

Introduction à la Biologie Moléculaire

Timothée Lionnet et Vincent Croquette
<http://pimprenelle.lps.ens.fr/biolps/>

19 mars 2010

Selon Michel Morange, la biologie moléculaire est "l'ensemble des techniques et découvertes qui ont permis l'analyse moléculaire des processus les plus intimes du vivant, de ceux qui en assurent la perennité et la reproduction"[1]. Cette introduction résume de manière historique les événements qui ont contribué à sa naissance et à son développement. On trouve une description bien plus complète des concepts et des techniques classiques de la biologie dans les ouvrages de référence[2, 3] ou de vulgarisation[4, 5].

- ▶ Les processus biologiques sont très anciens :(pain, brassage de la bière, production de fromage...
- ▶ En médecine aussi, utilisation par les chinois d'antibiotiques (extraits de champignons).
- ▶ L'histoire de la biologie moléculaire commence vraiment au milieu du XIX^{eme} siècle, avec l'abandon de la théorie vitaliste et du principe de génération spontanée.

- ▶ Naissance de la microscopie (Van Leeuwenhoek XVII^{ème} siècle); Une tranche de liège (Hooke XVII^{ème} siècle) montre des cellules.
- ▶ élaboration de la théorie cellulaire (Schleiden & Schwann 1839)
- ▶ La génération spontanée invalidée par Francesco Redi (1668). Mais la croyance l'emporte et il faut attendre Pasteur en 1862 pour mettre à bas cette théorie.
- ▶ Les liens entre micro-organismes et contagion sont observés mais non acceptés. (Semmelweis Vienne 1846). Koch observation de germes d'anthrax en 1870.
- ▶ Lien entre micro-organismes et fermentation par Pasteur. Il met au point la pasteurisation (1863) pour éliminer les ferments parasites.

- ▶ Le principe de la catalyse est découvert (1812).
- ▶ La synthèse de l'urée (Wöhler 1828), contredit la théorie vitaliste.
- ▶ la première enzyme est extraite du malt : la diastase catalyse amidon/glucose (Payen & Persoz 1830).
- ▶ fermentation des sucres (extrait acellulaire de levure Büchner 1897).
- ▶ Reconnaissance spécifique enzyme substrat (Ehrlich 1897);
- ▶ Echelle de pH, solutions "tampon"
- ▶ 1902-1909 Fonctionnement des enzymes (Henri, Michaelis & Menten)
- ▶ Première moitié du XX^{ème} siècle : Grandes voies métaboliques : glycolyse, cycle de l'urée, cycle de Krebs
- ▶ Découverte des macromolécules (1922 : Hermann Staudinger)
- ▶ Technique de séparation et d'analyse : Chromatographie, précipitation par des sels ultracentrifugation (1924)
- ▶ électrophorèse (1930-1945)

- ▶ Les "lois de l'hérédité" (Mendel 1866) . Cette théorie passe totalement inaperçue.
- ▶ observation de la chromatine (1879), des chromosomes et de la mitose (1882) par Fleming.
- ▶ de Vries redécouvre en 1900 la théorie de Mendel. Il observe l'apparition de mutations spontanées et propose un mécanisme d'évolution par mutations.
- ▶ Johannsen introduit les concepts de génotype (sans le lier à l'ADN) et phénotype (1909).
- ▶ Expériences de Morgan (1913) sur la transmission de caractères et de mutations chez la drosophile : preuve de l'existence des gènes et de leur localisation chromosomique (hérédité liée au sexe et première carte génétique).

Expériences fondatrices. Quelques dates

- ▶ la relation un gène - une enzyme : 1902 Garrod ; 1941 Beadle et Tatum ;
- ▶ 1943 : la naissance de la génétique bactérienne : le "test de fluctuation" (Luria et Delbrück). Il consiste à analyser comparativement des populations de bactéries infectées par un phage.
- ▶ 1944 : la nature chimique des gènes est démontrée par transformation de pneumocoques (Avery).
- ▶ 1949 : L'explication moléculaire de l'anémie falciforme (L. Pauling)
- ▶ 1952 : Confirmation de la nature chimique des gènes (Hershey et Chase) par marquage radioactif de l'ADN et des protéines d'un phage.

L'expérience de Beadle et Tatum

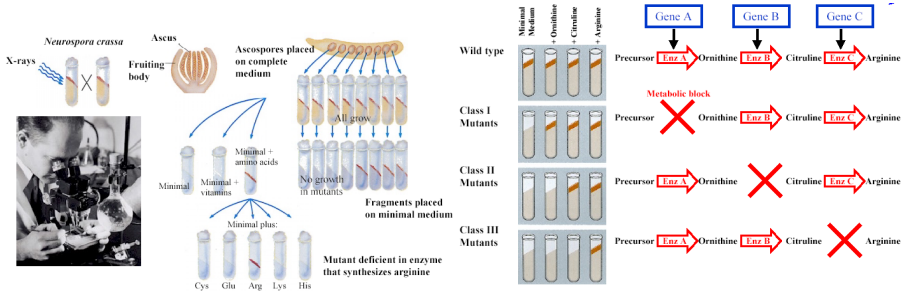


Figure: Les expériences de Beadle et Tatum (1941). En haut : Après exposition aux UV de *Neurospora*, Beadle et Tatum caractérisent des mutants déficients dans la synthèse de l'un des acides aminés, l'arginine. En bas : Parmi ces mutants, ils réussissent à les différencier suivant l'étape qui est déficiente dans la chaîne de synthèse de l'arginine. Chaque étape de la synthèse étant réalisée par une protéine spécifique, cette expérience démontre le lien univoque entre une mutation dans un gène et ses conséquences sur l'activité d'une enzyme donnée.

La nature chimique des gènes

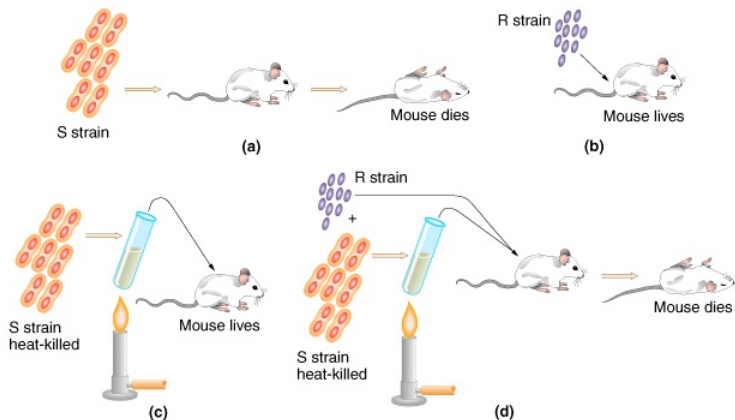


Figure: L'expérience de Griffith (1928) : l'injection de pneumocoques (S) entraîne la mort de la souris, alors que les pneumocoques (R) sont non pathogènes. La co-infection par des (R) vivants et des (S) tués par la chaleur entraîne la mort de la souris mettant en évidence la "transformation" des (R) par les (S).

En 1952, l'expérience de Hershey et Chase

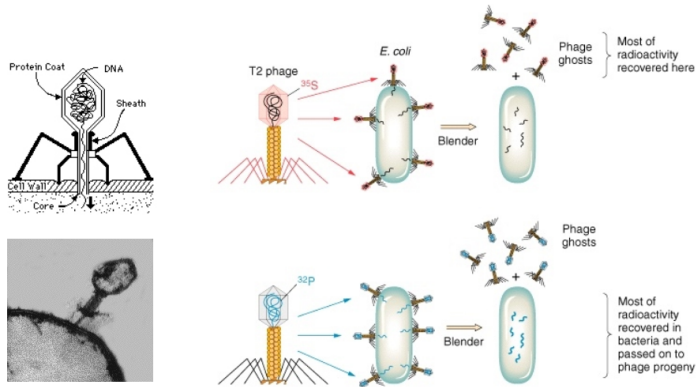
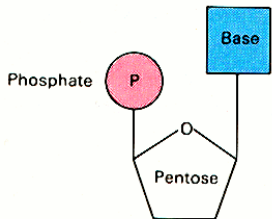


Figure: A gauche : Schéma et Image en microscopie électronique d'un bactériophage . A droite : l'expérience de Hershey et Chase (1952) : le marquage radioactif sélectif des protéines ou de l'ADN et le mécanisme particulier d'infection du bactériophage permet de déterminer la nature chimique du matériel génétique

De la structure de l'ADN à la régulation de l'expression des gènes

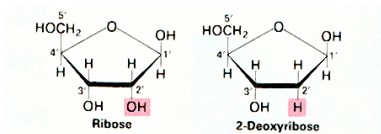
- ▶ 1953 : Découverte de la double hélice (Watson et Crick)
- ▶ 1953 : un modèle du code génétique (Gamow)
- ▶ 1956 : l'ARN est découvert.
- ▶ 1958 : modèle semi-conservatif de la réplication (Meselson et Stahl).
- ▶ 1960 : Découverte de l'ARN messager (Monod).
- ▶ 1961->1966 : le décryptage du code génétique (Nirenberg et Matthaei) grâce à la synthèse de protéines in vitro à partir d'un ARN poly-U.
Affinage du "dogme central" de la biologie moléculaire
- ▶ 1965 : L'opéron lactose (Jacob et Monod)
Etude de l'activité de la Beta-galactosidase
Modèle de la régulation de l'expression des gènes

Le monomère de la molécule d'ADN



Chaque monomère contient :

- ▶ Un phosphate.
- ▶ Un sucre : ribose.
- ▶ Une base : purine ou pyrimidine



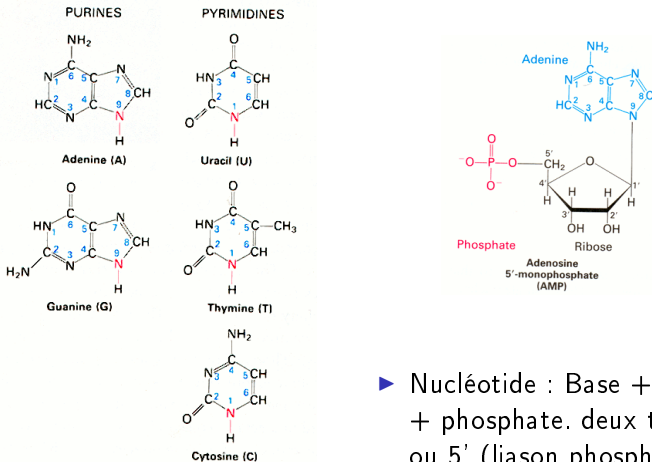
ARN

ADM

Sucre sous forme de cycle

Les atomes du cycle sont numérotés de 1' à 5'

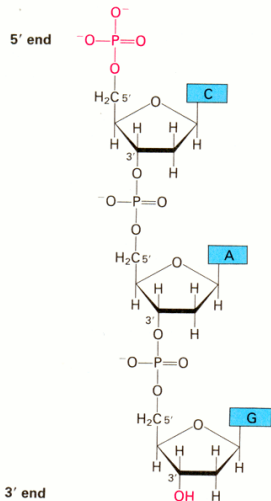
Les bases de l'ADN



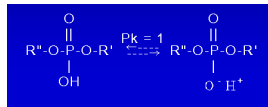
Les bases sont des molécules organiques plates, hydrophobes.

- ▶ Nucléotide : Base + Ribose + phosphate. deux types : 5' ou 3' (liaison phosphate)
- ▶ Nucléoside : Base + Ribose

l'ADN est un polymère

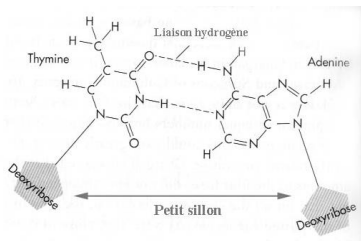


Le polymère est fait d'un "squelette" alterné de groupements phosphate et de riboses. Les bases sont accrochées sur le coté.



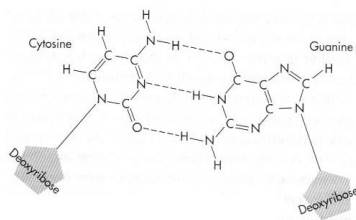
- ▶ pH = 7, la molécule est chargée négativement.
- ▶ Les liaisons se font en 3' et en 5'
- ▶ Le polymère d'ADN simple brin est orienté.

Les bases sont en interaction

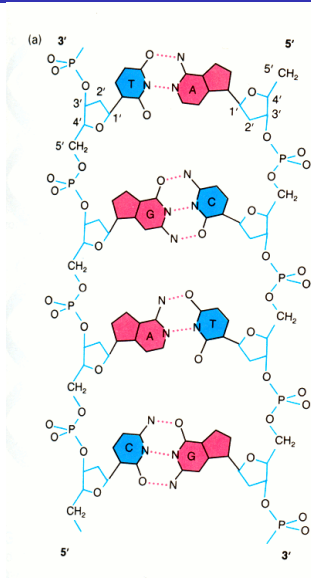


Les interactions A-A, G-A, T-C, etc. existent aussi mais ont des contributions énergétiques plus faibles, elles sont déstabilisées par l'agitation thermique.

- ▶ AT implique 2 liaisons hydrogènes ($1.5 k_B T$ par liaison).
- ▶ GC implique 3 liaisons hydrogènes ($3 k_B T$ par liaison).

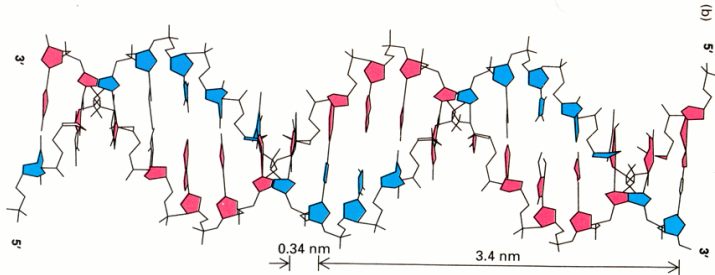


Deux chaînes complémentaires tête bêche



- ▶ Liaisons covalentes au sein des chaînes.
- ▶ Liaisons faibles entre chaînes : liaisons hydrogènes entre les bases complémentaires.
- ▶ Interactions hydrophobes entre bases : énergie d'empilement.

La double hélice droite de l'ADN-B



Diffraction de l'ADN-B par les rayons X

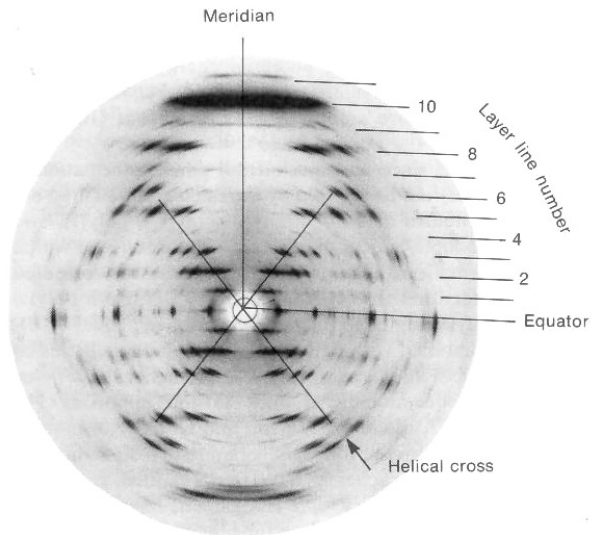
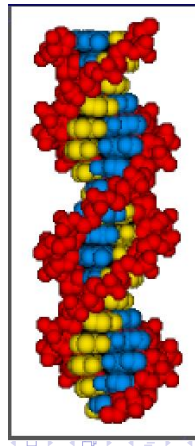
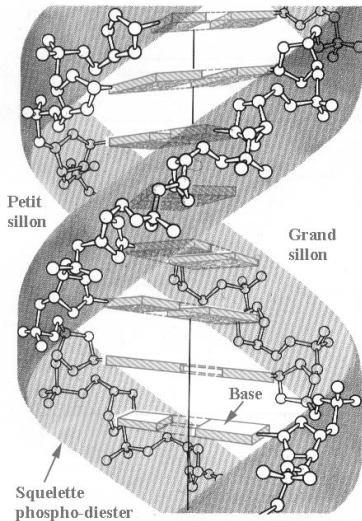


Figure: Diagramme de diffraction obtenu sur une fibre d'ADN avec des rayons X monochromatiques

Le petit et le grand sillon



Les différentes structures de l'ADN

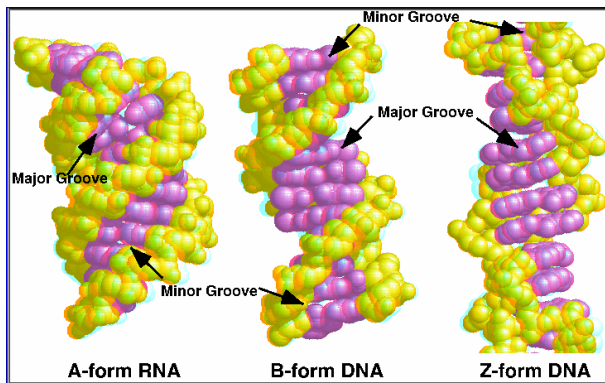


Figure: L'ADN A est obtenu en déshydratant l'ADN, c'est également la structure obtenue pour l'ARN double brin. L'ADN Z correspond à des séquences riches en GC et avec des conditions salines particulières.

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

¹ Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).

² Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).

³ Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).

⁴ Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey¹. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for

is a residue on each chain every 3.4 Å. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 Å. The distance of a phosphorus atom from the fibre axis is 10 Å. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally^{2,4} that the ratio

La réplication semi-conservative, Meselson et Stahl (1958)

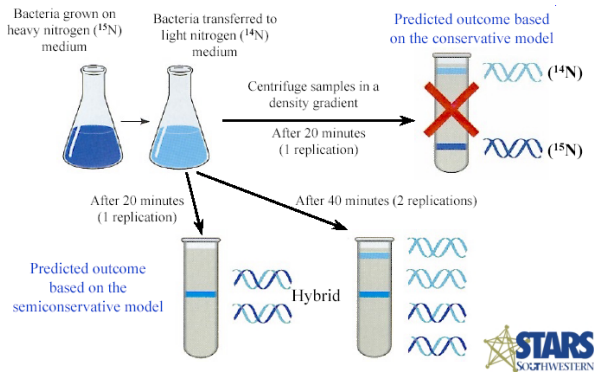


Figure: Expérience de Meselson et Stahl (1958) : séparation d'ADN marqué au ^{32}P suivant sa densité par centrifugation en CsCl et suivi au cours du temps de la proportion d'ADN lourd. Cette expérience démontre le caractère semi-conservatif de la réplication.

Le dogme central de la biologie moléculaire

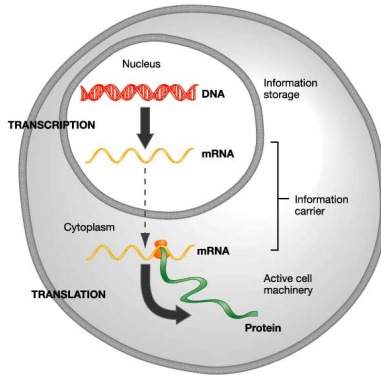
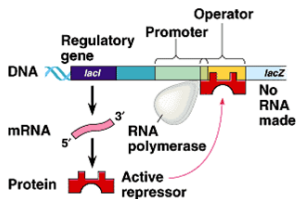
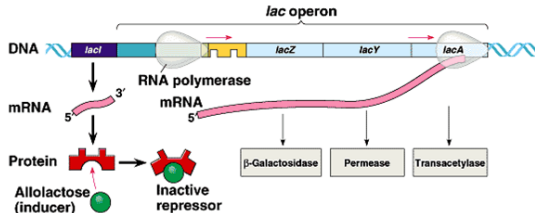


Figure: Le dogme central de la biologie moléculaire, énoncé par Crick en 1960. Il décrit l'expression de l'information génétique depuis son support, l'ADN, jusqu'aux protéines. Cette vision est cependant réductrice, voir texte.

La régulation de l'expression des gènes : l'opéron lactose



(a) Lactose absent, repressor active, operon off



(b) Lactose present, repressor inactive, operon on

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

Figure: L'opéron lactose : le modèle historique de régulation ; La β -galactosidase est une enzyme initie la chaîne du métabolisme du lactose en coupant le lactose en galactose et glucose. En l'absence de lactose dans le milieu, une protéine (répresseur) se fixe sur l'ADN en un site (opérateur) qui bloque la transcription du gène de la β -galactosidase. En présence de lactose dans le milieu, un isomère du lactose, l'allolactose, s'accroche sur le répresseur et diminue ainsi son affinité pour l'ADN. La transcription du gène est alors rendue possible.

Du génie génétique à la transgénèse

- ▶ 1972 : Début du génie génétique : recombinaison d'ADN in vitro (Jackson, Symons, Berg) Construction d'une molécule hybride d'ADN à partir d'une enzyme de restriction (EcoR1) et de ligase
- ▶ 1973 : Electrophorèse de l'ADN sur gel d'agarose
- ▶ 1974 : Clonage de l'ADN utilisant des banques d'ADN génomiques
- ▶ 1978 : 50 enzymes de restriction sont connues : cartographie de l'ADN
- ▶ 1977 : Construction du vecteur pBR322 (Bolivar)
- ▶ 1977 : Séquencage de l'ADN (Sanger , Maxam et Gilbert)
- ▶ 1978 : Expression de protéines humaines dans les bactéries
- ▶ 1980 : Expression de protéines humaines dans les cellules animales
- ▶ 1980 : Première plante transgénique
- ▶ 1981 : Première souris transgénique (Brinster)
- ▶ 1983 : Amplification spécifique d'ADN (PCR) (Mullis)

Electrophorèse de l'ADN

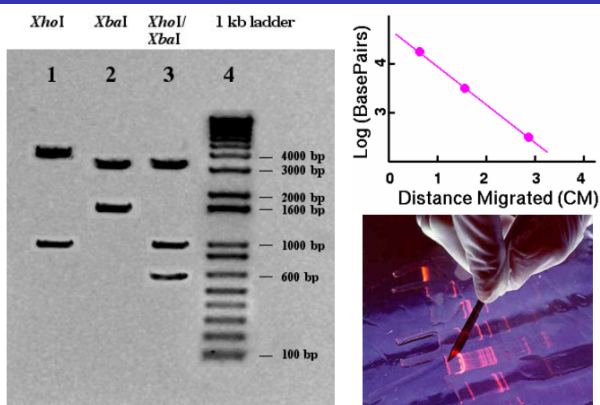
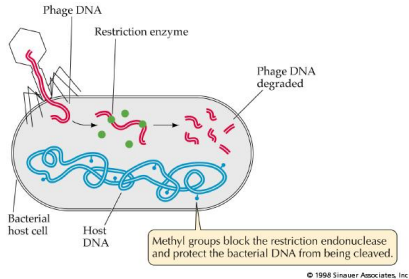


Figure: gauche : Photo du résultat d'électrophorèse d'ADN marqué par fluorescence sur gel d'agarose. Les trois premières colonnes montrent le même fragment coupé par les enzymes *XhoI* et/ou *XbaI*. La dernière colonne est une solution contenant des fragments de tailles connues. Droite haut : La distance de migration varie comme le logarithme de la taille de l'ADN ; droite bas : on peut extraire la bande voulue du gel et en récupérer l'ADN débarrassé des fragments parasites;

Les enzymes de restriction et autres enzymes



Type II:

Coupe au sein du site

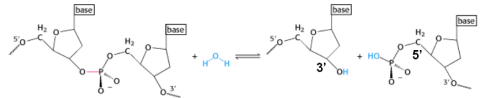
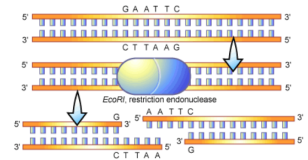


Figure: haut : Les bactéries ont développé le système de restriction/modification pour se protéger des phages. La bactérie modifie son ADN en des sites spécifiques (en général l'ADN est méthylé). Des enzymes de restriction sondent en permanence tout ADN présent dans la bactérie et coupent l'ADN non modifié aux sites de reconnaissance. Bas : exemple d'EcoRI, qui reconnaît la séquence 5'-GAATTC-3'. La coupure forme des extrémités 5'-P et 3'-OH.

Les enzymes de restriction et autres enzymes

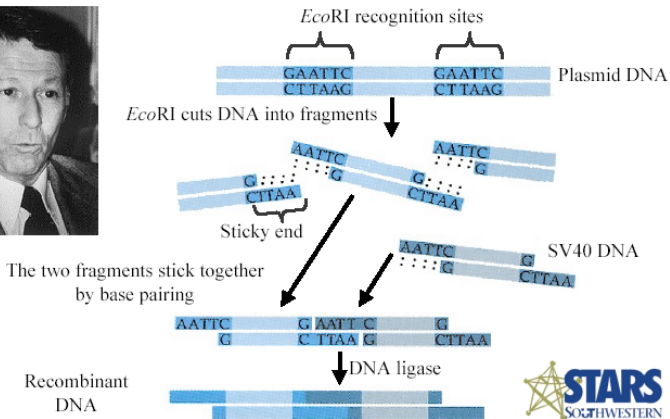


Figure: Expérience de recombinaison in vitro par Paul Berg en 1972 (simplifié) : deux molécules d'ADN d'origine différentes contenant le site de restriction *Eco*RI sont coupées par l'enzyme puis reliées ensemble à l'aide de la ligase

Vecteurs et transformation des cellules

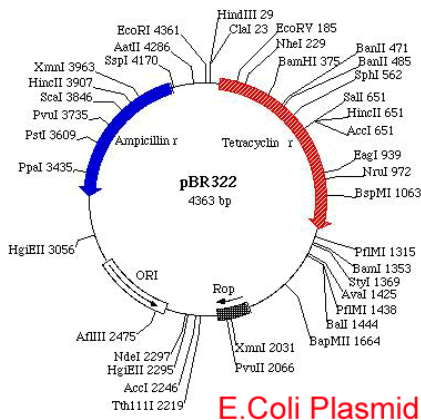
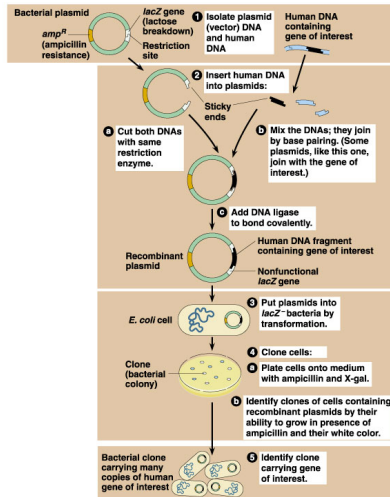


Figure: Carte de restriction du vecteur pBr322 (construit par Bolivar en 1977) : généralement, un plasmide contient une origine de réplication, un ou plusieurs gènes de résistance à des antibiotiques (gènes de sélection, ici ampicilline et tétracycline), et un grand nombre de site de restriction.

Clonage et expression de gènes chez la bactérie



©1999 Addison Wesley Longman, Inc.

Figure: Principe du clonage d'un gène humain dans une bactérie.

Le clonage d'un gène consiste à insérer un gène d'intérêt dans un

L'ADN polymérase

(A. Kornberg en 1956) les polymérases copient le brin complémentaire d'une molécule d'ADN simple brin. Elles travaillent uniquement dans le sens $5' \rightarrow 3'$. Taux d'erreur très faible de (10^{-4} à 10^{-7}) dû à un mécanisme de correction d'erreur, basée sur une activité exonucléase $5' \rightarrow 3'$ ou $3' \rightarrow 5'$.

L'utilisation *in vitro* de ces enzymes est à la base du principe d'amplification de l'ADN par PCR (voir plus loin).

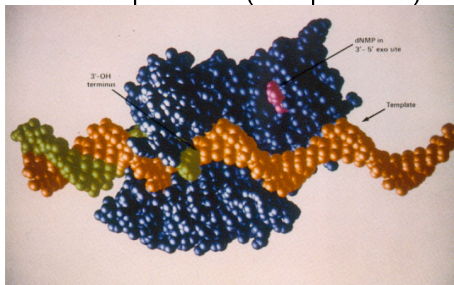


Figure: Structure de l'enzyme Klenow, polymérase de type I chez E.Coli.

Séquence de l'ADN (Méthode de Sanger)

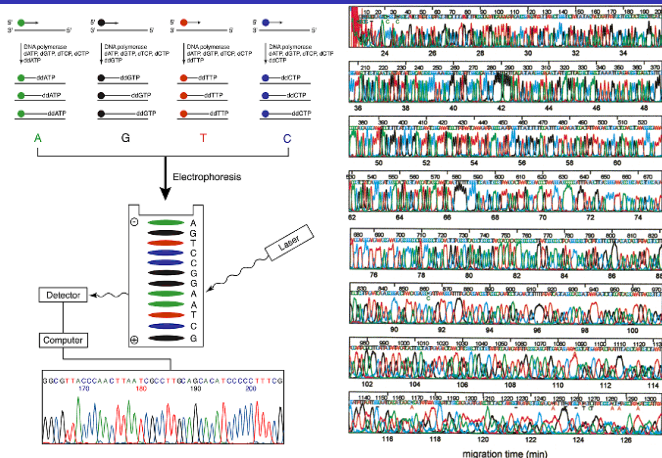


Figure: A gauche : Séquençage de l'ADN par la méthode de Sanger : 4 réactions de polymérisation en parallèle, Chacune contient les 4 nucléotides, qu'une faible quantité d'un des nucléotides modifié de manière à stopper la polymérisation (didéoxynucléotide). On fait ensuite un gel ultrarésolutif pour différencier les fragments différant d'une paire de bases : A droite : Analyse de séquence

Amplification de l'ADN (PCR)

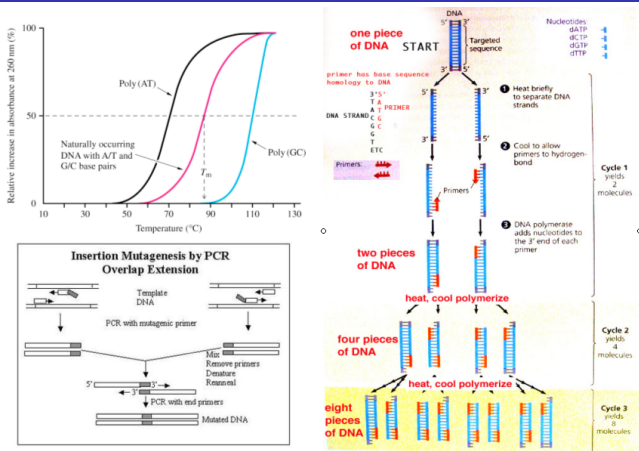
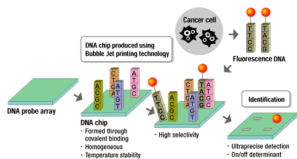




Figure: gauche haut : Dénaturation de l'ADN par la température ; A droite : Cycles d'amplification de l'ADN lors d'une PCR. A chaque cycle, la quantité d'ADN est x 2 ; gauche bas : mutagenèse dirigée par PCR.


Puces à ADN et réseaux génétiques





Les puces à ADN permettent de tester simultanément la transcription d'une grande quantité de gènes d'un organisme. La technique consiste à déposer sur une plaque de verre ou de nylon des ADN spécifiques d'un gène particulier. Ensuite, on vient hybrider sur cette plaque de l'ARN fluorescent préalablement exprimé par une cellule dans une condition particulière. Une plaque peut contenir plusieurs milliers de échantillons, ce qui donne accès à

 M. Morange.
Histoire de la Biologie Moléculaire.
Paris, 1994.

 B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson.
Molecular Biology of the Cell, 3rd.
Garland Publishing Inc., New-York, 1994.

 D. Voet and J.D. Voet.
Biochemistry, 3rd.
New York, 1995.

 F. Jacob.
La Logique du Vivant.
Paris, 1976.

 E. Schrödinger.
Qu'est-ce que la vie ?
Paris, 1993.