êtres sans exception, doit constituer l'unique et vaste objet d'une science particulière qui n'est pas encore fondée, qui n'a même pas de nom, et à laquelle je donnerai le nom de biologie." (1802)

Jean-Baptiste de Lamarck (1744-1829)

Botanique
Zoologie
Microbiologie

Ecologie
Taxonomie

Biologie cellulaire

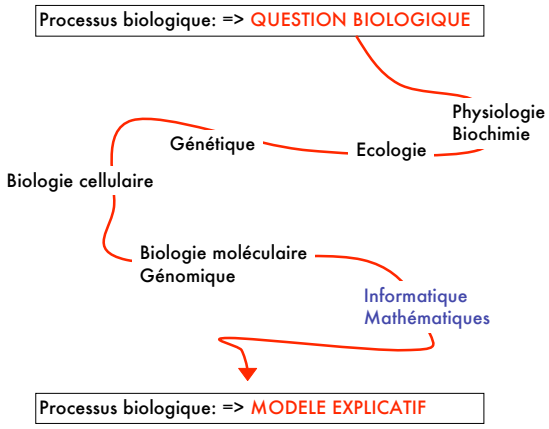
Physiologie
Biochimie
Génétique
Biologie moléculaire

Génomique

Biophysique
Bioinformatique
Modélisation
(biomathématiques?)

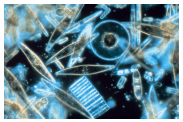
La biologie consiste en l'étude de | la classification / la distribution / la structure / le fonctionnement / la reproduction / l'évolution | de toutes les formes vivantes

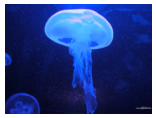
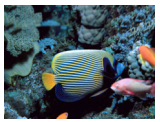
Comment transformer un informaticien ou un mathématicien en biologiste ?

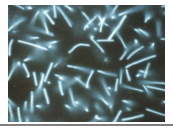


La biologie est vaste…

En raison de l'extrême diversité du vivant => biodiversité



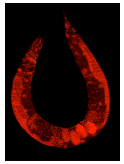





… la diversité de taille

Mycoplasma genitalium

Caenorhabditis elegans



Pan troglodytes

Armillaria ostoyae

≤ 1 µm

1 mm

1 m

≥ 1 km

> 9 ordres de grandeur !!!

Quels sont les points communs entre tous les être vivants?

Qu'est-ce que la vie ?

les principes fondamentaux :

- Reproduction : toute forme de vie se multiplie par division avec hérédité
- Hérédité : transmission des caractères innés d'un individu à sa descendance

- Evolution : Les caractères se modifient de générations en générations sous l'action de la sélection naturelle et de la dérive génétique

Qu'est-ce que la vie ?

les caractères universels :

-La théorie cellulaire : - l'unité de base de la vie est la cellule (problème des virus, sont-ils vivants?)
- Les cellules proviennent de cellules par division
- Toutes les cellules d'un organisme proviennent d'une cellule unique (œuf fécondé)

Mazzarello, P (1999). "A unifying concept: the history of cell theory". Nature Cell Biology 1: E13–E15

- Le carbone : élément qui sert de squelette à tous les composés organiques

-Les gènes : segments d'ADN (ou d'ARN) qui spécifient les caractères héréditaires

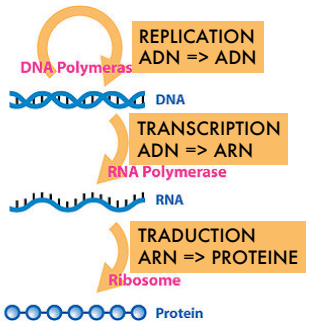
-L'énergie : attribut nécessaire à tout être vivant pour maintenir son état

1- LE FONCTIONNEMENT DES GENOMES

Dogme central de la biologie Moléculaire

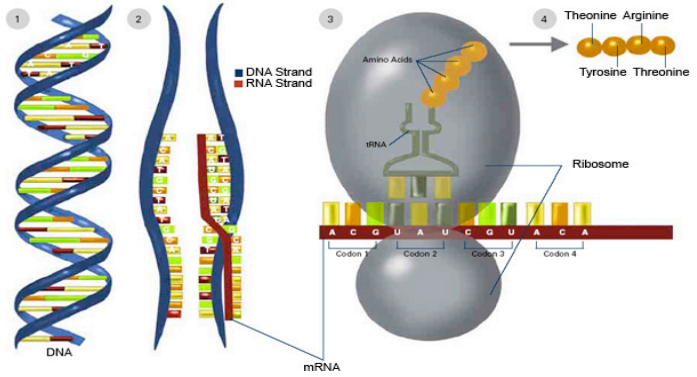
Francis Crick 1958 et 1970

Caractérisation de l'ARNm comme intermédiaire entre l'ADN et les protéines
Jacques Monod (Prix Nobel 1965)



Dogme central de la biologie Moléculaire

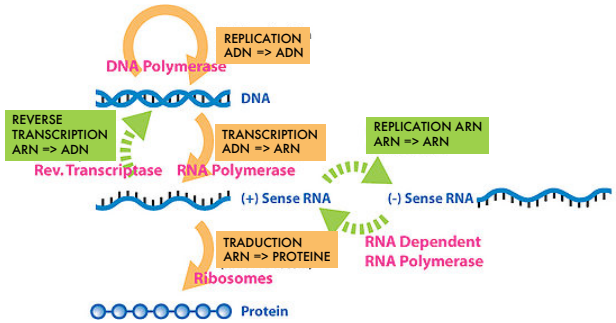
Francis Crick 1958 et 1970



Dogme central de la biologie Moléculaire

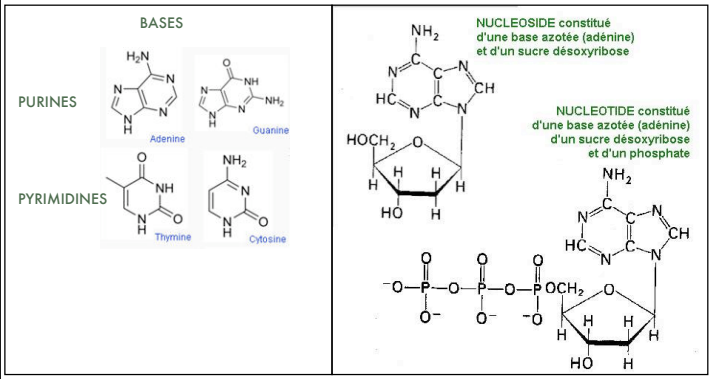
Francis Crick 1958 et 1970

Caractérisation de la reverse transcriptase : ARN = support de l'information génétique qui peut être retro-transcrit en ADN. Howard M. Temin (Prix Nobel 1975)



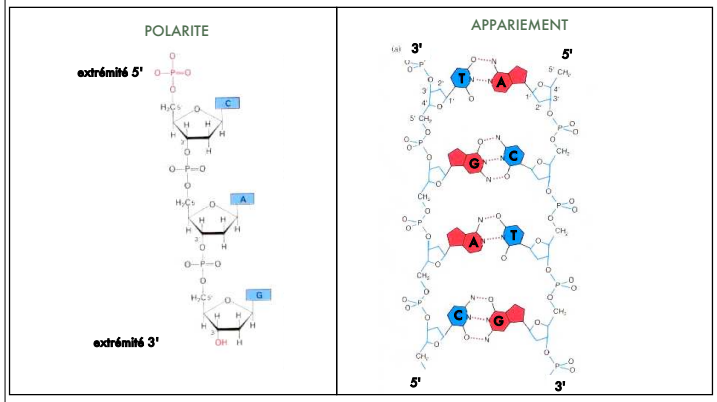
Structure de l'ADN

Watson et Francis Crick 1953



Structure de l'ADN

Watson et Francis Crick 1953



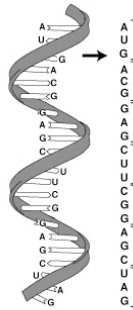
Code Génétique

Georges Gamow : code à 3 lettres car 3 est le plus petit entier tel que $4^n \geq 20$

Crick et Brenner (1961) : le codon est une succession de 3 bases

Nirenberg et Matthaei (1961) : première identification d'un codon (UUU Phe)

20 acides aminés, 3 codons stops, 64 possibilités



RNA

Ribonucleic acid

	U	C	A	G
U	UUU Phenylalanine UUC Phenylalanine UUA Leucine UUG Leucine	UCU Serine UCC Serine UCA Serine UCG Serine	UAU Tyrosine UAC Tyrosine UAA Stop UAG Stop	UGU Cysteine UGC Cysteine UGA Stop UGG Tryptophan
C	CUU Leucine CUC Leucine CUA Leucine CUG Leucine	CCU Proline CCC Proline CCA Proline CCG Proline	CAU Histidine CAC Histidine CAA Glutamine CAG Glutamine	CGU Arginine CGC Arginine CGA Arginine CGG Arginine
A	AUU Isoleucine AUC Isoleucine AUA Isoleucine AUG Methionine	ACU Threonine ACC Threonine ACA Threonine ACG Threonine	AAU Asparagine AAC Asparagine AAA Lysine AAG Lysine	AGU Serine AGC Serine AGA Arginine AGG Arginine
G	GUU Valine GUC Valine GUA Valine GUG Valine	GCU Alanine GCC Alanine GCA Alanine GCG Alanine	GAU Aspartate GAC Aspartate GAA Glutamate GAG Glutamate	GGU Glycine GGC Glycine GGA Glycine GGG Glycine

Code Génétique

Propriétes :

Redondant — AUU/AUC/AUA => Isoleucine

Non ambigu — AUU => toujours Isoleucine

Dégénéré — 1ere, 2eme ou 3eme positions

UCU Serine
UCC Serine
UCA Serine
UCG Serine

AGU Serine
AGC Serine

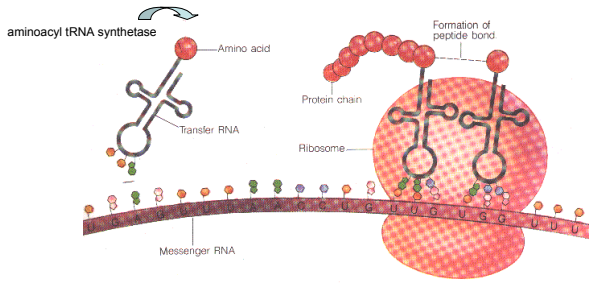
site dégénéré d'ordre 4
(3eme base de certains codons)

CCU Proline
CCC Proline
CCA Proline
CCG Proline

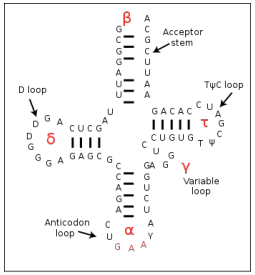
substitutions synonymes

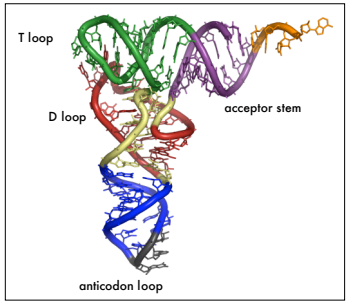
les ARN de transfert: tRNA

intermédiaire entre l'ARNm et la protéine



les ARN de transfert: tRNA





74-95 nucléotides

Chaque tRNA contient une boucle anticodon spécifique

La redondance du code génétique implique que plusieurs tRNA différents (anticodons différents) peuvent spécifier le même acide aminé.

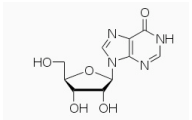
La nécessaire correspondance entre acides aminés et tRNA impliquerait l'existence de 61 types de tRNA, or la plupart des génomes contiennent moins de 61 types de tRNA (e.g. 49 chez l'homme)

les ARN de transfert: tRNA

Une boucle anticodon donnée peut reconnaître plusieurs codons différents

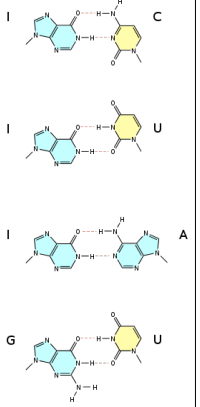
⇓

Les appariements WOBBLE



Inosine (hypoxanthine+ribose)

en première position de l'anticodon

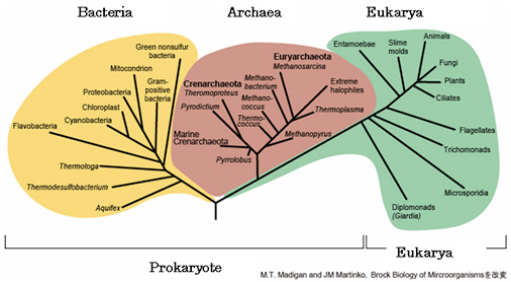


TOUS les être vivants partagent les mêmes propriétés

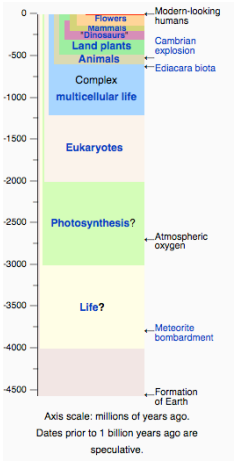
ADN	4 bases A T G C , double-brin, répliqué par une ADN polymérase
Code Génétique	64 codons de 3 lettres
Expression des gènes	ARN simple-brin (mRNA, tRNA, rRNA)
Expression des gènes	ARN simple-brin (mRNA, tRNA, rRNA) produits par des ARN polymérases ADN dépendante.
Synthèse des protéines	à partir de 20 acides aminés (forme L) par traduction des ARNm au sein des ribosomes
Les ribosomes	constitués de 2 sous-unités (grosse et petites) formés d'ARNr (catalytiques et de protéines
Le glucose	utilisé comme source de carbone et d'energie (glycolyse), seul l'isomere D est utilisée
L'energie	stockée sous forme d'ATP
La cellule	entourée d'une bi-couche lipidique
La division cellulaire	assure la reproduction et la croissance des organismes

TOUS les être vivants proviennent d'un ancêtre unique

la vie sur Terre a pour origine une seule et même forme de vie, désignée sous l'acronyme de LUCA (pour Last Universal Common Ancestor), qui serait apparue sur Terre il y a environ 3,5 milliards d'années.



Chronologie de l'évolution :



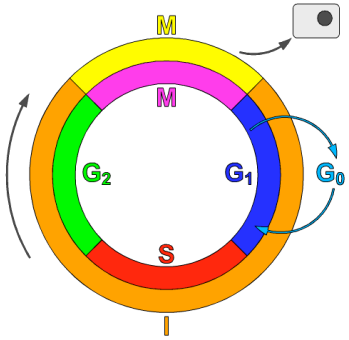
A l'échelle d'une journée :

00h00	formation de la terre
05h00	apparition de la vie
09h00	premiers eucaryotes
11h00	organismes pluricellulaires
18h00	premiers mollusques
21h40	premiers vertebrés (poissons)
22h55	premiers oiseaux
23h40	extinction des dinosaures
23h58	premiers hominidés
23h59m50s	maitrise du feu
23h59m57s	Néandertal et Cro-Magnon
23h59m58s	parole

1 sec = 50 000 ans
19 sec = 1 Ma
1 heure = 187 Ma

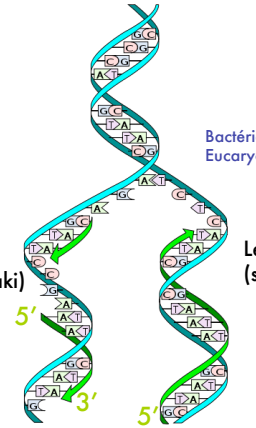
2 - LA DIVISION CELLULAIRE

Le cycle cellulaire eucaryote :



M : mitose
I : interphase
G1 : gap 1
S : syntèse ADN
G2 : gap 2
G0 : gap 0

La réplication de l'ADN :



Bactéries: 1000 nucléotides/sec
Eucaryotes : 50 nucléotides/sec

Lagging strand
(fragments d'Okazaki)

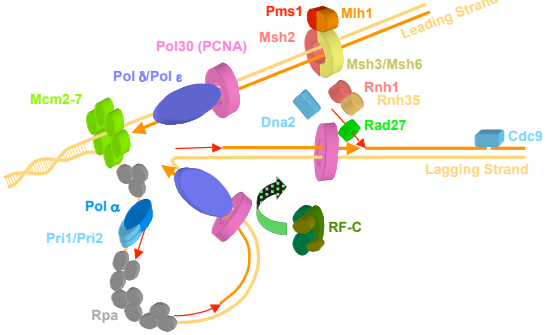
Leading strand
(synthèse continue)

5′ 3′ 5′

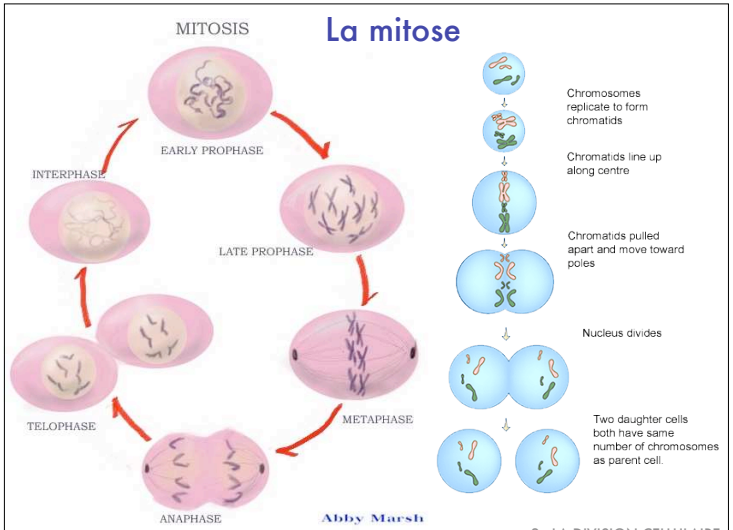
La replication est semi-conservative

La réplication de l'ADN :

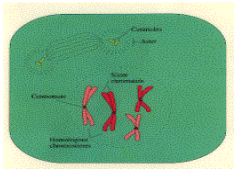
Anatomie d'une Fourche de Réplication



La mitose

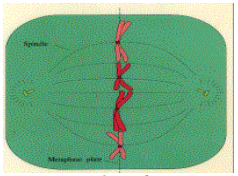


La mitose :



Prophase

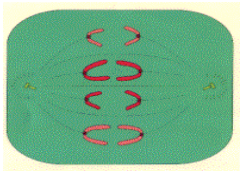
- condensation des chromosomes
- chromosome: 2 chromatides liées par leur centromère.
- mise en place du fuseau à partir du centrosome



Métaphase

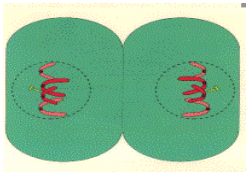
- le fuseau est en place, les chromosomes s'alignent sur la plaque équatoriale par leur centromère.

La mitose :



Anaphase

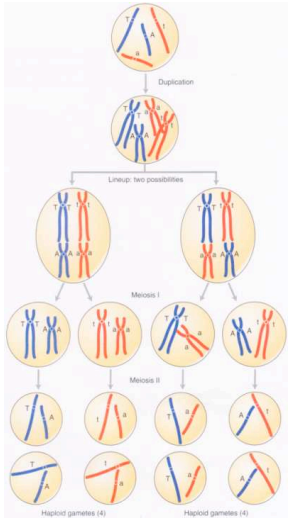
- les chromatides sœurs se séparent, chacune migrant vers l'un des pôles.



Télophase

- séparation en deux cellules filles
- les chromosomes se décondensent l'enveloppe nucléaire se reforme

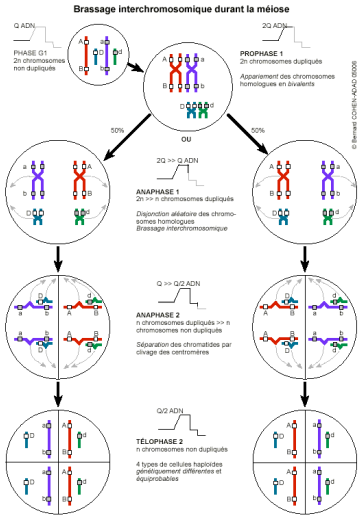
MITOSE: Racine d'ail



La méiose

Appariement des chromosomes homologues

Clivage des centromères à la deuxième division

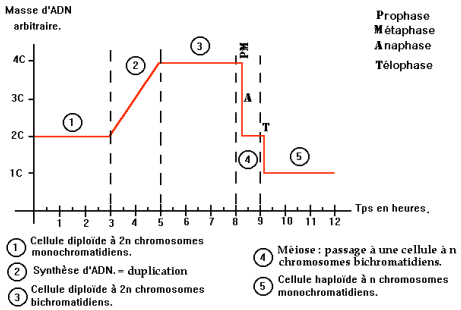


La méiose

23 paires de chromosomes chez l'humain

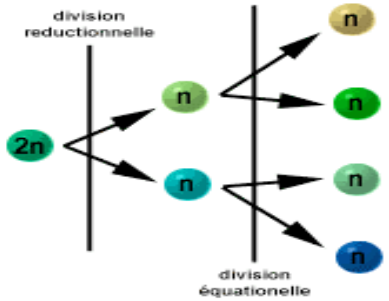
1 chance sur 2^{23} (8 millions) que deux spermatozoïdes ou ovules aient le même lot de chromosomes

La méiose : variation du contenu en ADN

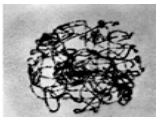


La méiose : 2 divisions cellulaires aboutissant aux cellules germinales

1 cellule 2n → 4 cellules n



La méiose : division reductionnelle



Durant la prophase I, on distingue 5 stades successifs:

Prophase I

- **Stade leptotène:** les chromosomes commencent à se condenser, ils sont longs et grèles.

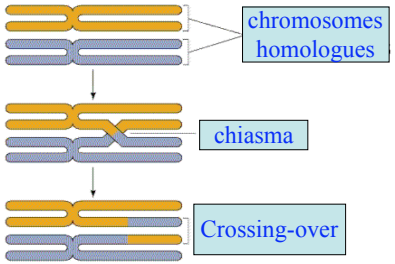
- Stade zygotène: les chromosomes homologues commencent à s'apparier

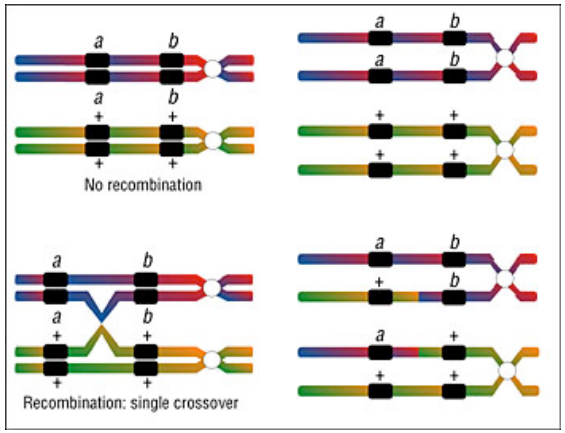
- **Stade pachytène:** les chromosomes homologues sont bien appariés, crossing-overs.

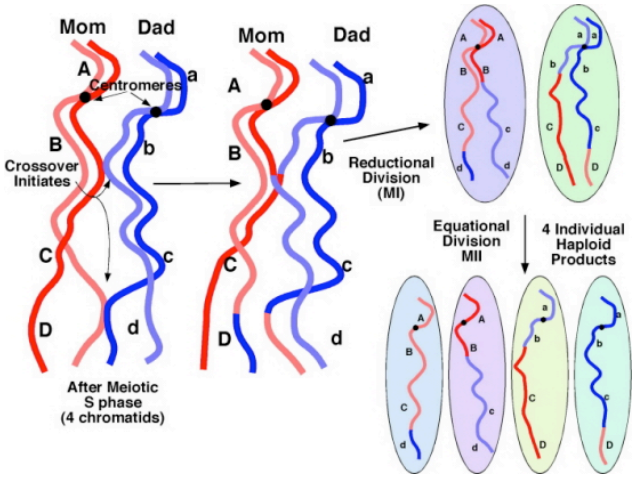
- **Stade diplotène:** les chiasmas deviennent visibles, désappariement progressif des chromosomes homologues

- **Stade diacinèse:** poursuite du désappariement, condensation maximale des chromosomes.

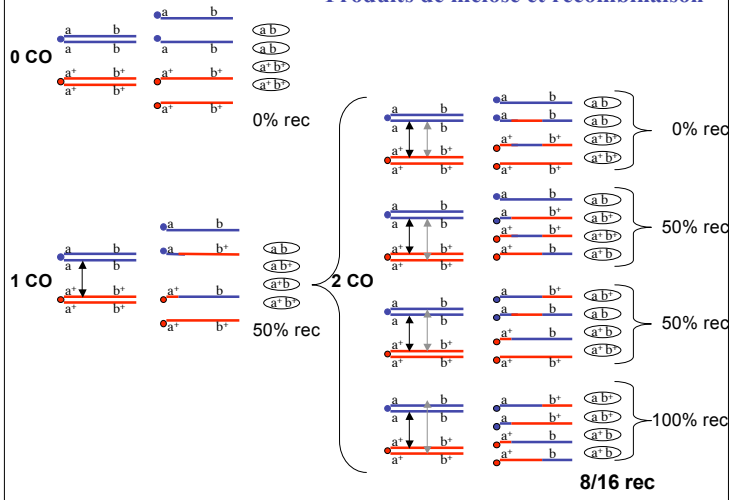
La méiose : recombinaison génétique







Produits de méiose et recombinaison



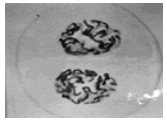
La méiose : division réductionnelle

Metaphase I

Les chromosomes se rangent par paires d'homologues sur un plan équatorial.
Disparition de la membrane nucléaire.

Anaphase I

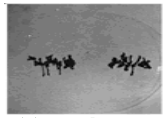
Les chromosomes de chaque paire se séparent et chaque chromosome migre vers l'un des pôles.



Télophase I

Les chromosomes s'organisent en deux lots haploïdes: fin de la **division réductionnelle**.

La méiose : division équationnelle



Métaphase II

Deux lots haploïdes de chromosomes s'organisent sur deux plans équatoriaux.

Anaphase II

Deux chromatides de chaque chromosome se séparent et migrent vers des pôles opposés: **division équationnelle.**

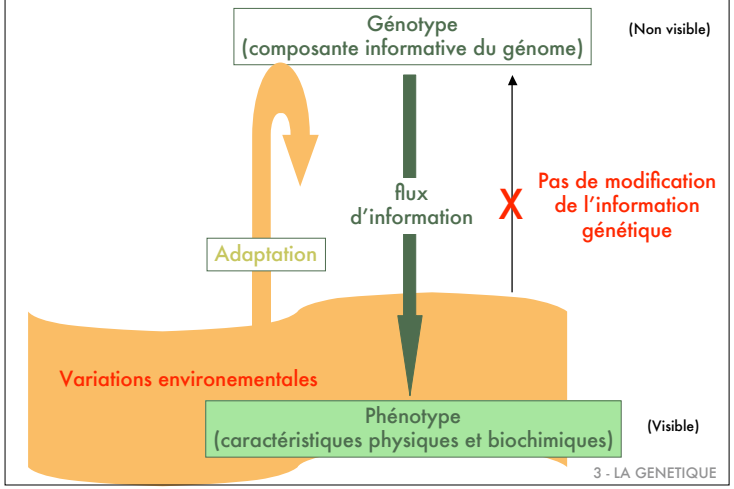
Comparaison entre mitose et méiose

Mitosis	Meiosis
In somatic cells	In cells in the sexual cycle
One cell division, resulting in two daughter cells (Parental cell → Daughter cells)	Two cell divisions, resulting in four products of meiosis (Meiocyte → Products of meiosis)
Chromosome number per nucleus maintained (e.g., for a diploid cell) (2n → 2n)	Chromosome number halved in the products of meiosis (2n → n)
One premeiotic S phase per cell division (e.g., for a diploid cell) (DNA per nucleus: 4, 3, 2, 1; G1 S G2 M G1)	One premeiotic S phase for both cell divisions (DNA per nucleus: 4, 3, 2, 1; S M_I M_{II})
Normally, no pairing of homologs	Full synapsis of homologs at prophase I
Normally, no crossovers	At least one crossover per homologous pair
Centromeres divide at anaphase	Centromeres do not divide at anaphase I but do at anaphase II
Conservative process: daughter cells' genotypes identical with parental genotype	Promotes variation among the products of meiosis
Cell undergoing mitosis can be diploid or haploid	Cell undergoing meiosis is diploid

3- LA GENETIQUE

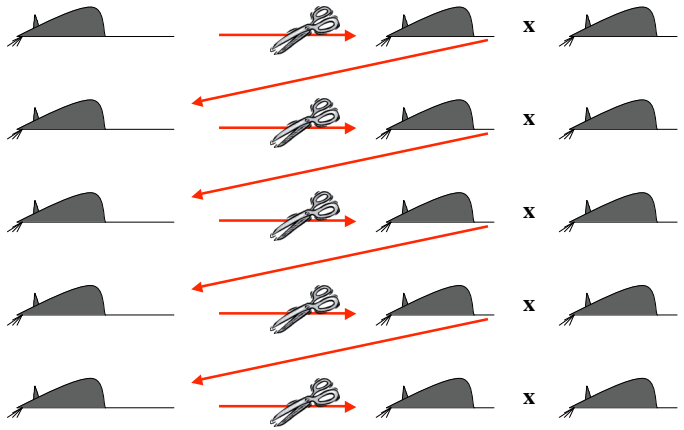
science de l'expression et de la transmission des caractères

Expression des caractères: Relation génotype - phénotype



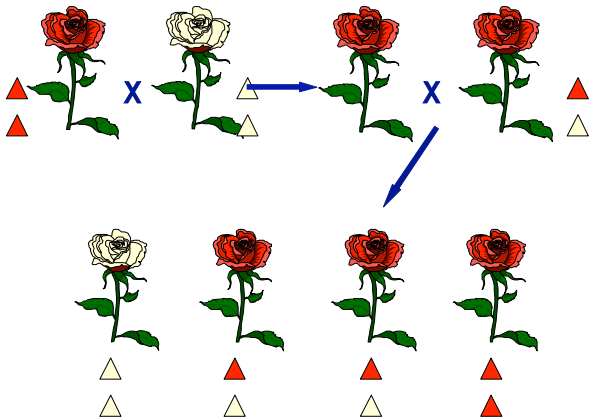
Transmission des caractères :

Pas d'hérédité des caractères acquis (A. Weissman, 1883)



Lignée somatique et lignée germinale

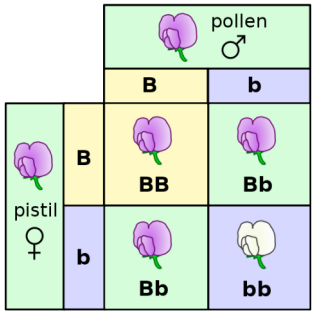
Relation génotype - phénotype :



Dominance/récessivité

Relation génotype - phénotype :

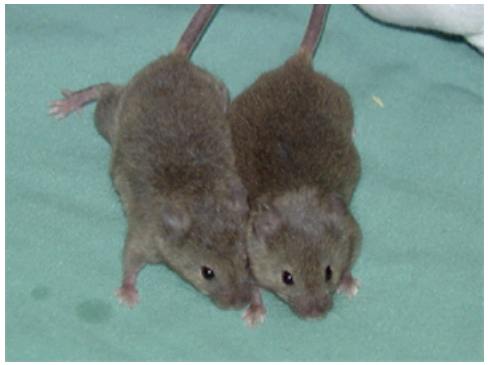
Hérédité Mendelienne



Les lois de Mendel (1865), redécouvertes début XXeme

Relation génotype - phénotype :

Un exemple pas si simple que ça : la couleur du pelage de la souris



Relation génotype - phénotype :

Le gène « tyrosinase » : une séquence d'ADN

```
caattaacagctggcatcaggggaaaagcagacatccaacaacactagctctgaaggagatcagc
agagaaaccttccagggattcatgctgtattttcatctgagcacccctgtcttctccatgcaaag
agcagcataggagacctgtgttctgaactcttgcttcgagaagaatgaaatcttacaacgtcctc
cccctagcctatatctccctttttcctgatgctgttttatcaggtttgggctcagtttccacgaga
gtgtgccaatattgaggctctgagacgtggggtgtgttgcccagacctgctcccttcctctggac
cggggactgacccttgtggctcatcatcaggaagaggcaggtgtgtggctgtgattgcagactcc
cgaccccacagccgccattatccccacgatggtaaagatgaccgagaagcctggcctctgaggtt
ctttaatagaacatgtcagtgcaatgataatttctcaggacacaactgtgggacttgccgtcctg
ggtggagaggagctgcatgcaaccagaaaattctcacagtcaggagaaatcttctagacttaagt
ccagaagaaaagagccactttgtcagggccttggatatggcgaagcgcacaactcaccctcaatt
tgtcattgccacaaggaggttagaagacatactgggaccagatggcaacacaccacaatttgaga
acatttccgtttataactactttgtttggacacactattattcagtcaaaaaaaaccttcctcggg
acaggacaggaaagctttggggatgtggatttctctcacgaaggacccgcttttctcacatggca
caggtaccatctgctgcagctggagagagacatgcaggagatgctgcaggagccttctttctccc
ttccttactggaattttgcaactgggaaaaacgtctgcgatgtctgcactgatgacttgatggga
tccagaagcaacttcgattctactcttataagccccaactctgtctttttctcaatggagagtggt
ctgtgaatccttggaagagtacgatacccctgggaacactttgtaacagcactgagggtggaccaa
tcaggagaaacccagctggaaatgtagggagaccagcagtgcagcgtcttcctgagccacaggat
gtcactcagtgcttggaggtccgtgtatttgacacacctcctttttattccaattctacagacag
ttttcgaaatacagtggaaggttacagtgctcccacgggaaaatatgaccctgctgttcgaagcc
ttcacaacctggcccacctcttcctgaatggaacgggaggacaaacccatttgtctcccaatgat
cctatttttgtcctcctgcacactttcactgatgcggtctttgacgaatggctaaggaggtataa
cgccgatatttctacctccccgttggaaaacgcacctattggacataacaggcaatacaacatgg
tgccattctggcctccagttaccaacacagaaatgtttgttactgctccagacaatctgggatat
gcttatgaagttcaatggccaggtcaggagtttactgtatctgaaatcattaccattgctgtagt
ggctgcgttgttacttgtagctgccattttcggggttgcttcttgtctgatccgttctagaagca
ccaagaatgaagccaaccagcctctcctcactgatcactatcaacgctatgctgaggactatgag
gagctcccgaatcctaaccactccatggtctgatagccgcctgcccatccccgcacgcagcaggt
tgggaagcaacctgattgaaacacagctgtcaaccgtatttcctttggcttccttcttttcttcc
caagatagatgtggtttggaattaaagcat...
```

Son produit : une protéine nécessaire pour fabriquer le pigment

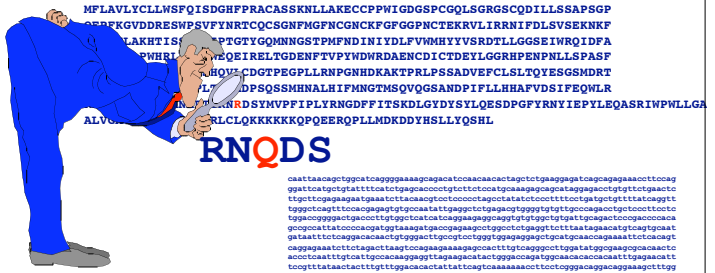
```
MFLAVLYCLLWSFQISDGHFPRACASSKNLLAKE
CCPPWIGDGSPCGQLSGRGSCQDILLSSAPSGP
QFPFKGVDDRESWPSVFYNRTCQCSGNFMGFNCG
NCKFGFGGPNCTEKRVLIRRNIFDLSVSEKNKF
FSYLTLAKHTISSVYVIPTGTYGQMNNGSTPMFN
DINIYDLFVWMHYYVSRDTLLGGSEIWRQIDFA
HEAPGFLPWHRLFLLLWEQEIRELTGDENFTVPY
WDWRDAENCDICTDEYLGGRHPENPNLLSPASF
FSSWQIICSRSEEYNSHQVLCDGTPEGPLLRNPG
NHDKAKTPRLPSSADVEFCLSLTQYESGSMDRT
ANFSFRNTLEGFASPLTGIADPSQSSMHNALHIF
MNGTMSQVQGSANDPIFLLHHAFVDSIFEQWLR
RHRPLLEVYPEANAPIGRNRDSYMVPFIPLYRNG
DFFITSKDLGYDYSYLQESDPGFYRNYIEPYLE
QASRIWPWLLGAALVGAVIAAALSGLSSRLCLQK
KKKKKQPQEERQPLLMDKDDYHSLLYQSHL
```


Relation génotype - phénotype :

Un mutant du gène tyrosinase

Relation génotype - phénotype :

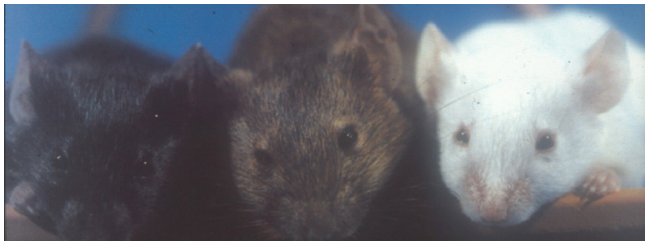
La protéine modifiée (mutante)

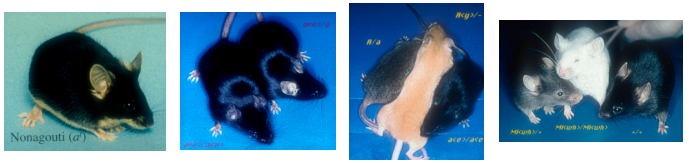


La mutation dans le gène cag

Relation génotype - phénotype :

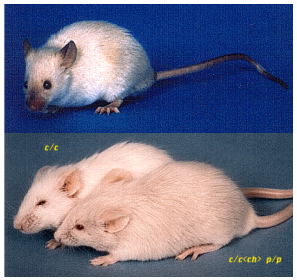
Environ 120 gènes nécessaires pour une couleur de pelage « normale »





Relation génotype - phénotype :

Et il n'y a pas que les gènes …



Froid

Chaud

Les mutations :

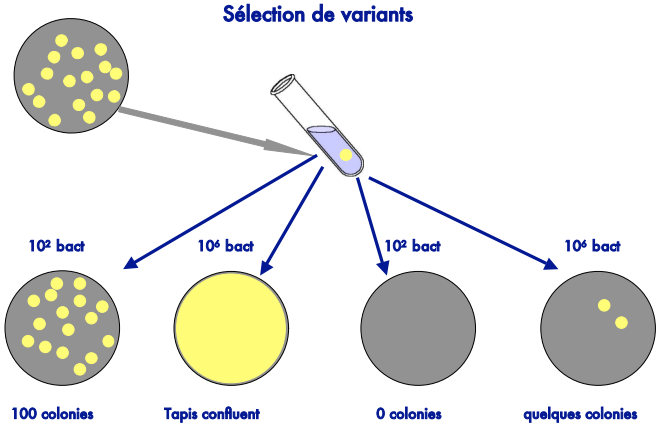
- Modifications irréversible de l'information génétique
- Caractère héréditaire (sauf pour le soma)
- A l'origine de la diversité biologiques des individus et des espèces

---

- Mutations ponctuelles
- Réarrangements chromosomiques

- Spontanées (chimie de l'ADN, dépurination, déamination)
- Induites par des agents mutagènes (Radiations, produits chimiques)
- Issues du métabolisme de l'ADN (Réplication, Recombinaison)

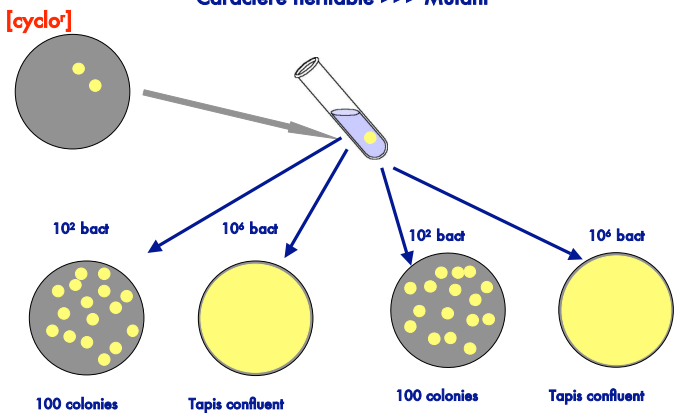
Sélection de variants



min + glu

min + glu + cyclo

Caractère héritable >>> Mutant



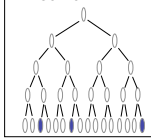
min + glu

min + glu + cyclo

La mutation préexiste à l'agent de sélection (2)

Mutation induite

CULTURE 1



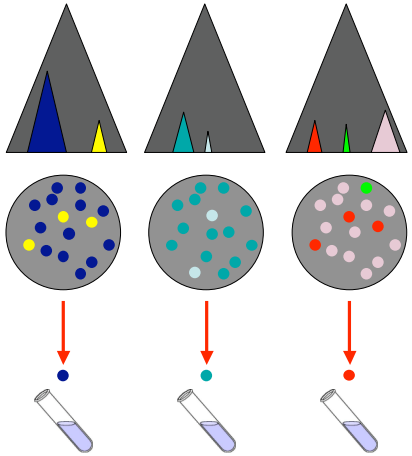
Mutation préexiste

La mutation préexiste à l'agent de sélection (2)

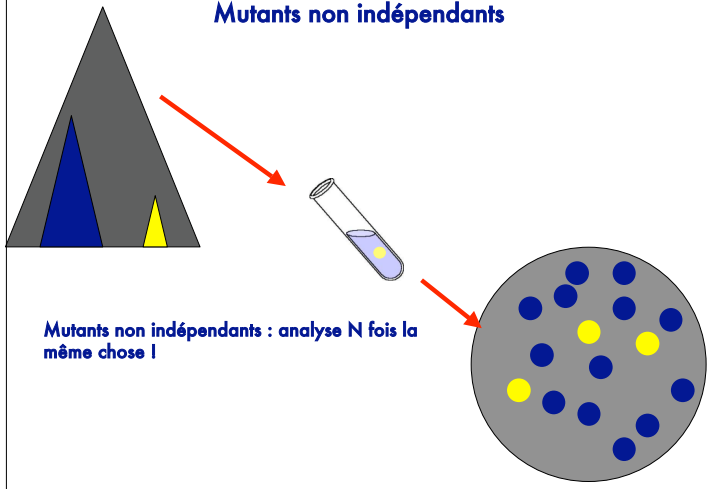
	Separate cultures		Large culture	
	Culture number	Nb résistants	Sample number	Nb résistants
	1	1	1	14
	2	0	2	15
	3	3	3	13
	4	0	4	21
	5	0	5	15
	6	5	6	14
	7	0	7	26
	8	5	8	16
	9	0	9	20
	10	6	10	13
	11	107		
	12	0		
	13	0		
	14	0		
	15	1		
	16	0		
	17	0		
	18	64		
	19	0		
	20	35		
mean		**11.3**		**16.7**
variance		**694**		**15**
variance / mean		**60.8**		**0.9**

Data from Luria, S. and M. Delbruck. 1943. Genetics 28: 491.

Mutants indépendants



Mutants non indépendants



Mutants non indépendants : analyse N fois la même chose !

Fréquence de mutants vs **Taux de mutation:**

- freq : proportion de mutants dans une population

- taux de mutation : événements par cellule et par division

=> test de fluctuation Luria et Delbruck : x cultures en parallèle

- Evènements rares : (loi binomiale fréquence p sur n divisions) qu'on approxime par une loi de Poisson (p<<0,1, n>50) ($P(X=k)=e^{-\lambda} \times \lambda^k/k!$ avec $\lambda = np$)

Fréquence (classe k=0) = e^{-np} (classe 0 = fréquence cultures sans révertants)

p = -ln(Fréquence classe 0)/n

taux de mutation: p

$$p = 1 - e^{Inf_0/n_{div}}$$

p # 10^{-7}/cellule/génération

p # 10^{-10}/base/génération

FIN DU 1er COURS

des question?