

Introduction à la Biologie Moléculaire

Timothée Lionnet et Vincent Croquette

<http://www.phys.ens.fr/~biolps/>

25 mars 2005

Résumé

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].

Table des matières

1	Les origines de la Biologie Moléculaire	2
1.1	La théorie cellulaire	2
1.2	La biochimie	2
1.3	La génétique	2
2	Expériences fondatrices	3
2.1	Quelques dates	3
2.2	L'expérience de Beadle et Tatum	3
2.3	La nature chimique des gènes	4
3	De la structure de l'ADN à la régulation de l'expression des gènes	6
3.1	Quelques dates	6
3.2	La double hélice d'ADN	6
3.3	La réplication semi-conservative	10
3.4	La régulation de l'expression des gènes : l'opéron lactose	11
4	Du génie génétique à la transgénèse	11
4.1	Quelques dates	11
4.2	Electrophorèse de l'ADN	12
4.3	Les enzymes de restriction et autres enzymes	13
4.4	Vecteurs et transformation des cellules	15
4.5	Clonage et expression de gènes chez la bactérie	16
4.6	L'ADN polymérase	17
4.7	Séquençage de l'ADN (Méthode de Sanger)	17
4.8	Amplification de l'ADN (PCR)	18
4.9	Puces à ADN et réseaux génétiques	19

1. Les origines de la Biologie Moléculaire

Les processus biologiques sont utilisées par les êtres humains depuis les temps préhistoriques en cuisine (pain, brassage de la bière, production de fromage...). En médecine aussi, des techniques biologiques ont été utilisées bien avant la compréhension des mécanismes sous-jacents (utilisation par les chinois d'antibiotiques extraits de champignons). L'histoire de la biologie moléculaire commence vraiment au milieu du XIX^{ème} siècle, avec la rationalisation de ces savoirs et observations. Cette période est marquée par l'abandon de la théorie vitaliste et du principe de génération spontanée. La première tente de séparer les êtres vivants du reste de la matière inerte, en leur prêtant une "force vitale", qui expliquerait leurs propriétés physiques (mouvement) et chimiques (fermentation par les micro-organismes). La génération spontanée, elle, postule l'apparition spontanée de la vie, par exemple l'apparition spontanée de micro-organismes sur un milieu fermentescible.

1.1. La théorie cellulaire

- Naissance de la microscopie (Van Leeuwenhoek XVII^{ème} siècle); Observation de "compartiments" dans une tranche de liège (Hooke XVII^{ème} siècle) nommés cellules, mais sans élaborer de théorie quant à leur nature.
- élaboration de la théorie cellulaire (Schleiden & Schwann 1839)
- La théorie de la génération spontanée est égratignée pour la première fois par Francesco Redi, qui prouve en 1668 que l'apparition d'asticots sur un morceau de viande en putréfaction n'a pas lieu si l'on prend soin de recouvrir les bords d'une fine mousseline. 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. Par exemple, Semmelweis en 1846 est assistant à la maternité de Vienne. Il fait passer la mortalité de 12% à 1% en imposant le lavage des mains avant opération. Lorsqu'il demande à son supérieur de se plier à la règle, il est renvoyé et devient fou. Il faudra attendre jusqu'à Koch pour que cette idée soit acceptée (observation de germes d'anthrax en 1870).
- Parallèlement, la compréhension du lien entre micro-organismes et fermentation s'affine: Pasteur, à la demande des industriels lillois, débute des recherches sur les fermentations ("alcoolique", "lactique" et "acétique") et découvre qu'à chaque fois, un micro-organisme est à l'oeuvre. Il met au point la pasteurisation (1863) pour éliminer les ferments parasites.

1.2. La biochimie

- Le principe de la catalyse est découvert (1812).
- La synthèse de l'urée en 1828 par Wöhler à partir de cyanate d'argent contredit la théorie vitaliste puisqu'il est possible de synthétiser un composé organique ("du vivant") à partir d'un composé minéral. Toutefois, les partisans du vitalisme répliquent que l'urée n'est qu'un produit de dégradation et ne sont pas convaincus.
- la première enzyme est extraite du malt: il s'agit de la diastase qui catalyse la transformation de l'amidon en glucose (Payen & Persoz 1830).
- fermentation des sucres à partir d'un extrait acellulaire de levure obtenu par broyage (Büchner 1897). La théorie vitaliste est contredite: les composants chimiques d'un être vivant suffisent seuls à remplir les fonctions de celui-ci.
- La notion de spécificité est découverte; reconnaissance enzyme substrat; étude des anticorps (Ehrlich 1897);
- Echelle de pH
- Solutions "tampon"
- 1902-1909 Principe de fonctionnement des enzymes (Henri, Michaelis & Menten)
- Premier moitié du XX^{ème} siècle: décryptage des grandes voies métaboliques: glycolyse, cycle de l'urée, cycle de Krebs
- Découverte des macromolécules (1922: Hermann Staudinger élabore le concept de polymère)
- Technique de séparation et d'analyse: Chromatographie, précipitation par des sels ultracentrifugation (1924) électrophorèse (1930-1945)

1.3. La génétique

- Mendel découvre en 1866 les "lois de l'hérédité" en étudiant la transmission de caractères dans des générations successives de pois. 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.
- Johanssen 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).

2. Expériences fondatrices

2.1. 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.

2.2. L'expérience de Beadle et Tatum

En 1902, Garrod découvre que la transmission d'une maladie humaine, l'alcaptonurie (une anomalie d'excrétion, affectant le métabolisme de la tyrosine et de la phénylalanine) se fait en accord avec les lois mendéliennes, et que cette maladie est liée à l'absence d'une enzyme dans le sérum des patients. Il est le premier à entrevoir le lien fondamental entre biochimie et génétique: le phénotype est de nature chimique, et ce sont les protéines qui en sont le support. Cette intuition fondamentale reste malheureusement non démontrée, car il s'agit d'une maladie rare et la collecte des statistiques nécessaires serait beaucoup trop longue. Il faut donc attendre que l'étude soit menée sur un système plus simple, et surtout au cycle de reproduction plus court, pour pouvoir apporter une preuve de la relation gène-enzyme. L'expérience de Beadle et Tatum (1941) démontre donc la relation univoque (vision simpliste cependant, qui a été affinée depuis) entre un gène et une protéine responsable d'une activité catalytique. Ils se penchent sur une moisissure, *Neurospora crassa*, facilement cultivable sur un milieu artificiel. On sait depuis Muller (1927) induire des mutations par exposition aux rayons X. Par irradiation, Beadle et Tatum obtiennent donc des mutants incapables de se développer sur ce milieu minimal, mais qui poussent sur un milieu complété avec tel ou tel métabolite. En particulier, les études de la synthèse du tryptophane, ou du cycle de l'ornithine, montrent que chacun de ces mutants est en réalité déficient en une des enzymes nécessaires à l'une des étapes des ces chaînes métaboliques. L'analyse génétique de ces mutants permet de montrer que chacune de ces déficiences segrège de façon mendélienne, et correspond donc à une mutation dans un unique gène. L'ensemble de ces observations aboutit donc à la conclusion que les gènes contrôlent la synthèse des enzymes, et que chaque protéine est codée par un gène différent.

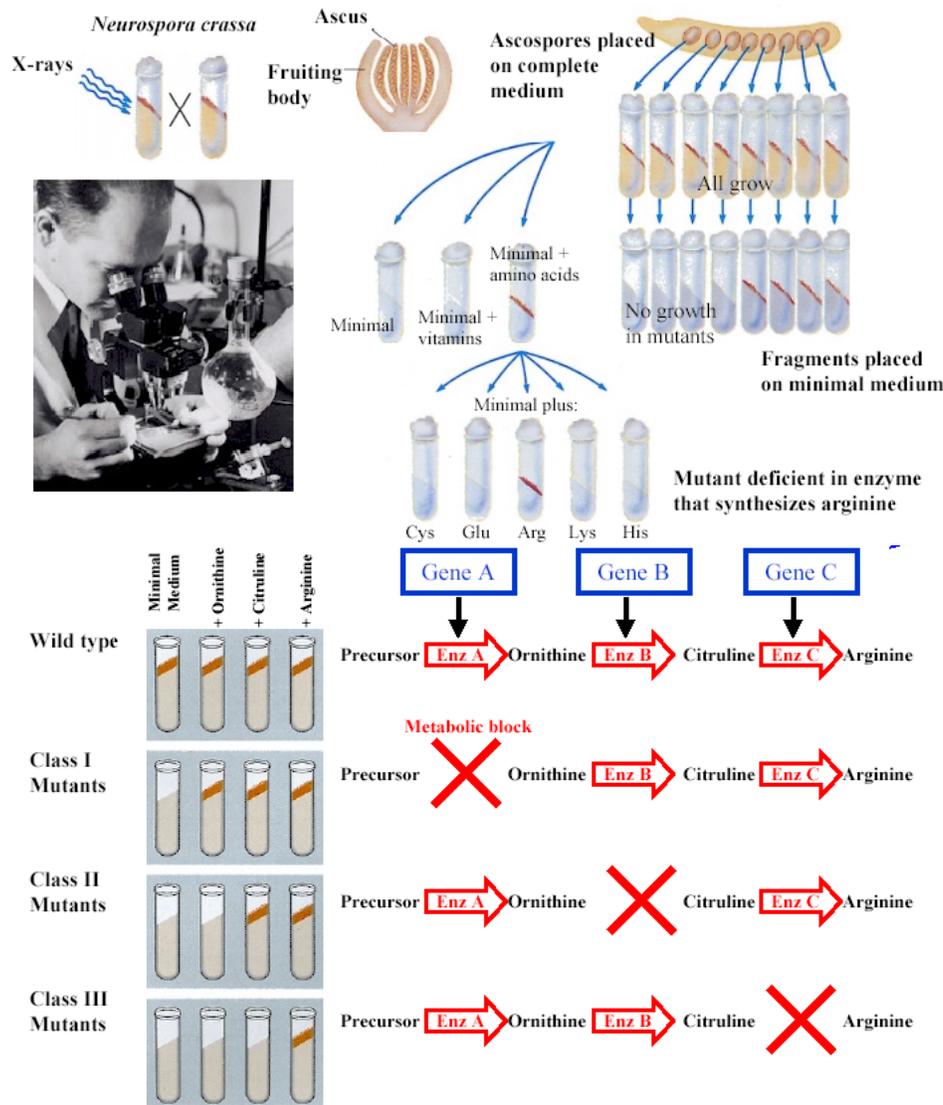


FIG. 1 – Les expériences de Beadle et Tatum (1941). En haut: Après exposition aux UV de *Neurospora*, Beadle et Tatum caractérisent des mutants déficients dans la synthèse de l'un des acides aminés, l'arginine. En bas: Parmi ces mutants, il réussissent à les différencier suivant l'étape qui est déficiente 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.

2.3. La nature chimique des gènes

En 1928, Fred Griffith découvre le phénomène de "transformation": une coinjection chez une souris de pneumocoques non pathogènes (R) et de pneumocoques virulents (S) mais tués préalablement entraîne la mort de l'animal. Les pneumocoques non pathogènes sont ainsi transformés par un facteur issu des pneumocoques virulents. En 1944, la purification du facteur transformant par Avery aboutit au résultat suivant: ce n'est pas une protéine, mais un acide desoxyribonucléique (ADN), mettant ainsi en évidence le rôle de l'ADN comme support de l'information génétique. Cependant, l'ADN est alors considéré comme un homopolymère dont l'élément de base serait ATGC (Takahashi 1932), et semble donc beaucoup trop rudimentaire pour être le support de l'information génétique.

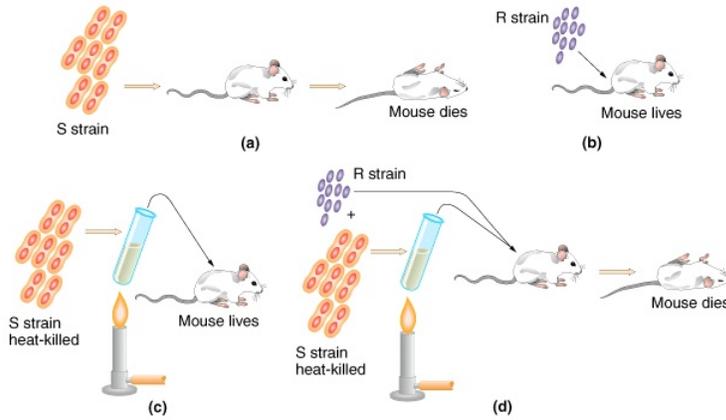


FIG. 2 – L'expérience de Griffith (1928): l'injection de pneumocoques (*S*) entraîne la mort de la souris, alors que les pneumocoques (*R*) sont non pathogènes. La co-infection par des (*R*) vivants et des (*S*) tués par la chaleur entraîne la mort de la souris mettant en évidence la "transformation" des (*R*) par les (*S*).

En 1952, l'expérience de Hershey et Chase confirme ce résultat. C'est une des premières expériences de biologie moléculaire où des isotopes radioactifs permettent de tracer des molécules. Ils utilisent un marquage isotopique différentiel de chacun des constituants du phage: du phosphore radioactif, incorporé à l'ADN, et du soufre radioactif, incorporé aux protéines de la capsid. Ces phages sont utilisés pour infecter des bactéries. Immédiatement après l'infection, il est possible de séparer par agitation mécanique les bactéries des phages qui les ont infectés, et de séparer par centrifugation les particules phagiques des bactéries infectées. Lorsqu'elles sont remises en culture, ces bactéries produisent des phages. Or l'analyse de la répartition de la radioactivité montre que ces bactéries ne contiennent que du phosphore radioactif: seul l'ADN a donc pénétré dans la cellule, et cette fraction est responsable à elle seule de la reproduction du phage. C'est donc l'ADN qui détient l'information génétique.

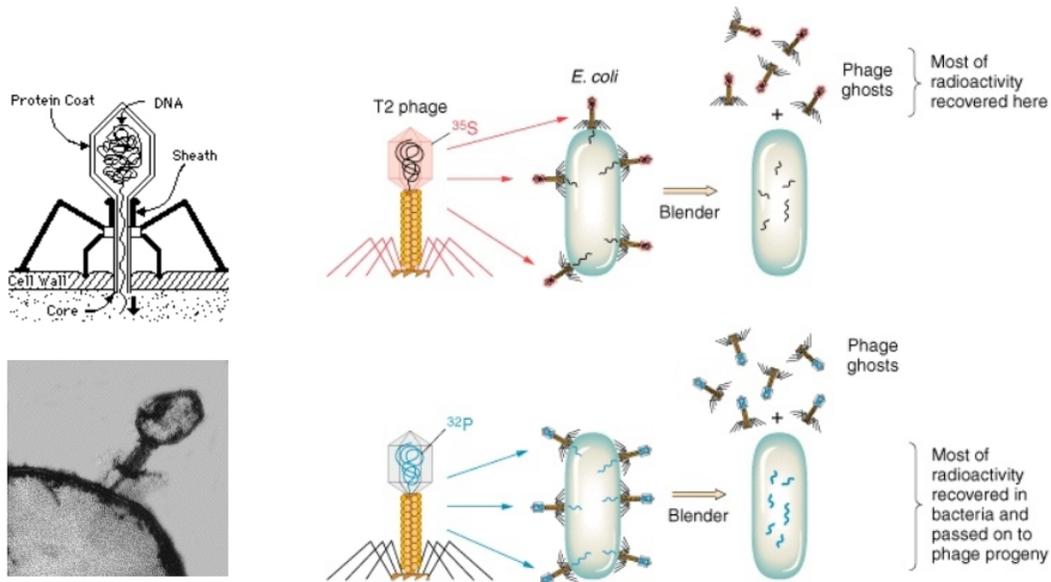


FIG. 3 – A gauche: Schéma et image en microscopie électronique d'un bactériophage (c'est une enveloppe constituée de protéines -la capsid- qui contient de l'ADN; il infecte les bactéries en injectant son ADN, laissant sa capsid à l'extérieur de la bactérie). A droite: l'expérience de Hershey et Chase (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

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

3.1. Quelques dates

- 1953 : Découverte de la double hélice (Watson et Crick)
Analyse de clichés de diffraction de rayon X par des cristaux d'ADN
- 1953: un modèle du code génétique (Gamow)
- 1956: l'ARN est découvert.
- 1958 : La démonstration du modèle semi-conservatif de la réplication de l'ADN (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

3.2. La double hélice d'ADN

La découverte de la structure de l'ADN bouleverse l'étude des phénomènes biologiques en introduisant la dimension moléculaire. Proposée par Watson et Crick en 1953, elle est obtenue non seulement à partir de l'interprétation de clichés de diffraction des rayons X réalisées par Franklin, mais aussi des travaux d'Erwin Chargaff (qui avaient montré que pour toute molécule d'ADN, le nombre de molécules d'adénine est égal au nombre de molécules de thymine, et que celui de cytosine est égal à celui de guanine) et enfin d'analyses en microscopie électronique, qui avaient montré que le diamètre de la molécule d'ADN est de 20 Å, ce qui suggérait que cette molécule comportait deux chaînes de désoxyribose-phosphate. Les deux points fondamentaux de la structure sont les suivants:

- les bases ou nucléotides (A,T,C,G) s'organisent en paires A-T et G-C
- Cet appariement permet un enroulement quasi-parfait en hélice droite des deux chaînes sucre-phosphate qui portent ces nucléotides; L'hélice possède un pas d'environ 10 paires de bases, qui sont espacées de 3.4 Å, donc le pas vaut environ 3.4 nm.

La structure est stabilisée par l'interaction (liaisons hydrogène) entre les bases et l'empilement successif des paires de nucléotides le long de la double-hélice ("stacking" du à des interactions hydrophobes).

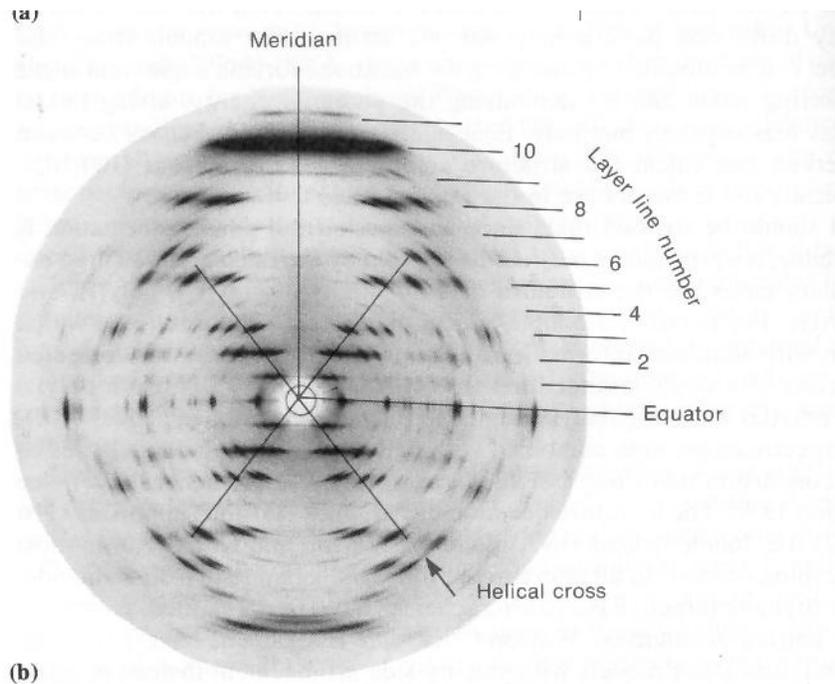


FIG. 4 – A droite: Image par diffraction des rayons X d'un cristal d'ADN.

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).

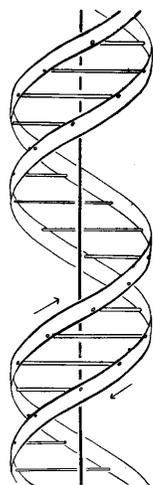
MOLECULAR STRUCTURE OF NUCLEIC ACIDS

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 this reason we shall not comment on it.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate-sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate di-ester groups joining β -D-deoxy-ribofuranose residues with 3',5' linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's² model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There

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^{3,4} that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data^{5,6} on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

FIG. 5 - Article original de Watson et Crick (30 Avril 1953);

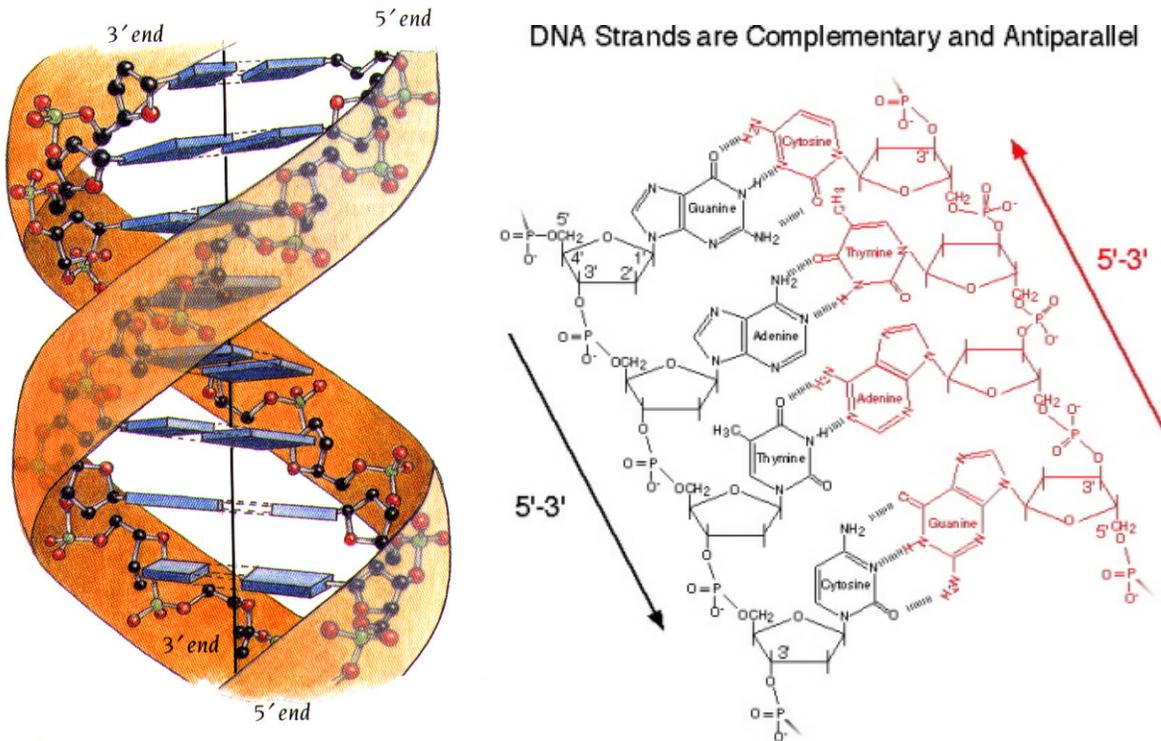


FIG. 6 – A gauche: structure primaire de l'ADN: les paires bases A-T et G-C sont stabilisées par respectivement 2 et 3 liaisons hydrogène. Chaque hélice est formée par un squelette sucre-phosphate orienté. Les deux squelettes sont orientés en sens inverse; A droite: la structure secondaire de l'ADN en double-hélice droite.

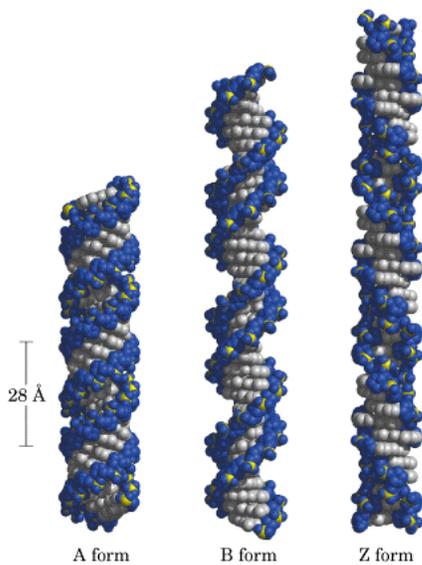


TABLE 28-1. STRUCTURAL FEATURES OF IDEAL A-, B-, AND Z-DNA

	A	B	Z
Helical sense	Right handed	Right handed	Left handed
Diameter	~26 Å	~20 Å	~18 Å
Base pairs per helical turn	11	10	12 (6 dimers)
Helical twist per base pair	33°	36°	60° (per dimer)
Helix pitch (rise per turn)	28 Å	34 Å	45 Å
Helix rise per base pair	2.6 Å	3.4 Å	3.7 Å
Base tilt normal to the helix axis	20°	6°	7°
Major groove	Narrow and deep	Wide and Deep	Flat
Minor groove	Wide and shallow	Narrow and deep	Narrow and deep
Sugar pucker	C3'-endo	C2'-endo	C2'-endo for pyrimidines; C3'-endo for purines
Glycosidic bond	Anti	Anti	Anti for pyrimidines; syn for purines

FIG. 7 – A gauche: il existe d'autres structures d'ADN que la forme (B) la plus répandue. La forme (A), de diamètre plus grand et de pas plus petit, est stable en milieu anhydre; La forme (Z), hélice gauche stable à haute concentration en sel; L'intérêt biologique de ces structures demeure incertain

3.3. La réplication semi-conservative

La découverte de la structure de l'ADN pose le problème de la réplication du code génétique lors de la division cellulaire. La compréhension du mécanisme de réplication fait un pas décisif avec l'expérience de Meselson et Stahl (1958).

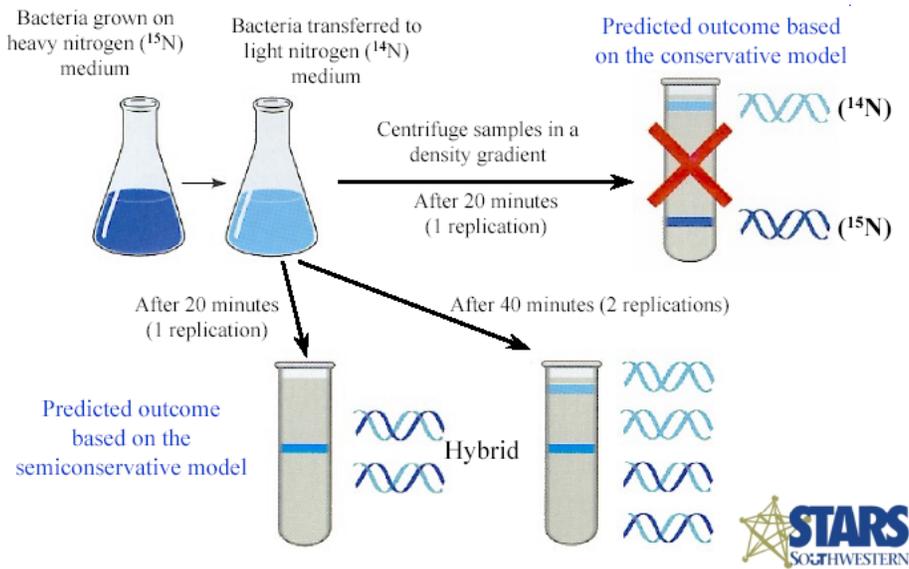


FIG. 8 – *Expérience de Meselson et Stahl (1958) : séparation d'ADN marqué au ³²P suivant sa densité par centrifugation en CsCl et suivi au cours du temps de la proportion d'ADN lourd. Cette expérience démontre le caractère semi-conservatif de la réplication.*

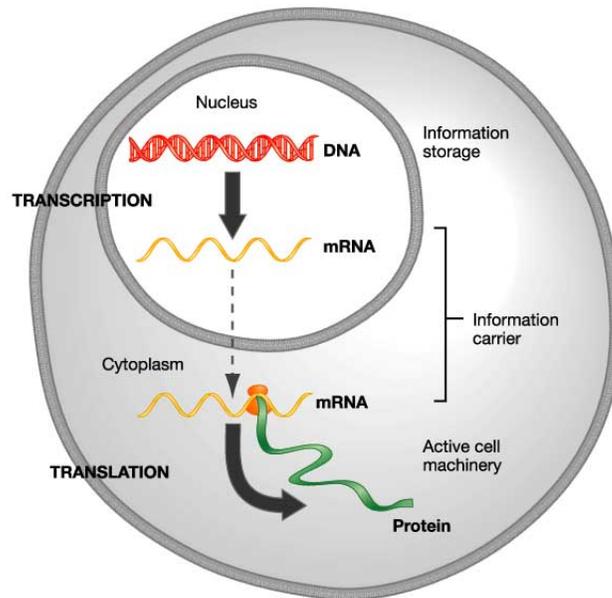


FIG. 9 – *Le dogme central de la biologie moléculaire, énoncé par Crick en 1960. Il décrit l'expression de l'information génétique depuis son support, l'ADN, jusqu'aux protéines. Cette vision est cependant réductrice, voir texte.*

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

Dans les années soixante, Jacob et Monod découvrent le premier système de régulation de l'expression des gènes, qu'ils nomment l'"opéron" lactose. Il s'agit du système qui gère la régulation de la production d'une enzyme (qui intervient au début de la chaîne de métabolisme du lactose) en fonction de la concentration de lactose dans le milieu.

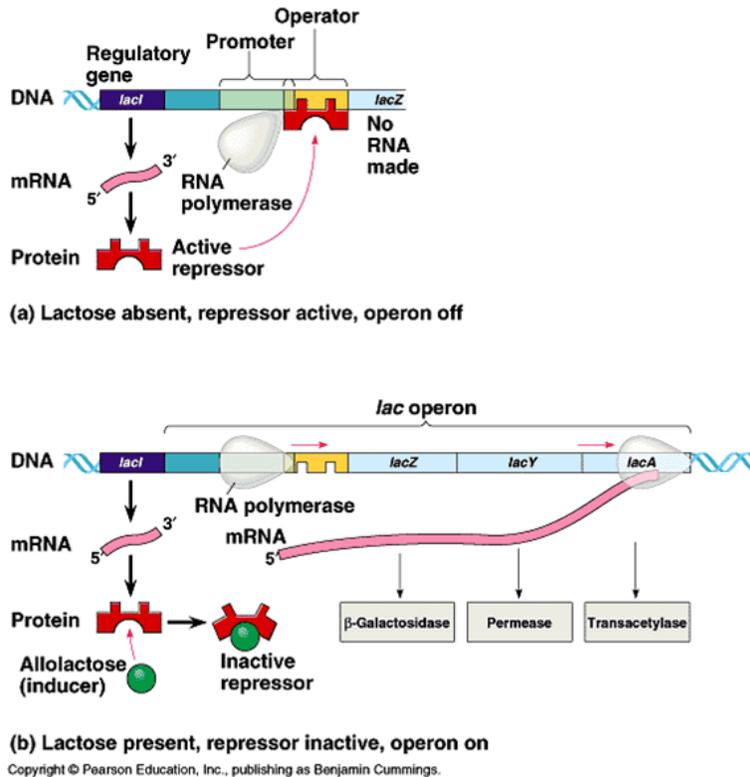


FIG. 10 – L'opéron lactose : le modèle historique de régulation; La β -galactosidase est une enzyme initie la chaîne du métabolisme du lactose en coupant le lactose en galactose et glucose. En l'absence de lactose dans le milieu, une protéine (répresseur) se fixe sur l'ADN en un site (opérateur) qui 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.

La découverte de la régulation permet d'affiner le dogme central de la biologie moléculaire. D'autre part, on sait aujourd'hui qu'à un gène ne correspond pas forcément une protéine unique. En effet, l'expression peut subir des modifications:

- post-transcriptionnelles: l'ARN messager transcrit à partir d'un gène, peut être recombiné (certaines parties sont coupées et éliminées, les introns, les autres sont "recollées" entre elles, les exons). C'est ce qu'on appelle l'épissage alternatif, qui peut être modulé en fonction du cycle cellulaire ou de stimuli extérieurs.
- post-traductionnelles: le repliement 3D d'une protéine peut être modifié, par exemple sous l'action de protéines dites chaperonnes. D'autre part, de nombreuses protéines subissent des modifications chimiques (formation de ponts disulfures, ajouts de groupements sucres pour former des glycoprotéines) après leur synthèse.

Ces variations d'expression sont à l'origine de la complexité de l'expression de l'information génétique. Certes, toute l'information génétique est contenue dans l'ADN, mais deux cellules au même contenu ADN peuvent être extrêmement différentes, en fonction du contenu de leur cytoplasme (cf différenciation des cellules dans un organisme).

4. Du génie génétique à la transgénèse

4.1. Quelques dates

- 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)

4.2. Electrophorèse de l'ADN

Le principe de l'électrophorèse, qui consiste à faire migrer dans un certain milieu des molécules chargées en présence d'un champ électrique, date des années trente. Cependant, elle est devenue une technique très répandue d'analyse et de séparation d'ADN (on l'utilisait initialement pour la migration de protéines) avec l'utilisation de gel d'agarose (1973). L'ADN est déposé dans un gel d'agarose en présence d'un marqueur qui devient fluorescent lorsqu'il s'intercale dans l'ADN. L'ADN est attiré vers l'électrode + (phosphates chargés négativement), mais les mailles du gel d'agarose permettent de séparer les molécules en fonction de leur taille. Cette méthode permet de vérifier la taille de fragments d'ADN, et de les séparer ensuite. On peut aussi l'utiliser pour détecter la présence d'une séquence spécifique dans des fragments donnés. C'est la technique du Southern blot: Après l'électrophorèse d'ADN double brin, le gel est transféré sur une membrane et l'ADN double brin dénaturé en simple brin. Cette étape permet de fixer l'ADN sur un support solide. La réplique du gel sur la membrane est ensuite mélangée à une sonde, c'est-à-dire un simple brin d'ADN portant un marqueur (radioactif ou fluorescent). Après rinçage, seules les bandes du gel ayant intégré la séquence voulue retiennent la sonde, et on peut les révéler (par autoradiogramme ou simple photo dans le cas d'un marqueur fluorescent). Cette technique est utilisable pour l'ADN (Southern), l'ARN (Northern), ou les protéines (Western).

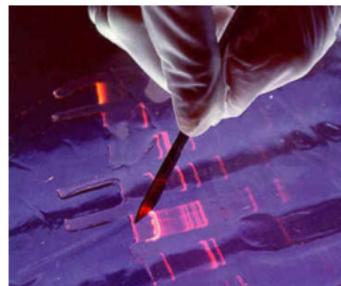
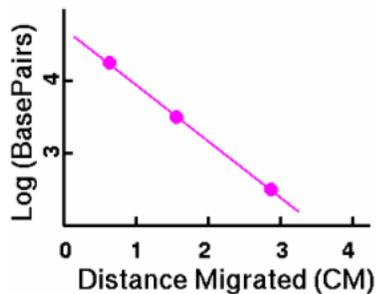
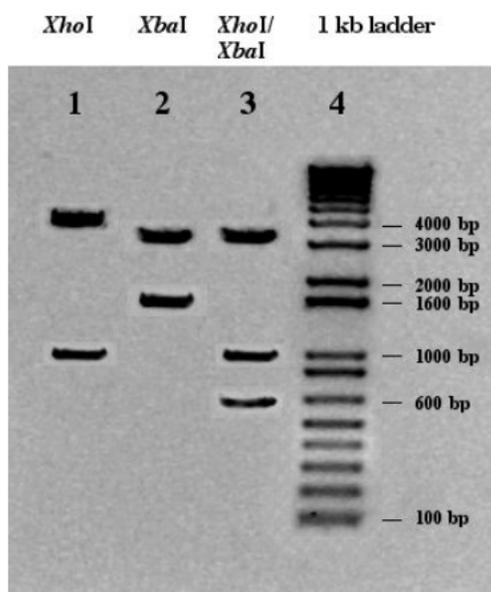


FIG. 11 – gauche : Photo du résultat d'électrophorèse d'ADN marqué par fluorescence sur gel d'agarose. Les trois premières colonnes montrent le même fragment coupé par les enzymes *XhoI* et/ou *XbaI*. La dernière colonne est une solution contenant des fragments de tailles connues. Ainsi, on peut mesurer la taille et séparer différentes molécules d'ADN; 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;

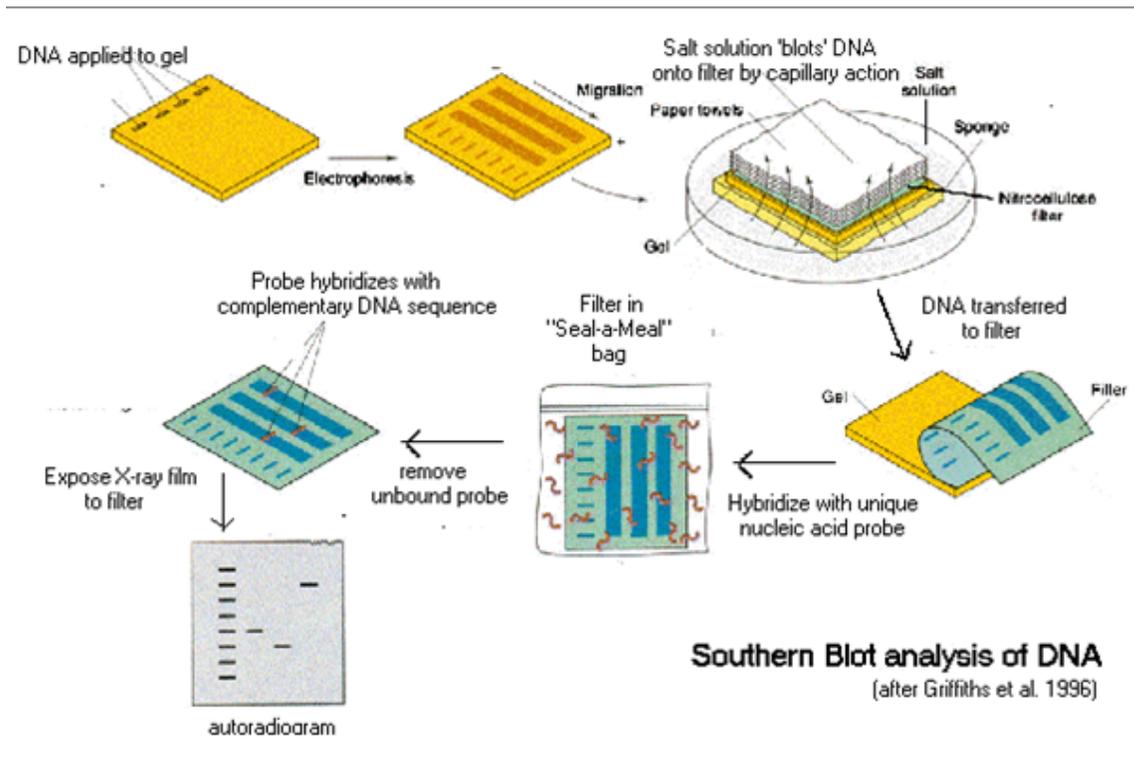


FIG. 12 – gauche : Principe de la technique d'hybridation sur gel ("blot"). L'ADN issu d'un gel est immobilisé sur une membrane et mis en présence d'une sonde pour tester la présence d'une séquence voulue dans les différents fragments;

4.3. Les enzymes de restriction et autres enzymes

La découverte des enzymes de restriction signe le début de l'ère du "génie génétique". Ces enzymes, qui constituent *in vivo* un mécanisme de défense contre des molécules d'ADN étrangères, sont capables de couper la molécule d'ADN en des sites bien spécifiques. La ligase, elle, catalyse la réaction inverse, c'est-à-dire la formation de liaisons phosphodiester pour recoller deux brins d'ADN. Les extrémités sont importantes, on distingue les fragments d'ADN à bouts francs (i.e. sans qu'un des deux brins ne dépasse) et les fragments à extrémités collantes (synonyme: cohésives) dont un des brins dépasse. La ligase ne peut que recoller des fragments dont les extrémités cohésives sont complémentaires (ou deux fragments à bouts francs, mais avec un moins bon rendement). En 1972, L'équipe de Berg utilise l'enzyme de restriction EcoRI pour obtenir *in vitro* une molécule d'ADN hybride qui contient à la fois l'ADN d'un virus (SV40) et une forme altérée d'un bactériophage (λ). C'est la première recombinaison génétique réalisée *in vitro*.

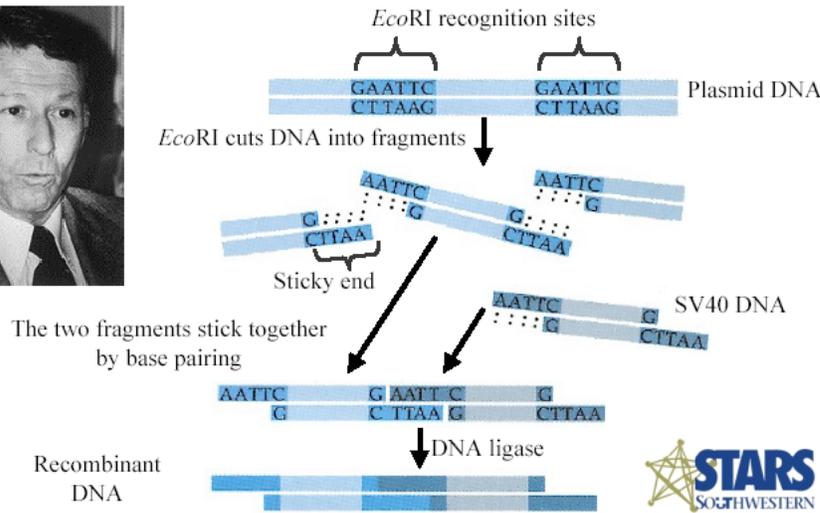


FIG. 14 – *Expérience de recombinaison in vitro par Paul Berg en 1972 (simplifié): deux molécules d'ADN d'origine différentes contenant le site de restriction EcoRI sont coupées par l'enzyme puis reliées ensemble à l'aide de la ligase*

4.4. Vecteurs et transformation des cellules

La découverte de nombreuses enzymes de restriction et la possibilité de recombiner l'ADN *in vitro* conduisent à l'utilisation de petites molécules circulaires d'ADN, les plasmides. Les plasmides sont présents naturellement dans le cytoplasme des bactéries et sont répliqués en parallèle du chromosome bactérien. La bactérie les tolère car ils portent en général un gène qui procure un avantage sélectif. On a depuis élaboré un grand nombre de plasmides artificiels, qui contiennent beaucoup de sites de restriction, et en général un ou deux gènes permettant de les sélectionner.

Après insertion d'une séquence d'ADN particulière dans un plasmide (qu'on appelle aussi vecteur), celui-ci peut être introduit dans une bactérie ("transformation"), soit pour obtenir un grand nombre d'exemplaire de cette séquence, soit pour modifier le patrimoine génétique de la bactérie. L'expression du plasmide peut être stable ou transitoire, et avec ou sans site de réplication. Une des limitations de la transformation est que la bactérie peut tolérer ou non la séquence introduite, et donc peut éventuellement éliminer le plasmide, ou le modifier.

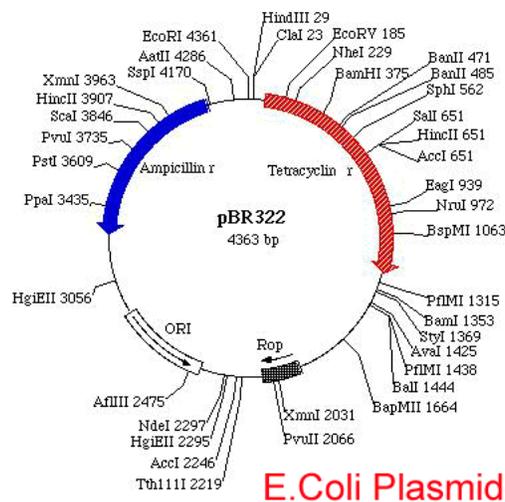
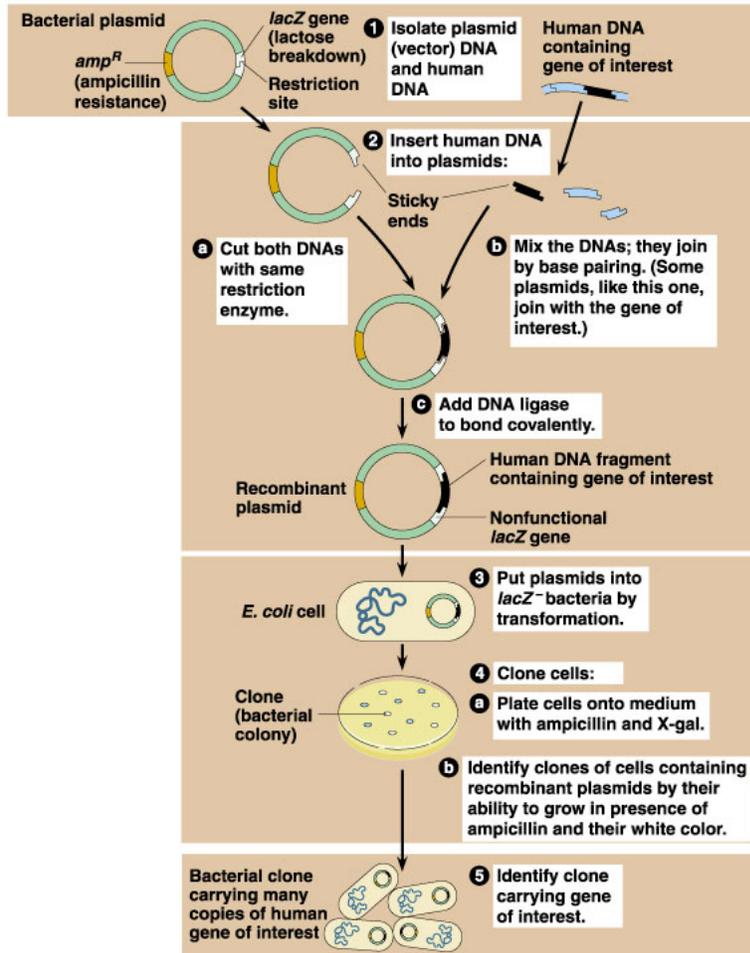


FIG. 15 – *Carte de restriction du vecteur pBr322 (construit par Bolivar en 1977): généralement, un plasmide contient une origine de réplication, un ou plusieurs gènes de résistance à des antibiotiques (gènes de sélection, ici ampiciline et tétracycline), et un grand nombre de site de restriction.*

4.5. Clonage et expression de gènes chez la bactérie

Le clonage d'un gène consiste à insérer un gène d'intérêt dans un vecteur et transformer des bactéries avec ce vecteur. Une fois inséré dans la bactérie, ce gène pourra, selon les besoins, être amplifié en vue d'un séquençage (voir plus loin), ou exprimé en vue de la production de la protéine associée à ce gène.

Historiquement, les premières expériences de production de protéines humaines dans les bactéries furent des protéines d'intérêt thérapeutique comme l'insuline (1979). La production de protéines ou l'élaboration d'organismes transgéniques ne sont pas des techniques limitées aux bactéries. D'autres méthodes de transformation peuvent aussi être employées, par exemple l'utilisation de vecteurs différents des plasmides (virus modifiés, phages, cosmides, phagemides, etc).



©1999 Addison Wesley Longman, Inc.

FIG. 16 – Principe du clonage d'un gène humain dans une bactérie.

4.6. L'ADN polymérase

Découverte par Arthur Kornberg en 1956, les polymérase assurent la copie du brin complémentaire d'une molécule d'ADN simple brin lors de la réplication. Elles travaillent uniquement dans le sens 5'→3'. Certaines présentent un taux d'erreur remarquablement faible (10^{-4} et jusqu'à 10^{-7}) du à un mécanisme de correction d'erreur, basée sur une activité exonucléase 5'→3' ou 3'→5'.

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

4.7. Séquençage de l'ADN (Méthode de Sanger)

Le séquençage d'une molécule d'ADN consiste à déterminer l'enchaînement des bases qui le compose. En 1977 sont mises au point deux techniques concurrentes: celle de Gilbert, d'un part, qui est une méthode chimique de coupure des bases, et celle de Sanger, d'autre part, qui utilise un procédé enzymatique requérant l'ADN polymérase. Les deux techniques requièrent l'utilisation d'électrophorèse ultra-résolutive, permettant la séparation de molécules d'ADN ne différant que d'un seul nucléotide.

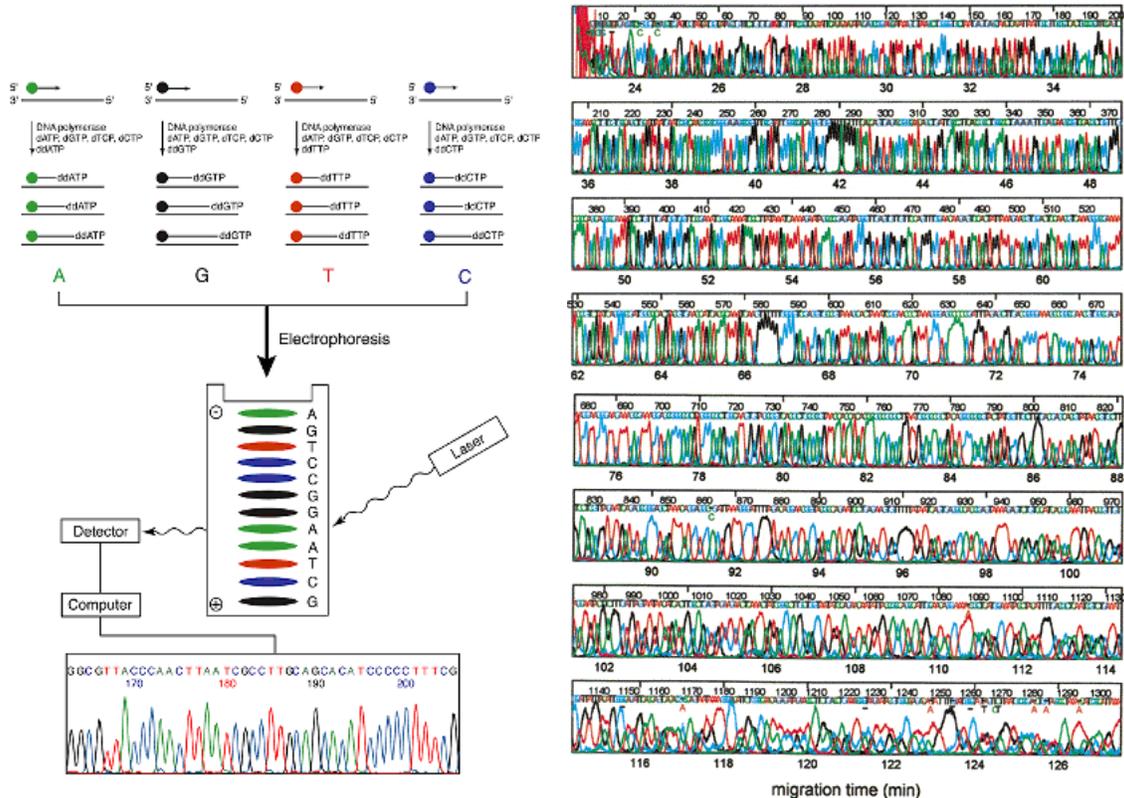


FIG. 17 – A gauche: principe du séquençage de l'ADN par la méthode de Sanger modernisée. 4 réactions de polymérisation sont menées en parallèle. Chacune contient les 4 nucléotides, un marqueur coloré spécifique, ainsi qu'une faible quantité d'un des nucléotides modifié de manière à stopper la polymérisation (didéoxynucléotide). On fait ensuite un gel ultrarésolutif pour différencier les fragments différant d'une paire de bases; A droite : Analyse de séquence

4.8. Amplification de l'ADN (PCR)

La technique d'amplification de l'ADN par Réaction de Polymérisation en Chaîne (PCR) vient compléter l'arsenal du génie génétique en 1983. La PCR permet d'amplifier de manière très spécifique et très fidèle une séquence d'ADN d'intérêt.

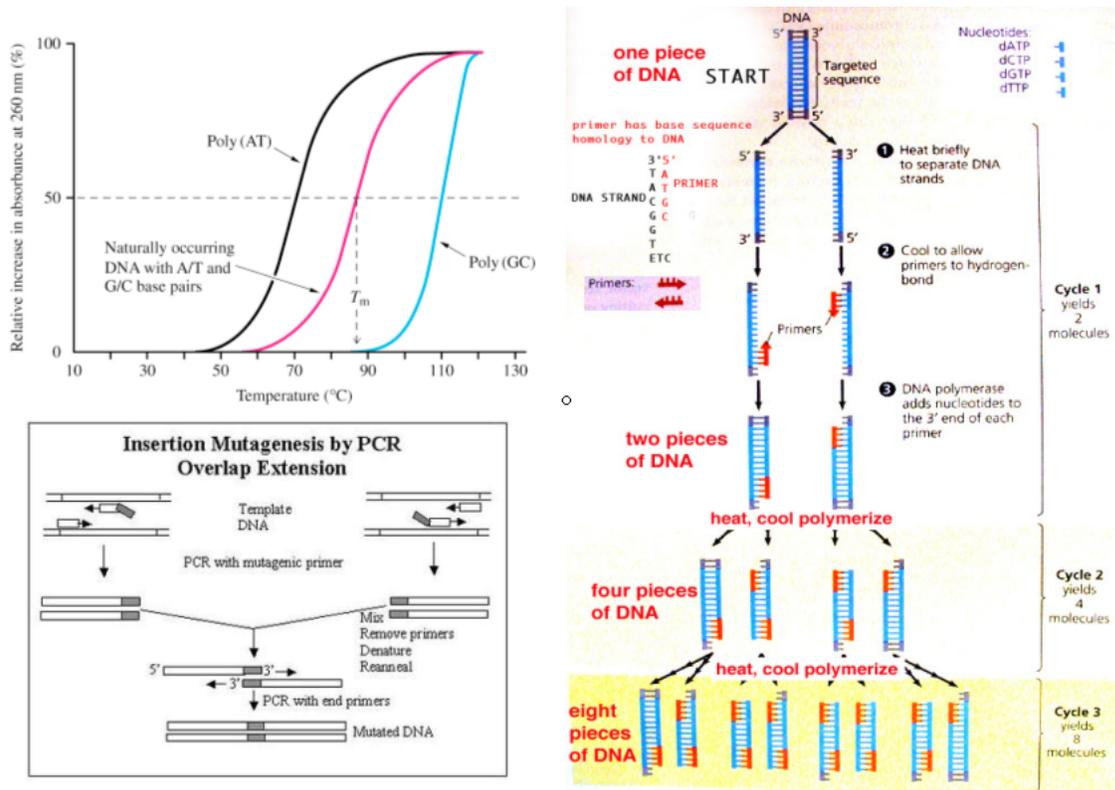


FIG. 18 – gauche haut: Dénaturation (melting) de l'ADN en fonction de la température et la composition de l'ADN; A droite: Cycles d'amplification de l'ADN lors d'une PCR. La PCR est un processus exponentiel: à chaque cycle, la quantité d'ADN est théoriquement multipliée par deux; gauche bas: mutagenèse dirigée par PCR. Une base modifiée dans les primers permet de synthétiser un fragment dans lequel la mutation est localisée à l'extrémité. On peut recombiner deux de ces PCR pour reformer par PCR le fragment complet avec la mutation voulue à l'intérieur.

4.9. 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 d'échantillons, ce qui donne accès à de nombreux gènes et permet de mettre en évidence des corrélations et donc les réseaux génétiques qui sous-tendent le fonctionnement de la cellule.

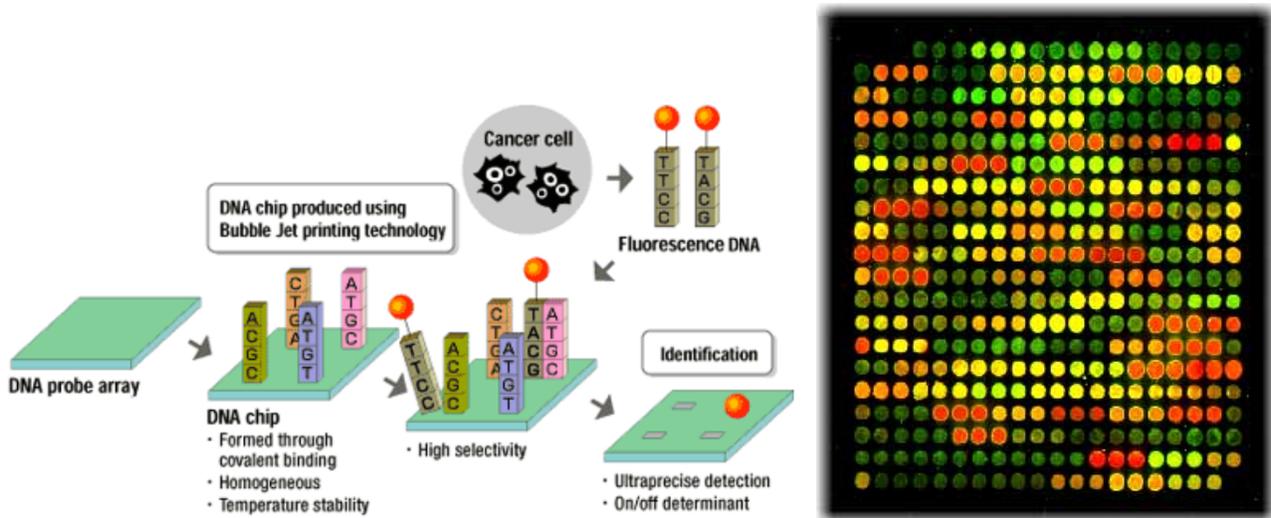


FIG. 19 – Principe et exemple de puce à ADN

Références

- [1] M. Morange. *Histoire de la Biologie Moléculaire*. Paris, 1994.
- [2] 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.
- [3] D. Voet and J.D. Voet. *Biochemistry, 3rd*. New York, 1995.
- [4] F. Jacob. *La Logique du Vivant*. Paris, 1976.
- [5] E. Schrödinger. *Qu'est-ce que la vie ?* Paris, 1993.