

Traitement de l'information génétique



"Le code de la vie (*L'organisme est inscrit dans l'œuf*)"
James Gleick

Plan

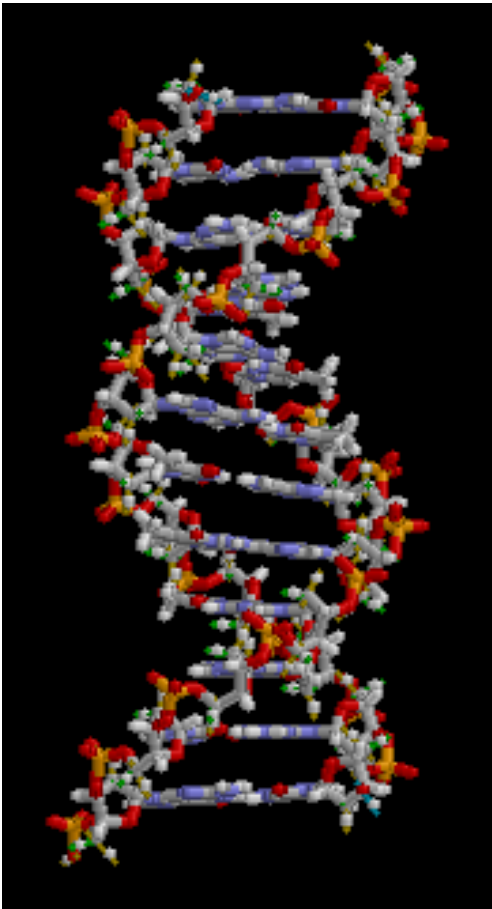
- La découverte du code génétique
 - Rappels : de l'ADN aux protéines...
 - Code génétique théorique vs. expérimental ?
 - Le « dogme central » de la génétique et ses limitations
- Le contrôle du développement embryonnaire
- Epigénétique
 - Définition
 - Exemples d'hérédité épigénétique
- Autres exemples d'information biologique
- Ce qu'il faut retenir...

Découverte du code génétique

Un pionnier : Erwin Schrödinger

- Dans son livre « *What is Life* » publié en 1944, Schrödinger écrit :
 - « *La fibre chromosomique contient, chiffrée dans une sorte de **code miniature**, tout le devenir d'un organisme, de son **développement**, de son **fonctionnement*** »
- Il était ainsi le premier à poser le problème central de la biologie en termes d'**information codée**
- Selon les mémoires de James Watson (« *L'ADN, le secret de la vie* »), le livre de Schrödinger lui a donné l'inspiration pour rechercher le gène, ce qui l'a conduit à la découverte de la structure en double hélice

La découverte de la double hélice

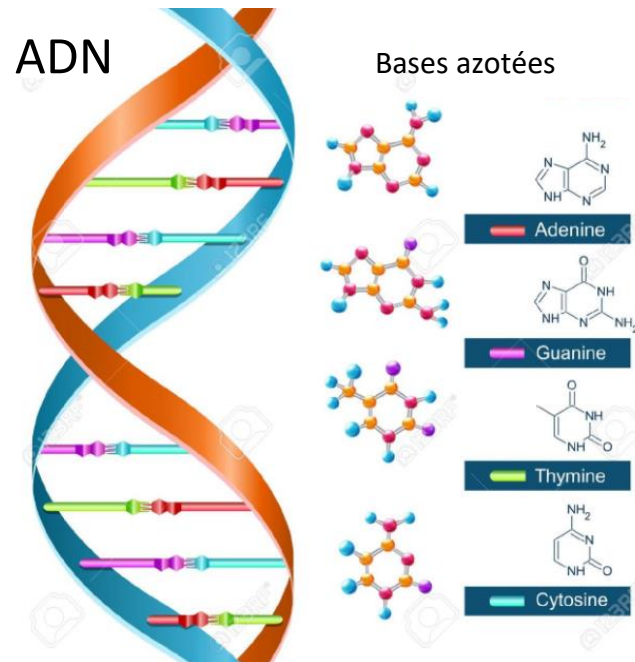
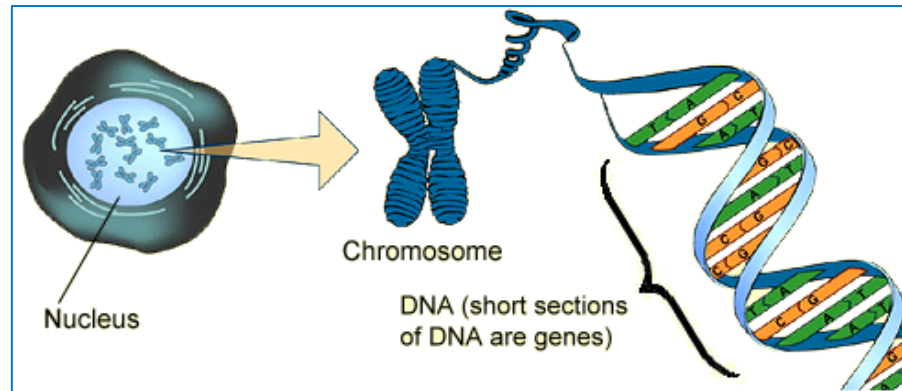


- Au début des années 50, *l'ADN n'est plus une inconnue*. Les biologistes de l'époque ont compris que cette grosse molécule était le *support de l'hérédité*
- Ils savent aussi qu'elle se compose de *quatre types de molécules* plus petites, appelées *nucléotides* ou *bases* et qui se distinguent par leur base azotée : *A (adénine), T (thymine), C (cytosine) et G (guanine)*
- Ce qu'ils ignorent totalement c'est sa *forme*. *Crick et Watson* vont montrer que l'ADN possède une structure hélicoïdale en forme de *double hélice*. Au sein de cette double hélice les bases s'apparient par *paires: A-T et C-G*
- Ces caractéristiques permettront de comprendre comment *se copie et se transmet l'information génétique*

Rappel : de l'ADN aux protéines

☞ Résumé de ce qui allait être découvert...

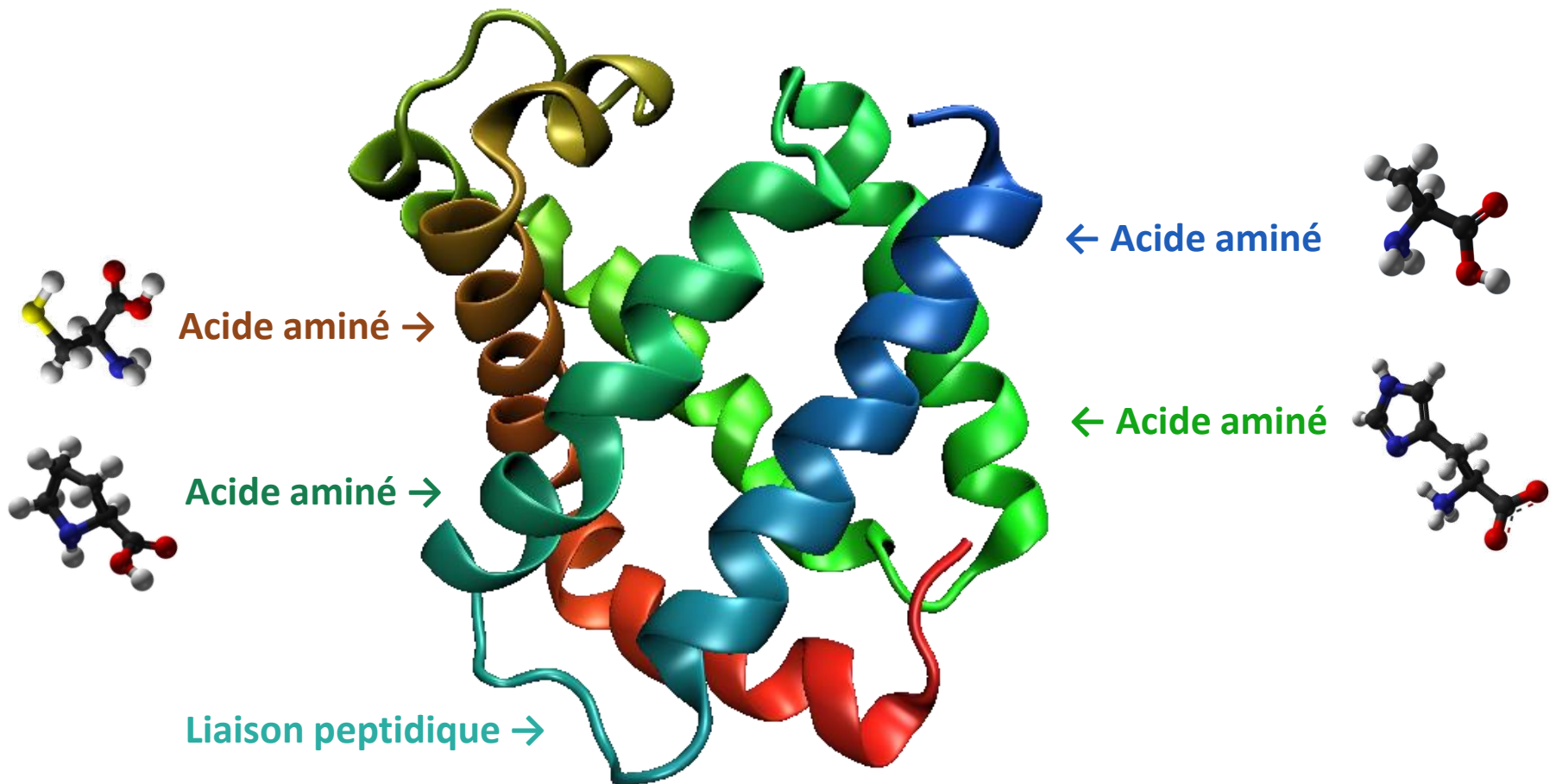
Gène = fragment d'ADN



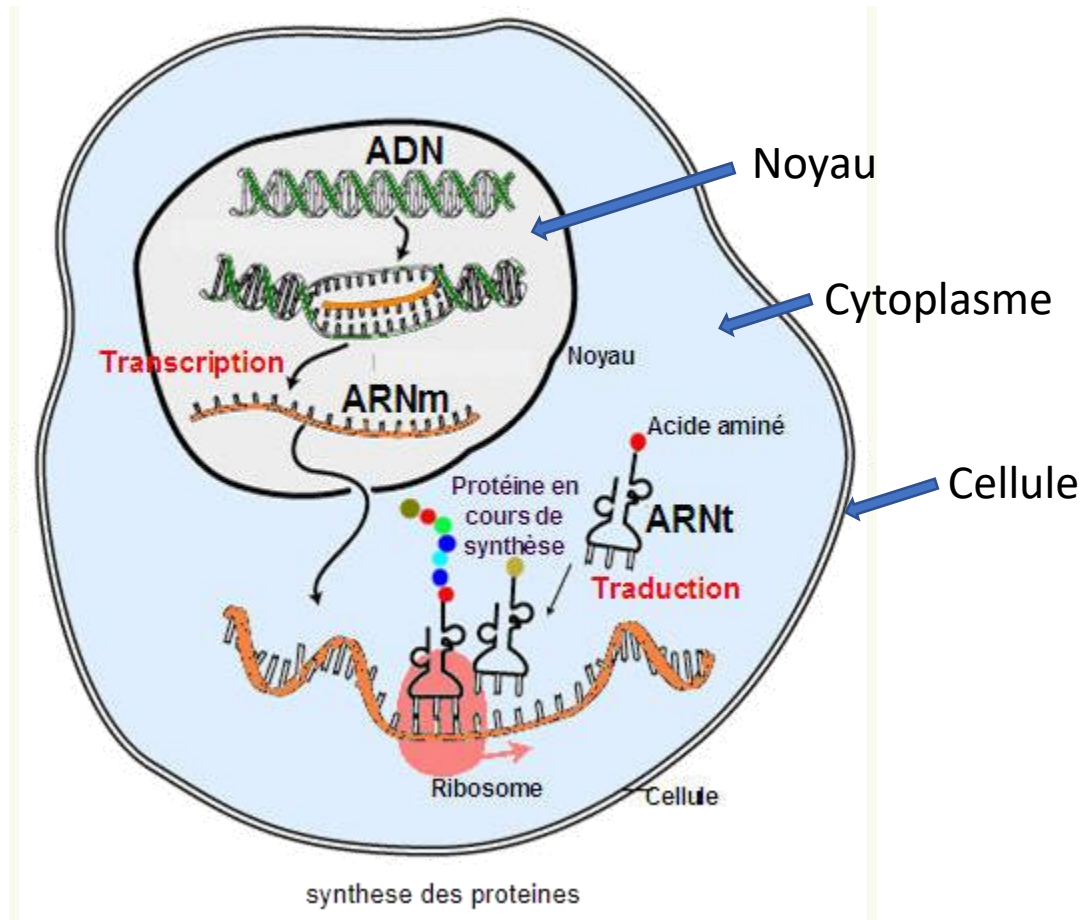
ADN humain :
Plusieurs milliards de bases
 ± 2 m (déroulés) par cellule

Structure d'une protéine

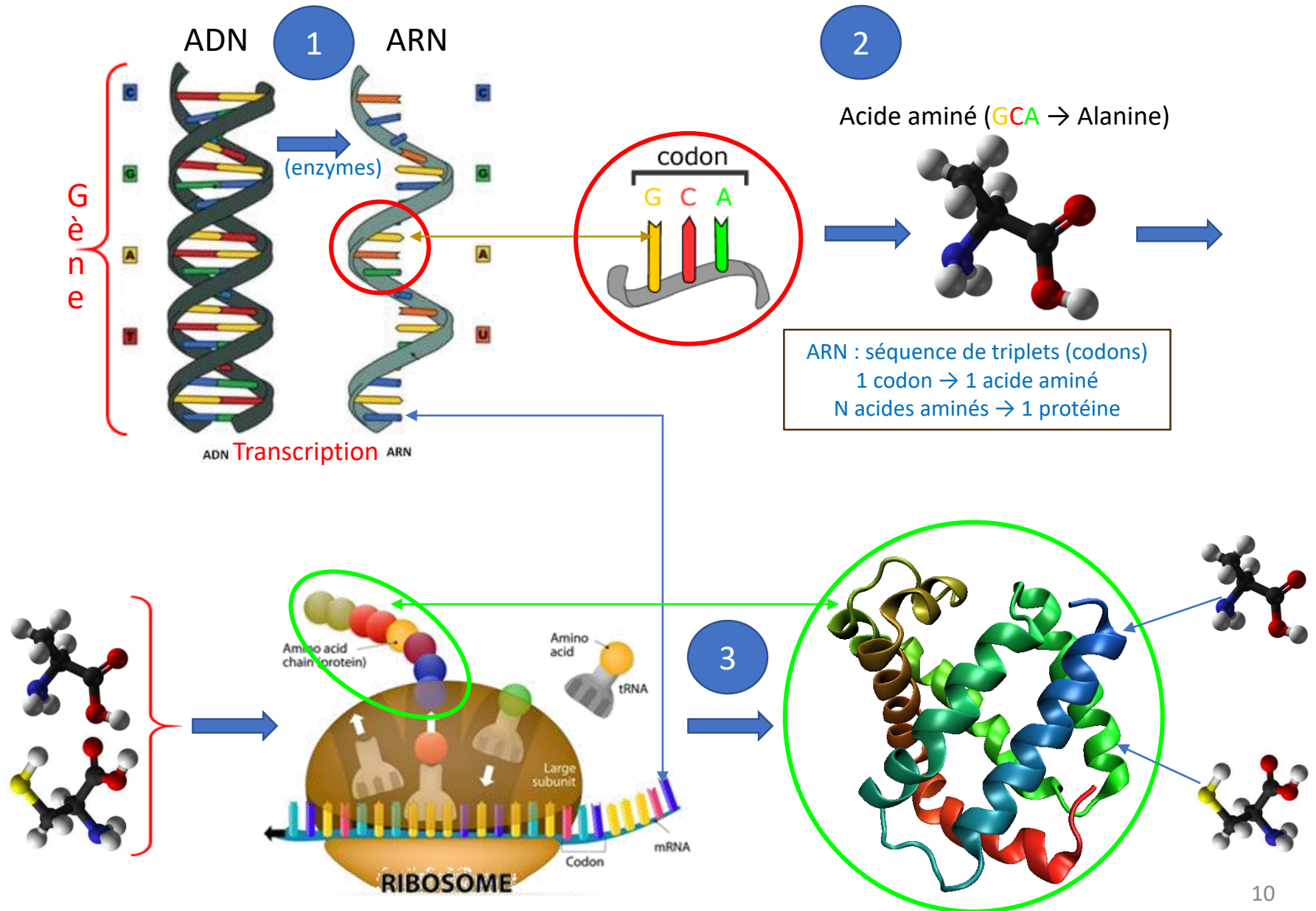
Une protéine = une séquence d'acides aminés, enroulée en 3D



Du gène à la protéine (résumé)

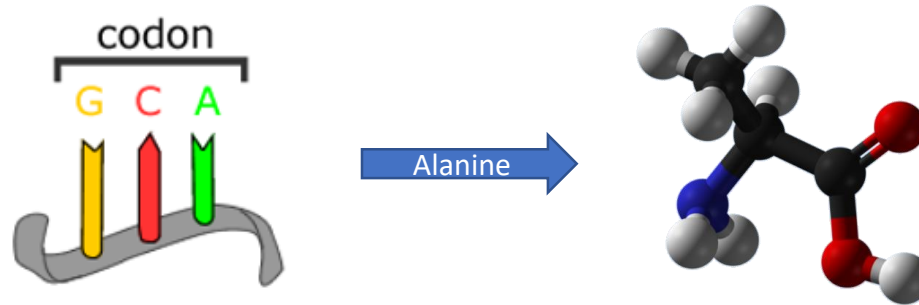


Du gène à la protéine, en 3 étapes



Etape 2. De l'ARN aux acides aminés

(Le « code génétique », raconté ci-après)



		Deuxième acide aminé du codon					
		U	C	A	G		
Premier acide aminé du codon	U	Phé	Ser	Tyr	Cys	U	Troisième acide aminé du codon
		Phé	Ser	Tyr	Cys	C	
		Leu	Ser	STOP	STOP	A	
		Leu	Ser	STOP	Trp	G	
	C	Leu	Pro	His	Arg	U	
		Leu	Pro	His	Arg	C	
		Leu	Pro	Gln	Arg	A	
		Leu	Pro	Gln	Arg	G	
	A	Ile	Thr	Asn	Ser	U	
		Ile	Thr	Asn	Ser	C	
		Ile	Thr	Lys	Arg	A	
		Met	Thr	Lys	Arg	G	
	G	Val	Ala	Asp	Gly	U	
		Val	Ala	Asp	Gly	C	
		Val	Ala	Glu	Gly	A	
		Val	Ala	Glu	Gly	G	

Phé	Phénylalanine
Leu	Leucine
Ile	Isoleucine
Met	Méthionine
Cys	Cystéine
Trp	Tryptophane
Val	Valine
Thr	Thréonine
Pro	Proline
Tyr	Tyrosine
Ala	Alanine
Ser	Sérine
Gly	Glycine
His	Histidine
Gln	Glutamine
Arg	Arginine
Asn	Asparagine
Lys	Lysine
Asp	Aspartate
Glu	Glutamate

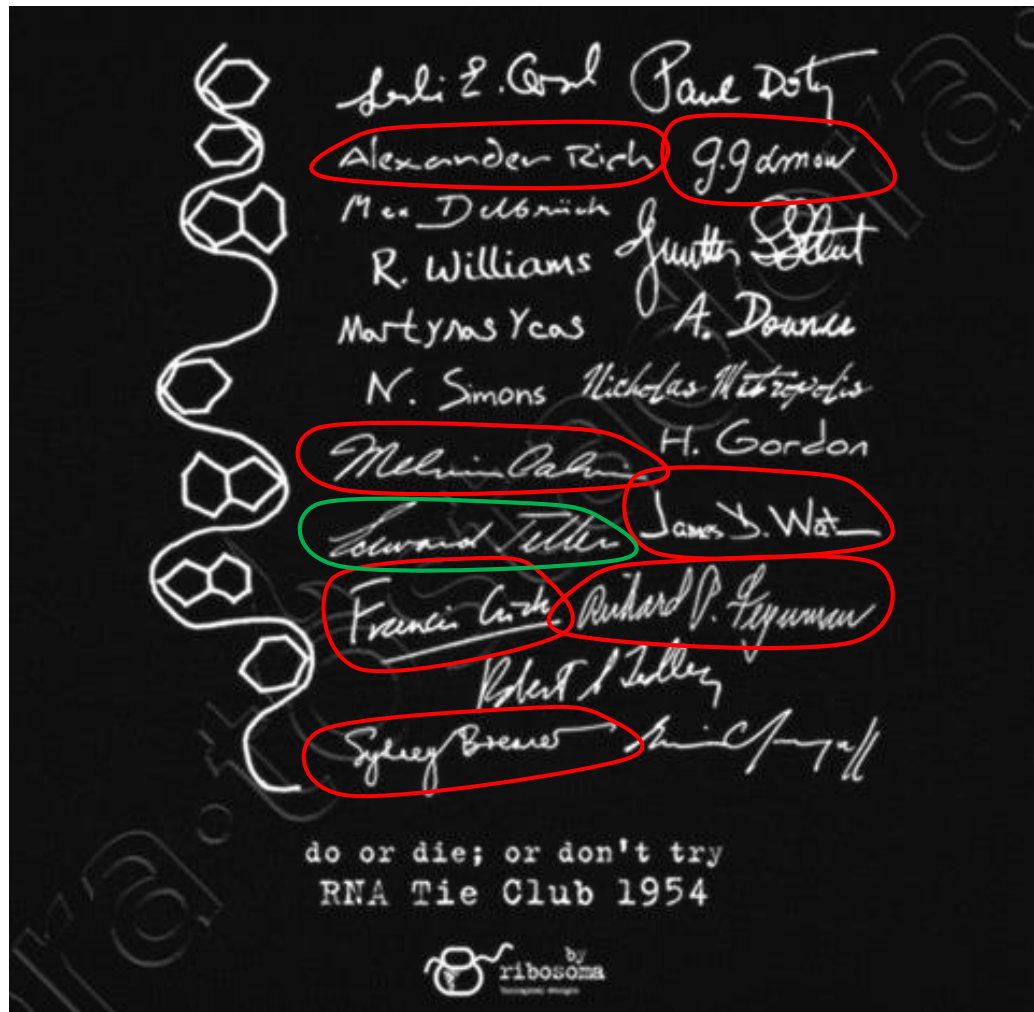
A la découverte du code génétique

Le RNA Tie Club

- Les premières recherches sur le décodage du code génétique ont concerné l'étape 2 de *traduction de l'ARN « messenger » vers les acides aminés*
- Cette recherche a été impulsée en 1954 par *James Watson*, rapidement rejoint par *George Gamow* que le challenge mathématique intéressait
- Ne disposant d'aucune donnée expérimentale, Gamow a proposé de fonder le « *RNA Tie Club* », composé de 20 membres (*comme le nombre d'acides aminés*), raisonnant à un niveau *purement théorique*
- Sur ces 20 membres, 8 avaient ou devaient obtenir le prix Nobel
- Malgré certaines avancées (*comme la proposition par Gamow des triplets de bases ou codons*), ces recherches, faute de support expérimental, devaient *échouer*
- Mais « en même temps », *Marshall Nirenberg*, qui ne faisait pas partie du Club, entreprenait des travaux expérimentaux qui, eux, devaient *aboutir*

Le RNA Tie Club

Une idée de George Gamow



\$4.00 each
(payable to George Gamow)

Le code génétique effectif

		Deuxième acide aminé du codon				
Premier acide aminé du codon	a	U	C	A	G	Troisième acide aminé du codon
		Phé	Ser	Tyr	Cys	
U		Phé	Ser	Tyr	Cys	U
		Phé	Ser	Tyr	Cys	C
		Leu	Ser	STOP	STOP	A
		Leu	Ser	STOP	Trp	G
C		Leu	Pro	His	Arg	U
		Leu	Pro	His	Arg	C
		Leu	Pro	Gln	Arg	A
		Leu	Pro	Gln	Arg	G
A		Ile	Thr	Asn	Ser	U
		Ile	Thr	Asn	Ser	C
		Ile	Thr	Lys	Arg	A
		Met	Thr	Lys	Arg	G
G		Val	Ala	Asp	Gly	U
		Val	Ala	Asp	Gly	C
		Val	Ala	Glu	Gly	A
		Val	Ala	Glu	Gly	G

		b
	Hydrophobe	
Phé		Phénylalanine
Leu		Leucine
Ile		Isoleucine
Met		Méthionine
Cys		Cystéine
Trp		Tryptophane
Val		Valine
Thr		Thréonine
Pro		Proline
Tyr		Tyrosine
Ala		Alanine
Ser		Sérine
Gly		Glycine
His		Histidine
Gln		Glutamine
Arg		Arginine
Asn		Asparagine
Lys		Lysine
Asp		Aspartate
Glu		Glutamate
	Hydrophile	

- *Plus de 10 ans* après le RNA Tie Club, le « véritable » code génétique a été découvert en *1966* par trois *chercheurs expérimentaux* *, qui en ont été récompensés par le prix Nobel
- Ce code semble *dépourvu de toute logique*, mais à y regarder de plus près, il est *étonnamment robuste*. Ainsi, lorsqu'on s'intéresse aux propriétés des acides aminés, ici leur affinité pour l'eau, on se rend compte qu'une substitution d'un acide par un autre, suite à la *mutation d'une base*, n'entraîne que peu de changements, car les deux molécules auront la *même affinité pour l'eau*. *Laurence Durst et Stephen Freeland* ont étudié cette propriété pour *un million de codes* créés au hasard : *un seul est plus robuste* que celui mis au point par la nature.

L'ADN, résultat d'une évolution ?

- Cette « robustesse » et cet « illogisme » sont typiques du résultat d'une *évolution darwinienne* par *essais-erreurs* : les variations se font au hasard, et les éléments les moins robustes sont *éliminés* par la sélection naturelle
- Cela semble confirmer que le code génétique est bien le fruit d'une évolution, plutôt qu'une donnée *a priori* conçue par un esprit supérieur, ou une équipe de savants
- Une branche de la recherche s'intéresse à cette l'évolution ; une hypothèse est que l'ARN ait évolué avant l'ADN
- *NB : Sur l'évolution naturelle du code génétique, on pourra consulter le récent livre de Michael Yarus : "Life from an RNA World: The Ancestor Within" (Harvard University Press, 2010).*

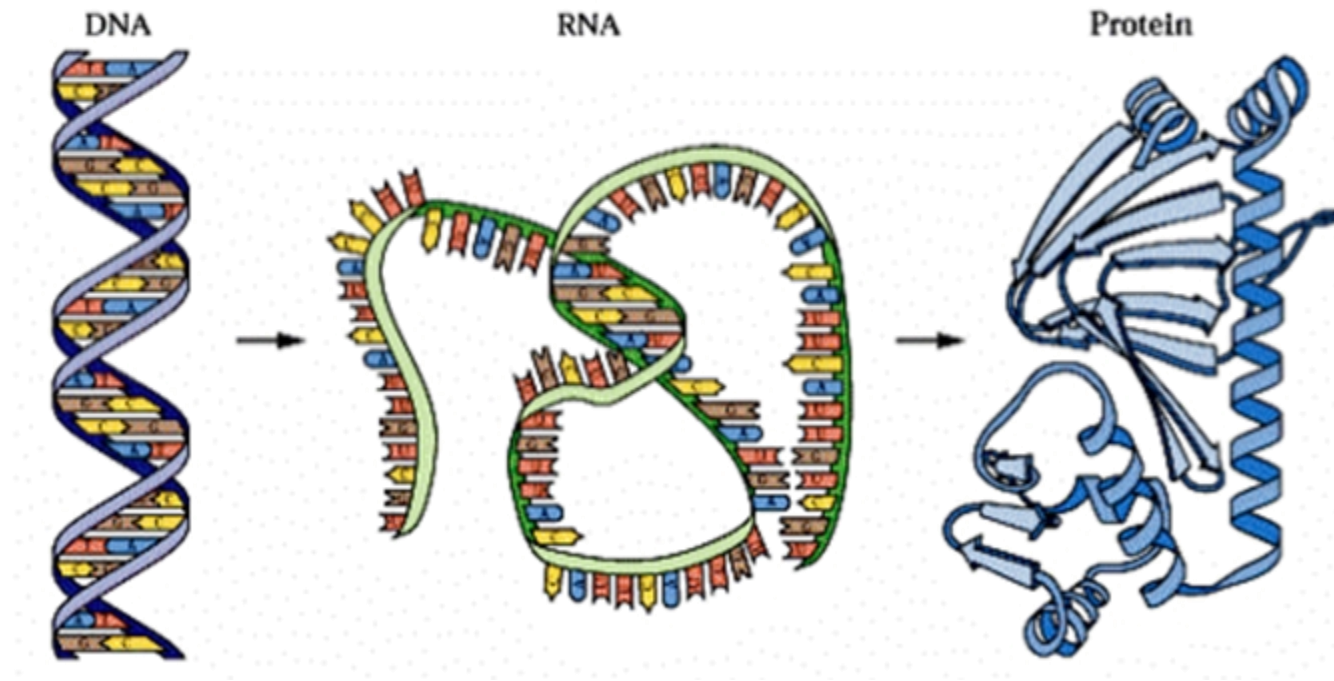
Conception humaine vs. système naturel

<i>Un système récemment conçu par l'homme</i>	<i>Un système issu de l'évolution naturelle</i>
• Top-Down	• Bottom-Up
• Beau, élégant	• Bricolage, usine à gaz
• Adapté à une fonction préétablie	• Adapté à sa propre survie à LT
• Efficace dans un environnement déterminé	• Efficace dans un environnement variable
• Simple, simulable sur ordinateur	• Complexe, trop pour être simulé
• Composants bien identifiés et relativement peu nombreux	• Enormément de composants divers
• Pas toujours très flexible	• Robuste, auto-organisé
• Il faudra attendre les versions suivantes (évolution future nécessaire)	• Déjà adapté (évolution passée, sélection)

Le « dogme central » de la génétique et ses limites

Le « dogme central » de la génétique

(Rappel)



Le dogme central

Tout ne s'arrête pas là...

Une citation du Pr. Michel Morange, auteur du livre : « La vie expliquée, 50 ans après la double hélice » (2010)

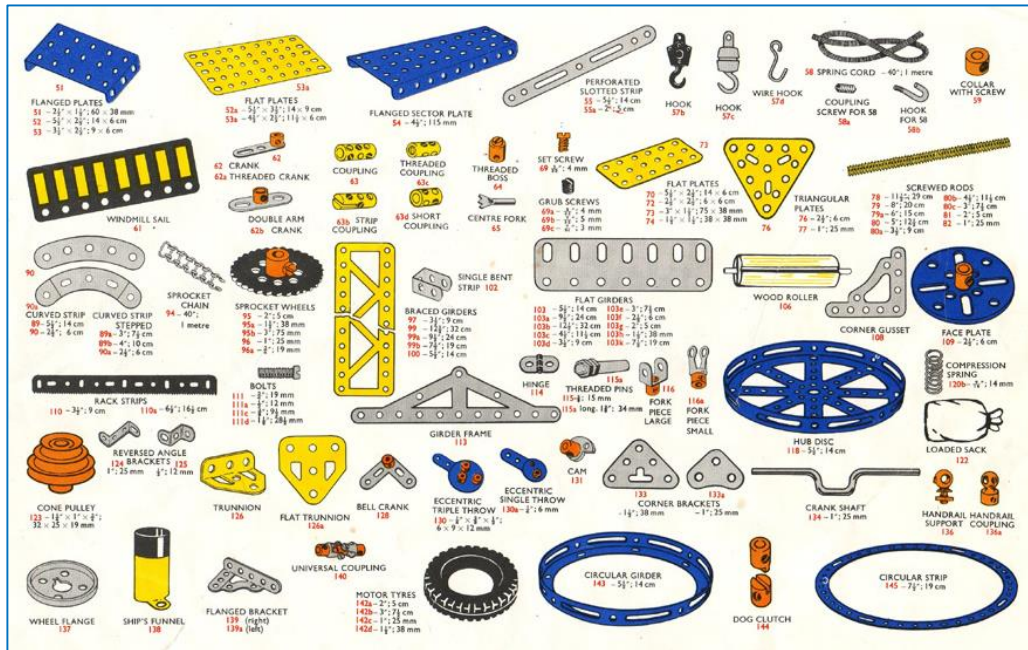
« Il y a cinquante ans, Francis Crick et James D. Watson découvraient la structure en double hélice de l'ADN, constituant des gènes, support de l'hérédité. Ils pensaient avoir ainsi, selon l'expression même de Francis Crick, "*percé le secret de la vie*".

[Mais il convient] de prolonger cette histoire *au-delà de la double hélice* et d'en interpréter les évolutions récentes par les transformations intervenues en biologie pendant les *cinquante dernières années*

Ces transformations sont *souvent passées inaperçues auprès du grand public*, comme si la biologie moléculaire était *restée figée autour de la double hélice* » M. M.

Limites du dogme central

L'organisme est un système complexe
Les composants individuels ne le définissent pas



Les protéines

*Dans un système complexe, la structure et les interactions
sont plus importantes que les composants pris séparément*



Gènes codants et gènes « poubelle »

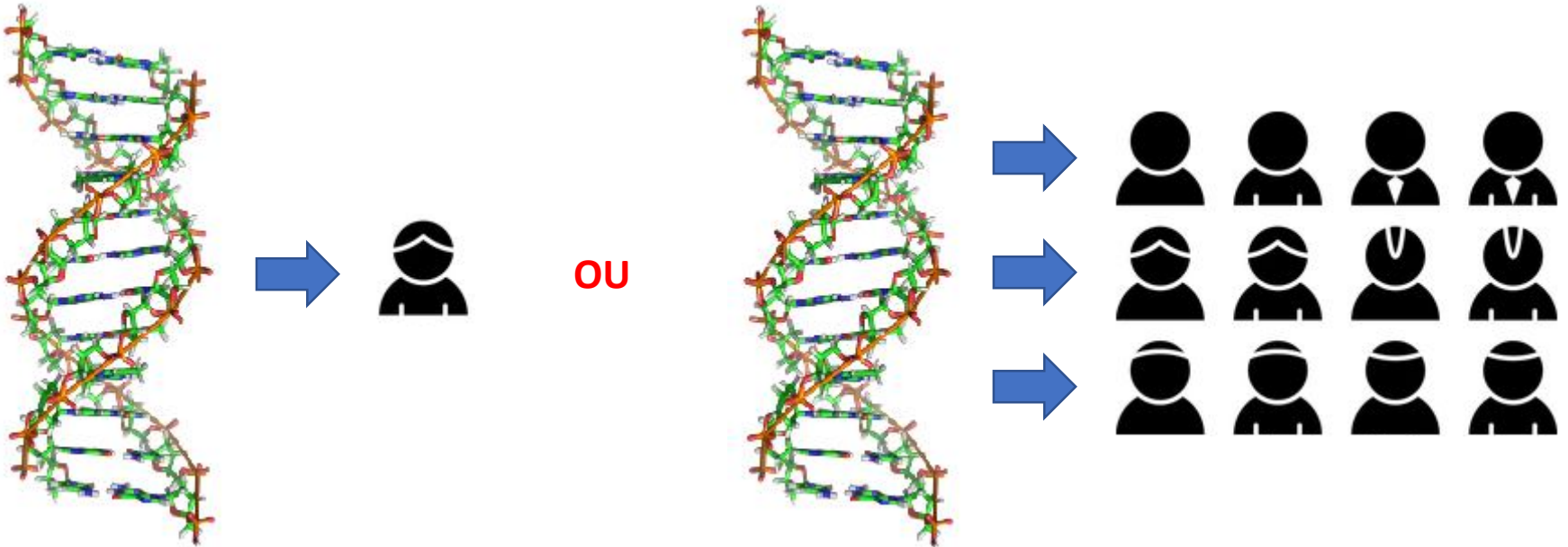
- *Seulement 3 à 10 %* des gènes *codent effectivement* pour une protéine
- On a longtemps dénommé les autres par l'appellation péjorative de « *gènes-poubelle* » (1972 ++, encore en 2000...)
- Or ces sections non codantes de l'ADN contribuent à définir la *structure* de nos organismes et le *fonctionnement* de nos organes !
- C'est dire *l'illusion* qui a longtemps prévalu que *la seule fonction « intéressante »* du génome était de *fabriquer des protéines*
- C'était dans l'enthousiasme qui a suivi la découverte de la *double hélice* et du « *dogme central* » (années 60, 70...)
- Mais des spécialistes en *systémique* ne sauraient se contenter d'un rôle aussi limité...

Ce que le « dogme central » *n'a pas résolu*

- Toutes les cellules d'un organisme portent *le même génome*
 - Or elles ne synthétisent pas toutes *les mêmes protéines*
 - Leur « *expression* » dépend de plusieurs facteurs
 - Le *temps* (*le stade d'avancement de l'embryon*)
 - Le *lieu* (*telle partie de l'organisme, tel tissu...*)
 - La *forme* à construire (*un nez, un cœur, un vaisseau...*)
 - Suivant ce qu'ils *ont à faire*, les mêmes gènes s'expriment différemment
- *Comment les gènes « savent-ils » ce qu'il doivent faire précisément, à chaque instant et dans chaque lieu ?*

Une autre question se pose

- « *L'organisme est-il inscrit dans l'œuf* » ?
- **OU** : plusieurs phénotypes peuvent-ils résulter d'un même génotype ?





Le contrôle du développement

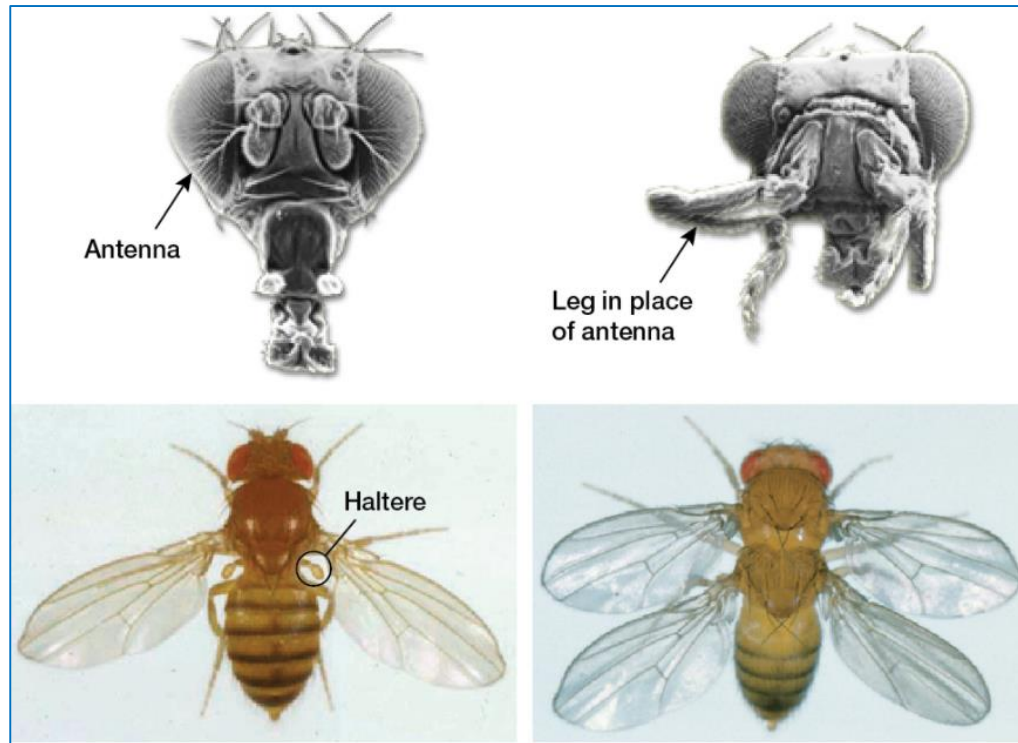


Protéines et gènes de régulation

- La *régulation d'un gène* est assurée par des « *facteurs de transcription* » présents dans *l'environnement*
 - Des *protéines « régulatrices »*, dont la *présence* et la *concentration stimulent, ralentissent ou bloquent* l'expression de certains gènes
 - D'autres facteurs présents dans *l'environnement cellulaire* jouent le même rôle :
*Agents chimiques ou organiques, température, etc.**
- Des *gènes de régulation* codent pour ces *protéines régulatrices*
 - Mais *eux-mêmes sont contrôlés* par leur environnement !
 - Il s'agit d'un *système d'information* extrêmement *complexe* (*nombreux agents, nombreuses interactions, réactions en cascade, rétroactions...*)

Gènes homéotiques et plan d'organisation

- Les *gènes homéotiques* sont des gènes qui déterminent le *plan d'organisation* des êtres vivants, c'est-à-dire la *place des organes* les uns par rapport aux autres : *HOX, HOM...*



- Certains *déplacements, duplications* ou *mutations* de gènes homéotiques dans l'embryon modifient la *structure* du futur organisme
- Ici une drosophile *Antennapedia* et une *Bithorax*

L'épigénétique

L'épigénétique

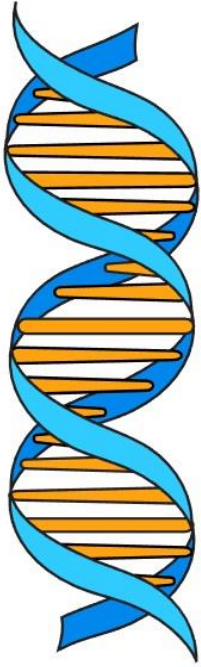
- Selon Wikipedia :

- L'épigénétique (*du grec ancien ἐπί, « au-dessus de », et de génétique*) est la discipline de la **biologie** qui étudie les *mécanismes moléculaires* qui *modulent l'expression* du **patrimoine génétique** en fonction du *contexte*
- Alors que la génétique correspond à l'étude des **gènes**, l'épigénétique s'intéresse à une *couche d'informations complémentaires* qui définit comment ces gènes vont être utilisés par une cellule ou... ne pas l'être
- C'est un concept qui dément en partie la « *fatalité* » des *gènes* (*Morange*)
- Il s'agit en définitive du *rôle de l'environnement cellulaire* (*protéines, température, agents chimiques...*) sur *l'expression des gènes* de chaque cellule
- Mais, selon la manière dont elle intervient, l'épigénétique va avoir des *implications philosophiques* très différentes :

Trois perceptions de l'épigénétique

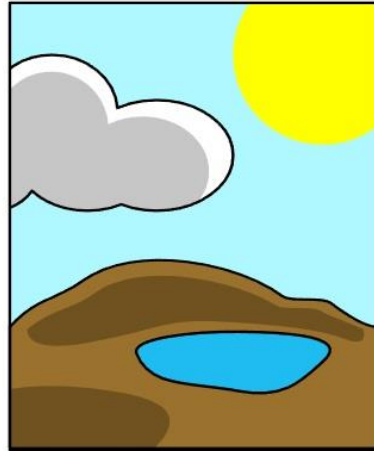
- L'« **épigénétique** » caractérise, de manière **très générale**, le rôle de **l'environnement** sur l'expression des gènes ; mais la perception de ce rôle a évolué avec la science, selon **3 hypothèses successives** :
 1. Un environnement **déterministe**, **contrôlé** par des **gènes de régulation** (**avant 1990**)
 2. Un environnement comportant **une part de variations** : ces variations peuvent dépendre du **milieu de vie** (**placenta ou écosystème**) ou du **comportement de l'organisme** après la naissance ; selon cette 2^{ème} hypothèse, il n'y a **pas de transmission** de ces variations à la descendance (**depuis 1990 env.**)
 3. Dans certains cas, mis en évidence récemment (**années 2000+**), les **variations acquises** d'un environnement cellulaire peuvent être **transmise de la mère à l'enfant** sur une ou parfois **plusieurs générations**
- Les implications « **philosophiques** » de chaque hypothèse :
 1. **Tout est prévu dans le génome**, les gènes de régulation façonnent un phénotype **unique** : « **l'organisme est inscrit dans l'œuf** » (**sous strict contrôle épigénétique**)
 2. Le génome **ne définit plus l'individu de manière univoque** : plusieurs phénotypes peuvent être issus d'un même génotype : **le milieu et le hasard** jouent un rôle
 3. Les cas de **transmission** de la mère à l'enfant de **caractère acquis** pendant la vie peuvent faire évoluer la théorie **néodarwinienne** en y intégrant des éléments de « **lamarckisme** » (**domaine de recherche actif**)

Résumé



Genotype

+



Environment

=



Phenotype

Quelques exemples...

- Différentiation cellulaire
- Plan d'organisation de l'embryon
- Effets à court terme (*horloge circadienne*)
- Effets à long terme (*irréversible pour un individu*)
 - *Susceptibilité à certaines maladies (cancer, diabète, Alzheimer...)*
 - *Comment l'abeille devient reine...*
- « **Epimédicaments** » (*une voie de recherche active*)
 - *Ils agissent sur l'environnement des gènes (ex : sur la méthylation)*
 - *Certains agissent sur l'organisme traité (ex : cancer ?)*
 - *D'autres n'agissent que sur sa descendance*
- Hérité sur une ou plusieurs générations

Exemples d'« hérédité » épigénétique

Des souris obèses...



- Des souris hypernutries dans les premiers jours de la vie restent *obèses toute leur vie*, même si le régime *redevient normal*, et peuvent même *donner naissance à des souriceaux qui deviendront obèses*
- Ce dernier cas est un exemple *d'hérédité de caractères acquis ! (mais pas souhaités !)*

Le gène « agouti » (1998)



- Un autre exemple est celui du *gène agouti*, impliqué dans la détermination de la *couleur du pelage chez la souris* : dans un groupe d'animaux portant tous la *même version* de ce gène, certains ont un pelage *brun chiné* et d'autre un pelage *jaune*. Ces derniers ont en outre une susceptibilité accrue à *l'obésité*, au *diabète* et à certains *cancers*
- Qu'est-ce qui les différencie ? Il ne s'agit *pas d'une mutation* affectant la séquence de leur ADN, mais bien d'une *marque épigénétique* portée par les souris brunes, qui *éteint le gène agouti*
- Or on observe que la proportion de souriceaux bruns est plus importante dans la descendance des mères brunes que dans celle des mères au pelage jaune : ceci suggère que les mères brunes peuvent *transmettre à leur descendance* la marque épigénétique (*hyperméthylation*) qui éteint le gène agouti

Chez l'humain : la famine hollandaise

- Des femmes hollandaises *sous-nutries* pendant la guerre (*hiver 1944-45*) ont donné naissance à des *enfants plus petits* que la moyenne
- Cela est dû à une *carence* en certains nutriments, agissant sur l'expression des gènes embryonnaires
- Mais si cela a eu lieu dans ce sens, cela pourrait avoir lieu aussi *dans le sens inverse* :
 - Des femmes bien alimentées pourraient donner naissance à des *générations de taille supérieure*, et cela pourrait se conserver, voire s'amplifier à travers les générations, *sans passer par le patrimoine génétique*...
 - *Une explication de la taille croissante des humains ?*

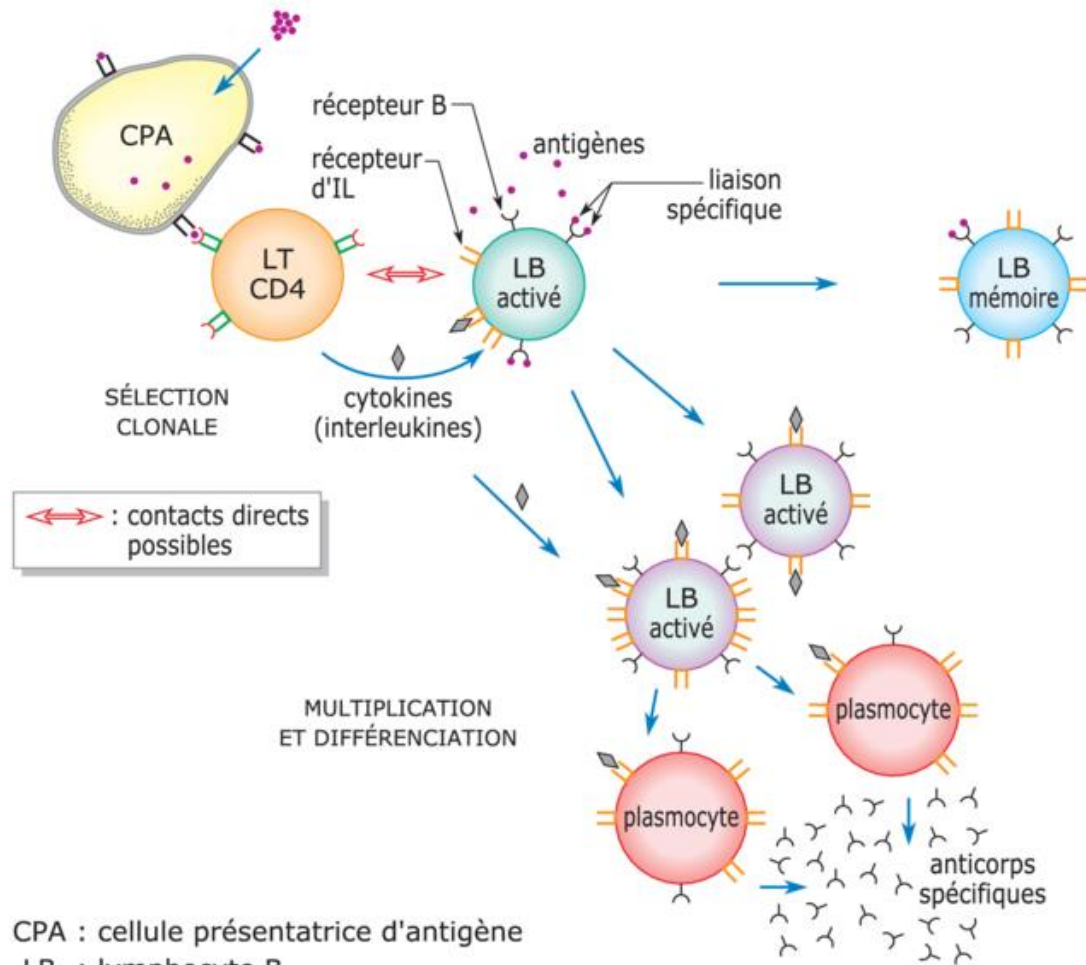
Biologie et Information

Autres systèmes d'informations biologiques

- **Cerveau** : système de traitement de l'information
- **Systèmes nerveux périphériques**
 - Douleur (*système nociceptif*)
 - Moteur
 - Autonome (*sympathique, parasympathique...*)
- **Systèmes sensoriels**
- **Système immunitaire**
- **Système endocrinien...**

Système d'information immunitaire

Stimulation antigénique et production d'anticorps



CPA : cellule présentatrice d'antigène
LB : lymphocyte B
LT : lymphocyte T
IL : interleukine

Ce qu'on pourra retenir...

Résumé

- La grande révolution des années 1950 et 1960 a été la découverte de la *double hélice de l'ADN*, ainsi que les mécanismes de *synthèse des protéines*, les « briques » des organismes vivants : le « *code génétique* » et le « *dogme central de la biologie* »
- L'histoire anecdotique de ces découvertes a permis d'illustrer une base de la méthode scientifique : rien ne peut se faire sans un *fondement expérimental*
- Après une *période de « gel »* grand public sur ces notions, les scientifiques sont partis à la découverte d'un nouveau monde : les mécanismes complexes du *développement embryonnaire*, et *les gènes de régulation*
- Ces mécanismes mettent en œuvre le concept d'*épigénétique* : influence de *l'environnement* cellulaire sur l'expression des gènes
- Les propriétés épigénétiques de l'environnement sont parfois susceptibles d'être *transmises de la mère à l'enfant*, sans toutefois affecter les gènes
- Concernant l'évolution des espèces, cela ouvre de *nouvelles perspectives* (qu'il faut considérer avec *prudence*) par l'introduction d'une certaine dose de « *lamarckisme* »...
- Au fondement de l'ensemble de ces mécanismes biologiques complexes, *l'information* apparaît comme le lien *incontournable* et *omniprésent*

