

Bertrand Jordan

Au commencement était le Verbe

Une histoire personnelle
de l'ADN



AU COMMENCEMENT ÉTAIT LE VERBE

Une histoire personnelle de l'ADN

Crédits photos

Image de couverture : *DNA-Music of Life* (© Chris Madden).

Page 34 : dessin © Marc Chalvin.

AU COMMENCEMENT ÉTAIT LE VERBE

Une histoire personnelle de l'ADN

Bertrand Jordan



Du même auteur

Voyage autour de Génome, le tour du monde en 80 labos. Paris : Inserm/John Libbey Eurotext, 1993.

Voyage au pays des gènes. Paris : Les Belles Lettres, 1995.

Génétique et Génome : la fin de l'innocence. Paris : Flammarion, 1996.

Les Imposteurs de la génétique (Prix Roberval 2000). Collection Science ouverte. Paris : Seuil, 2000.

Le Chant d'amour des concombres de mer. Collection Points sciences. Paris : Seuil, 2002 et 2006.

Chroniques d'une séquence annoncée. 1992-2002 : dix ans de programmes Génome. Paris : Editions EDK, 2003.

Les marchands de clones. Collection Science ouverte. Seuil, 2003.

Le clonage : fantasmes et réalité. Paris : Milan, 2004.

Thérapie génique : espoir ou illusion ? (Prix Jean Rostand 2007). Paris : Odile Jacob, 2007.

L'humanité au pluriel, La génétique et la question des races (Prix La science se livre 2009). Collection Science ouverte. Paris : Seuil, 2008.

Autisme, le gène introuvable. Paris : Seuil, 2012.

EDP Sciences
17, avenue du Hoggar
PA de Courtabœuf
91944 Les Ulis Cedex A, France
Tél. : 01 69 18 75 75
Fax : 01 69 86 06 78
www.edpsciences.org

© EDP Sciences, 2015
ISBN : 978-2-7598-1710-8

Il est interdit de reproduire intégralement ou partiellement le présent ouvrage – loi du 11 mars 1957 – sans autorisation de l'éditeur ou du Centre Français d'Exploitation du Droit de Copie (CFC), 20, rue des Grands-Augustins, 75006 Paris.

SOMMAIRE

1. Sur la route de Meyrin	7
2. Au commencement était le Verbe	10
3. Feu la Biologie Moléculaire ?.....	20
4. Les <i>fingerprints</i> de Sanger	24
5. Le printemps des cloneurs de gènes	31
6. On apprend à lire	37
7. Les vingt glorieuses de la nouvelle génétique médicale	44
8. Une idée folle	54
9. L'invasion des « puces à ADN »	61
10. Le Graal n'est plus ce qu'il était.....	67
11. Puces et maladies complexes	76
12. Le réveil des séquenceurs	84
13. Une médecine enfin personnalisée ?	93
14. Acteur ou témoin ?	99
Glossaire.....	103
Index des acronymes.....	109

Vj ku'r̥ ci g'kpvgpvkqpcm̥ 'ighv'dn̥pm

1

SUR LA ROUTE DE MEYRIN

Printemps 1965. La vieille 2CV – ma première voiture – s'époumone sur un faux plat, dans la longue ligne droite qui mène de Meyrin à Genève. J'ai vingt-cinq ans, je termine un doctorat de physique sur « La production de particules dans les collisions proton-proton à 19,6 GeV/c », au sein d'une équipe travaillant au grand accélérateur de particules du CERN (Centre européen pour la recherche nucléaire)¹. Depuis quelques mois, j'ai pris conscience que ce type de recherche ne me convient qu'à moitié, que ce travail ne correspond pas vraiment à mes aspirations. Certes, on y emploie des techniques de pointe : j'y ai fait mes débuts en informatique, alors dans ses prémisses, et dans bien d'autres secteurs (électronique, calculs de mécanique quantique...); certes, l'ambiance du groupe dans lequel je travaille est très sympathique, et la vie de « fonctionnaire international » à Genève tout à fait agréable. Mais la physique des particules a déjà un côté *Big Science* affirmé (cela ne s'est pas arrangé depuis), avec des équipes nombreuses dans lesquelles le rôle de chaque chercheur s'apparente à celui d'un petit engrenage dans une grosse machine, et une coupure presque totale entre les expérimentateurs et les théoriciens. Bref, je me sens motivé pour la recherche, mais une recherche dans laquelle je puisse moi-même interpréter mes expériences, en réaliser plusieurs chaque mois (au lieu d'une par année au CERN), et où je puisse espérer que mon travail et ma réflexion jouent un réel rôle dans le développement de la science. J'envisage donc de me réorienter vers un autre domaine scientifique, et, depuis quelques mois déjà, m'intéresse à une nouvelle venue, la « biologie moléculaire ». Quelques lectures m'ont confirmé dans cette voie, et m'ont convaincu que cette nouvelle manière d'aborder l'étude du vivant, en cherchant à identifier les molécules et les mécanismes et en privilégiant l'étude du message héréditaire dont on sait depuis peu qu'il est inscrit dans la molécule d'ADN², va sans nul doute révolutionner la biologie et ouvrir des opportunités de recherches passionnantes. De plus, beaucoup des grandes figures de cette science récente sont des physiciens reconvertis, apportant à des études jusque-là très descriptives la rigueur et, parfois, le formalisme mathématique qui leur manquaient encore. J'ai appris qu'à l'université de Genève avait été fondé un tout nouvel « Institut de biologie moléculaire », et son directeur, Edouard Kellenberger – un ancien physicien, lui aussi – a accepté de me donner un rendez-vous pour cet après-midi.

Kellenberger me reçoit en effet, et parle de ses travaux sur les « bactériophages », ces virus qui parasitent certaines bactéries et dont l'étude, lancée par Max Delbrück (encore un physicien reconvertis), a posé, grâce à un système expérimental simple et facilement manipulable, les bases de la nouvelle biologie. C'est notamment grâce à ces « phages » que Hershey et Chase [1] ont pu confirmer, en 1952, que l'ADN est bien la molécule porteuse de l'hérédité – ce qu'Avery, Macleod et McCarthy [2] avaient déjà démontré en 1944 sans que leur découverte rencontre un grand écho. Mais ce directeur est un homme très occupé, et je me retrouve rapidement à discuter dans un laboratoire

1. Voir l'*Index des acronymes*, page 109.

2. Voir le *Glossaire*, page 103.

souterrain, presque une cave, avec un petit homme rond au fort accent suisse-allemand, plus accueillant et plus prêt à me consacrer du temps. À travers mes lectures, je sais que l'une des questions centrales qui préoccupent les chercheurs de ce domaine est celle du « code génétique ». La formule des protéines qui assurent le fonctionnement de tous les organismes est, on le sait alors depuis quelques années, inscrite dans l'ADN par la suite des bases (ou nucléotides) T, A, G ou C qui se succèdent le long de chaque chaîne de cette très longue molécule. Mais les protéines sont formées d'une suite d'acides aminés, dont il existe vingt variétés (de l'alanine à la lysine en passant par la méthionine) : il doit donc exister une correspondance entre le langage de l'ADN, qui ne comprend que quatre lettres, et celui des protéines qui en emploie vingt. Une correspondance, un code, comme le code Morse qui représente chacune des vingt-six lettres de notre alphabet par une suite de « longues » et « brèves »³. Diverses expériences, dont l'interprétation est encore discutée, laissent penser que chaque acide aminé est représenté dans l'ADN par un groupe de trois bases. Mais lesquelles ?

On pense généralement que ce problème va occuper les chercheurs pendant bien des années : il va falloir déterminer la suite des acides aminés dans une protéine et celle des bases dans le gène correspondant et comparer les deux pour, peu à peu, en déduire le code qui les fait correspondre. La tâche n'est pas facile : on sait à peu près, et non sans mal, séquencer une protéine, à condition qu'elle ne soit pas trop complexe ; mais on ne sait pratiquement pas lire la suite des bases dans un ADN... Peut-être y a-t-il eu des progrès récents, car mes informations reposent sur des livres ou des articles scientifiques, et ont donc quelques mois de retard sur les avancées des laboratoires. Mais Werner Arber (c'est le nom de mon interlocuteur) ne me parle pas de cela, il me raconte ses travaux sur des enzymes bizarres, extraits de certaines bactéries, et qui semblent avoir pour rôle principal de couper l'ADN d'autres bactéries. Cela paraît extrêmement compliqué, avec des résultats difficiles à interpréter, et pour tout dire j'ai de la peine à le suivre dans ce sujet qui visiblement le passionne... Je profite finalement d'un « blanc » dans la conversation pour lui demander où on en est du déchiffrage du code génétique. Son visage s'éclaire et, avec un large sourire, il sort de la poche de sa blouse blanche une feuille ronéotée (nous sommes en 1965, avant les ordinateurs et même avant que les photocopieuses fassent partie du matériel de bureau courant) sur laquelle est inscrit ce code génétique ! D'ingénieuses expériences menées aux États-Unis par un chercheur talentueux, Marshall Nirenberg, ont fourni un raccourci pour déterminer le code, et celui-ci est maintenant en voie d'être entièrement élucidé. Voilà donc résolue une question centrale sur laquelle on imaginait que les chercheurs allaient trimer une bonne dizaine d'années – encore un exemple des progrès foudroyants de cette nouvelle biologie...

Je reprends le volant de ma guimbarde, retrouvant son odeur familière (essence, vieille huile, et légers relents de moisI – la capote n'est plus étanche depuis longtemps, et il pleut beaucoup à Genève), et reprends la route de Meyrin en ayant conscience d'avoir assisté presque en direct à une avancée décisive. J'ignore, bien sûr, que ces obscures enzymes dont m'a parlé Arber vont, dix ans plus tard, être un des éléments essentiels de la révolution du Génie génétique, et lui vaudront le prix Nobel en 1978... Mais ce que j'ai perçu suffit à me séduire, et je suis plus décidé que jamais à entamer ma réorientation. Quel contraste avec la physique des particules que je pratique depuis trois ans, au sein d'une équipe de vingt personnes qui travaille toute une année pour préparer une expérience (concevoir des appareils, des circuits électroniques, faire de fastidieux calculs avec des machines à calculer mécaniques ou avec les premiers ordinateurs), la réaliser (un mois « sur le faisceau » de l'accélérateur de particules, à faire des mesures jour et nuit)

3. Ce code Morse est un peu oublié de nos jours – mais le lecteur se souvient sans doute du signal SOS (*Save Our Souls*) figuré par trois brèves, trois longues et trois brèves, lancé pour la première fois par le Titanic en 1912 et encore utilisé aujourd'hui...

puis enfin dépouiller les résultats... tout cela pour préciser un peu la durée de vie d'une particule instable ou évaluer plus exactement sa probabilité d'apparition. Dans cette nouvelle biologie, on touche aux secrets les plus intimes de la vie, les progrès sont quasi quotidiens, et ils sont le fait d'individus ou de petites équipes, pas de monstres un peu anonymes comme le CERN. Oui, décidément, c'est là que je veux travailler... Cet après-midi d'avril 1965 fut un tournant majeur dans ma vie professionnelle, et je me suis toujours félicité d'avoir suivi le sentiment qui m'habitait alors sur la route de Meyrin : j'ai pu ainsi, durant les quarante années de ma carrière de chercheur, participer à une véritable révolution qui a transformé la biologie et a fait de cette science un acteur essentiel de notre xx^e siècle [3]. C'est un peu de cette histoire que je veux raconter dans ce livre, à travers l'éclairage tout personnel de ma vie avec l'ADN...

Références et lectures conseillées

1. Hershey A, Chase M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* 1952 ; 36 : 39-56.
2. Avery OT, MacLeod CM, McCarty M. Studies of the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J Exp Med* 1944 ; 79 : 137-58.
3. Morange M. *Histoire de la biologie moléculaire*. Paris : La Découverte, 2013.
Une excellente histoire de la biologie moléculaire, par un expert reconnu. Apporte de très nombreuses informations et une mise en perspective rigoureuse – contrairement à la vision personnelle, forcément parcellaire et incomplète, que développe le présent ouvrage.

2

AU COMMENCEMENT ÉTAIT LE VERBE

« Au commencement était le Verbe, et le Verbe était avec Dieu, et le Verbe était Dieu [...] Toutes choses ont été faites par lui, et rien de ce qui a été fait n'a été fait sans lui » : c'est ainsi que débute dans notre Bible l'évangile selon Jean. Le Verbe, la parole, le message... nous voici tout près du cœur de la biologie actuelle, la molécule d'ADN et l'information dont elle est porteuse ! Interrogation millénaire sur notre origine, interrogation plus concrète sur la manière dont se construit un être humain : comment, de l'union d'un homme et d'une femme, apparaît un enfant à la fois semblable et différent de ses deux parents ? Comment cet embryon formé au départ de quelques cellules toutes pareilles arrive-t-il à se complexifier au cours de son développement en suivant un plan très précis, pour aboutir à un organisme comportant des dizaines d'organes spécialisés, des centaines de tissus différents ? Il n'y a pas si longtemps régnait l'idée de la préformation, de la présence dans le spermatozoïde d'un « homoncule », sorte d'être humain miniature (*Figure 1*) qui allait se développer dans l'utérus de la mère, selon la métaphore de la graine et de la terre nourricière. Conception qui, tout en niant le rôle maternel, débouchait sur l'absurde, puisque dans cette hypothèse l'homoncule aurait dû lui-même contenir un homoncule qui lui-même...

Les débuts de la génétique

Une fois cette idée abandonnée pour celle d'une contribution de chaque parent aux caractères de l'enfant, restaient bien des mystères : n'était-il pas logique de penser que la descendance était un « mélange » à parts égales des deux géniteurs, et ne devait-on pas dans ce cas s'attendre à une certaine uniformisation, à tendre vers un type « moyen » au sein de nos populations ? L'apparition d'enfants aux yeux bleus, ou aux cheveux roux alors qu'aucun des parents ne possédait ces particularités restait une énigme. C'est Mendel qui éclaira en 1865 cette question en montrant, avec ses fameux petits pois, que sont transmis des déterminants – nous disons aujourd'hui des gènes – « discrets » (au sens de « non continus ») et que certains de ceux-ci sont « dominants », d'autres « récessifs ». Ainsi le croisement d'une plante à pois lisses avec une autre à pois ridés donne uniquement, à la génération suivante, des plantes hybrides portant des pois lisses – et non des pois un peu lisses, un peu ridés. En fait, ces plantes ont le « génotype » (le jeu de caractères héréditaires) lisse *et* ridé, mais le caractère ridé est « récessif » et n'est pas visible sur la plante, ne fait pas partie de son « phénotype ». Si ensuite l'on croise ces hybrides entre eux, on trouve parmi leurs descendants trois quarts de plantes à pois lisses (celles qui possèdent les déterminants lisse/lisse ou lisse/ridé) et un quart de plantes à pois ridés (celles qui possèdent le déterminant ridé/ridé) : l'assortiment des déterminants au sein de la descendance obéit aux lois de la statistique. Soulignons que Mendel n'avait aucune idée de la nature du support de l'hérédité, et que le mot gène lui-même ne devait être inventé qu'une cinquantaine d'années plus tard. Charles Darwin, qui publie en 1859 sa théorie de l'évolution, « *L'origine des espèces* », et a eu connaissance des travaux de Mendel sans apparemment leur accorder d'importance, n'en sait pas plus que

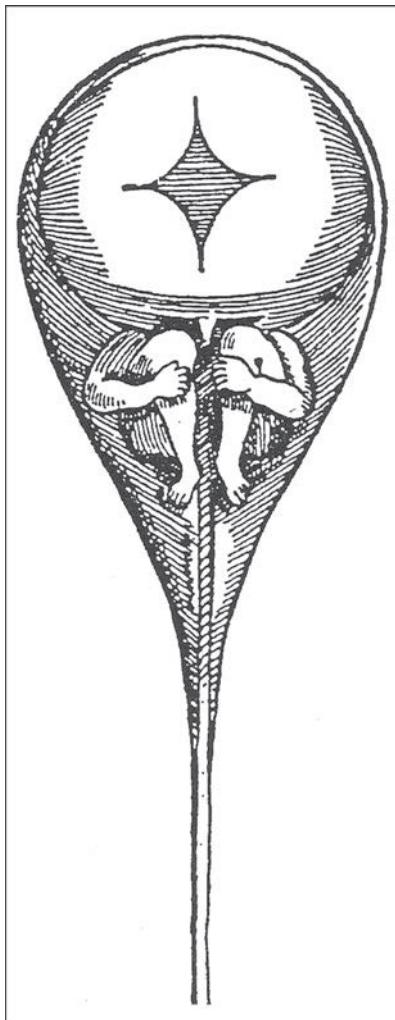


Figure 1. Représentation d'un spermatozoïde contenant un homoncule (Nicolas Hartsoeker, *Traité de dioptrique*, 1694).

ce dernier sur les mécanismes sous-jacents à l'évolution. Il ne tranche pas sur le fait de savoir si les variations qui, avec la sélection par l'environnement, sont le moteur de l'évolution, apparaissent au hasard ou sont liées au déroulement de la vie de l'individu concerné, comme le soutient Lamarck. C'est seulement après la mort de ces deux savants, et la redécouverte des travaux de Mendel au début du xx^e siècle, que se fait la synthèse que l'on appellera le néo-darwinisme. Elle reste le fondement de la biologie contemporaine même si, bien sûr, de nombreux aménagements lui ont été apportés au fil des années. Néo-darwinisme dont on peut résumer le message essentiel de la manière suivante : chaque espèce produit potentiellement beaucoup plus d'individus qu'il n'en peut survivre ; de petites variations génétiques aléatoires se produisent d'une génération à la suivante ; certaines de ces variations rendent l'individu qui les porte mieux adapté à l'environnement

présent à cette époque ; les individus ainsi favorisés ont une descendance plus nombreuse (un meilleur « succès reproductif ») ; les versions de gènes correspondantes s'imposent donc progressivement au sein de la population. C'est ainsi, par exemple, que l'on comprend aujourd'hui pourquoi les Européens, dont les ancêtres, venus d'Afrique il y a 60 ou 70 000 ans, étaient sûrement noirs de peau, ont un épiderme clair : les mutants dépigmentés apparus par hasard il y a des dizaines de milliers d'années ont prospéré, grâce à leur meilleure capacité à synthétiser la vitamine D malgré la faible luminosité¹, et ont progressivement supplanté leurs congénères à la peau foncée.

De quoi sont faits les gènes ?

Jusqu'au milieu du xx^e siècle et malgré les très grands progrès de la génétique, la nature des gènes, leur support matériel, leur mode de transmission et de mise en œuvre restaient un grand mystère. Des travaux menés notamment sur la drosophile (la mouche du vinaigre des généticiens) indiquaient que les gènes étaient disposés de manière linéaire sur des filaments que l'on pouvait parfois voir dans le noyau des cellules et que l'on appela les chromosomes. On avait pu construire une « carte génétique », une sorte de balisage des gènes, en analysant de nombreux croisements entre mouches appartenant à des souches différentes ; l'observation des chromosomes géants présents dans certains organes de la mouche (*Figure 2*) permit alors de relier cette carte génétique avec la morphologie des chromosomes et les bandes plus ou moins foncées observées au microscope.

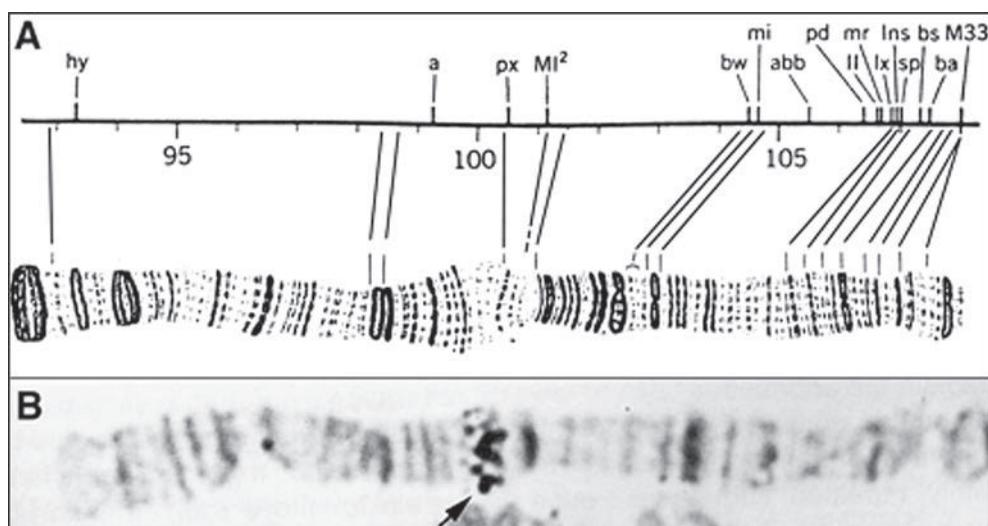


Figure 2. Gènes et chromosomes chez la drosophile. **A.** Carte génétique (établie par l'étude de croisements entre mouches appartenant à différentes souches) indiquant la position relative de différents gènes (par exemple, *hy* désigne le gène Humpy dont la mutation entraîne une déformation du thorax de la mouche), et (en dessous) sa correspondance avec les « bandes » que l'on peut observer sur les chromosomes géants des glandes salivaires de la mouche. **B.** Microphotographie d'un tel chromosome : on y distingue effectivement des bandes, mais elles sont bien moins nettes que sur leur représentation idéalisée.

1. La phase finale de la synthèse de la vitamine D se fait dans la peau et est favorisée par la lumière. L'insuffisance en vitamine D entraîne le rachitisme. Une alimentation abondante et variée apporte suffisamment de vitamine D pour que sa synthèse par l'organisme ne soit plus nécessaire.

Mais en quoi consistaient ces gènes ? On savait qu'ils renfermaient chacun la formule de fabrication d'une protéine, selon l'adage célèbre « un gène, une protéine, une fonction »². On avait aussi compris que les protéines, ces grosses molécules très complexes qui assurent l'intendance au sein de nos cellules et de notre corps, ces enzymes qui dégradent les sucres pour fournir de l'énergie, qui transportent l'oxygène à travers nos vaisseaux, qui assurent la coagulation du sang en cas de blessure... que ces protéines étaient formées d'une suite d'acides aminés, une sorte de collier de perles linéaire se repliant sur lui-même pour prendre une forme caractéristique lui permettant d'assurer son rôle dans l'organisme. Les chromosomes devaient donc, d'une manière ou d'une autre, renfermer la formule des milliers de protéines nécessaires à la construction et au fonctionnement d'un être vivant – bactérie, végétal, animal ou homme.

L'ADN et la double hélice : le tournant décisif

L'analyse chimique ne donnait pas la réponse à cette question : les chromosomes contiennent, en quantité à peu près équivalente, des protéines et de l'ADN (acide désoxyribonucléique). La composition de cet ADN, connu depuis sa découverte par Friedrich Miescher, au dernier quart du XIX^e siècle (il l'appelait « nucléine »), n'en faisait pas, *a priori*, une molécule très intéressante : il était formé d'une suite monotone de quatre entités chimiques (les « bases » adénine, guanine, cytosine et thymine, A, G, C et T), et l'on pensait que sa structure était celle d'une répétition d'unités formées d'un assemblage de ces quatre éléments. Une telle répétition de motifs identiques pouvait difficilement renfermer un message, et, très logiquement, l'opinion générale situait le support de l'hérédité au niveau des protéines, bien plus complexes et multiformes. Et lorsque Avery, McCarthy et McLeod, en 1944 [4], démontrèrent que c'est l'ADN qui porte l'information génétique, ils se heurtèrent à beaucoup d'incrédulité et leurs travaux eurent peu d'écho. Ils étaient pourtant parfaitement concluants : ces chercheurs avaient pu transférer un caractère génétique d'une souche de bactéries à une autre à l'aide d'un extrait bactérien dont ils pouvaient éliminer soit les protéines, soit l'ADN. Et c'est seulement lorsque l'extrait contenait l'ADN que l'expérience réussissait. Malgré leur qualité technique, ces expériences restèrent largement ignorées jusqu'à ce qu'elles soient reproduites dans un autre système par Hershey et Chase, en 1952 [5], et surtout jusqu'à ce que la structure en double hélice découverte par Watson et Crick en 1953 [6] soit publiée et éclaire tout ce domaine en montrant comment l'ADN pouvait renfermer de l'information et également assurer sa propre reproduction, sa duplication.

La double hélice est donc bien une de ces avancées qui illuminent un secteur de la science (en l'occurrence, toute la biologie) et changent fondamentalement notre manière de penser la nature. Sa découverte, racontée sans détour dans le livre éponyme de Jim Watson [7], fut le fruit de rencontres, d'imagination créative, d'approches sans tabous et parfois sans scrupules, et aboutit à ce superbe modèle d'une double hélice formée de deux brins « complémentaires », la suite des bases T, A, G ou C présente sur un brin s'appariant avec les bases complémentaires (A, T, C ou G) présentes sur l'autre (*Figure 3*).

2. Adage dont on sait aujourd'hui qu'il est faux en ce sens qu'il constitue une simplification abusive de la réalité – mais une simplification qui a bien servi à faire avancer la science durant plusieurs décennies.

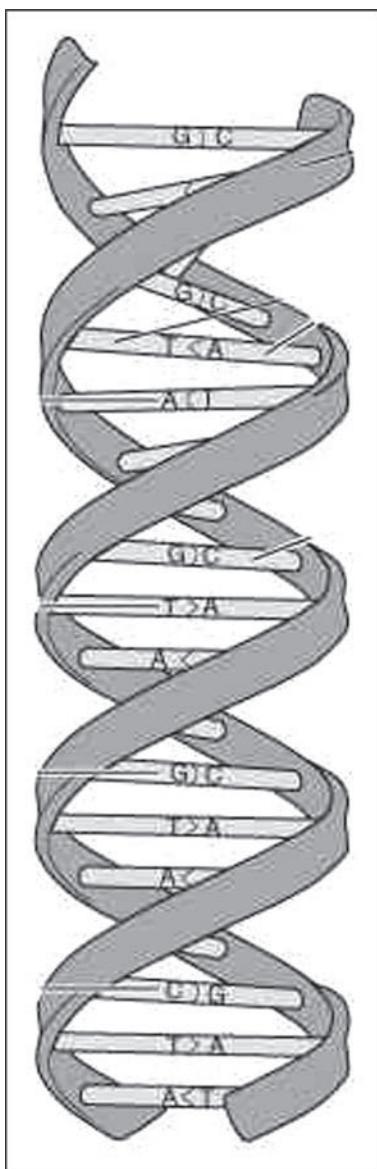


Figure 3. La double hélice de l'ADN (représentation très schématique). Les bases T, A, G, C constituent les « barreaux » de l'hélice, appariées selon la règle T-A et G-C. La séquence (suite des bases) portée par un brin détermine donc celle qui est portée par l'autre. Les deux chaînes externes constituent le « squelette ». Cette structure suggère que pour la réPLICATION les deux chaînes se séparent et que chacune sert de « moule » pour en fabriquer une nouvelle – à l'arrivée on a deux doubles hélices identiques à cette du départ (voir *Figure 4*).

Du coup la forme de la molécule est parfaitement régulière quelle que soit la suite des bases le long d'une chaîne : elle peut donc renfermer n'importe quel message, et les deux brins portent la même information : la séquence de l'un peut être déduite de celle de l'autre puisqu'en face d'un C on a obligatoirement un G (et vice versa), en face d'un A un T. On imagine immédiatement comment une telle double hélice peut être dupliquée (voir *Figure 4*) : comme l'indiquent Watson et Crick à la fin de leur bref article, dans un magistral *understatement* à la britannique, « Il ne nous a pas échappé que ce modèle suggère immédiatement un mode de reproduction de l'information génétique » *Indeed...*



Figure 4. RéPLICATION DE L'ADN (lors de la division d'une cellule). La double hélice s'entrouvre et une nouvelle chaîne est formée sur chacun de ces deux brins. À l'arrivée, on a deux doubles hélices identiques à celle du départ.

Le « dogme central »

C'est donc à cette date que l'on peut situer le vrai début de la « biologie moléculaire », qui allait constituer le secteur le plus dynamique des sciences du vivant pendant deux décennies avant de les envahir toutes entières. Les années qui suivirent la double hélice virent l'élucidation, dans les grandes lignes, du mécanisme de la réplication de l'information biologique, à savoir comment une cellule transmet l'intégralité de son patrimoine génétique aux deux cellules filles qu'elle forme en se divisant.

On put aussi comprendre les modalités de l'expression génique, comment la formule d'une protéine inscrite dans l'ADN est lue pour diriger l'assemblage de la protéine correspondante (*Figure 5*). Sans rentrer dans les détails, la réplication implique la séparation des deux brins de la double hélice, la synthèse sur chaque brin d'un brin

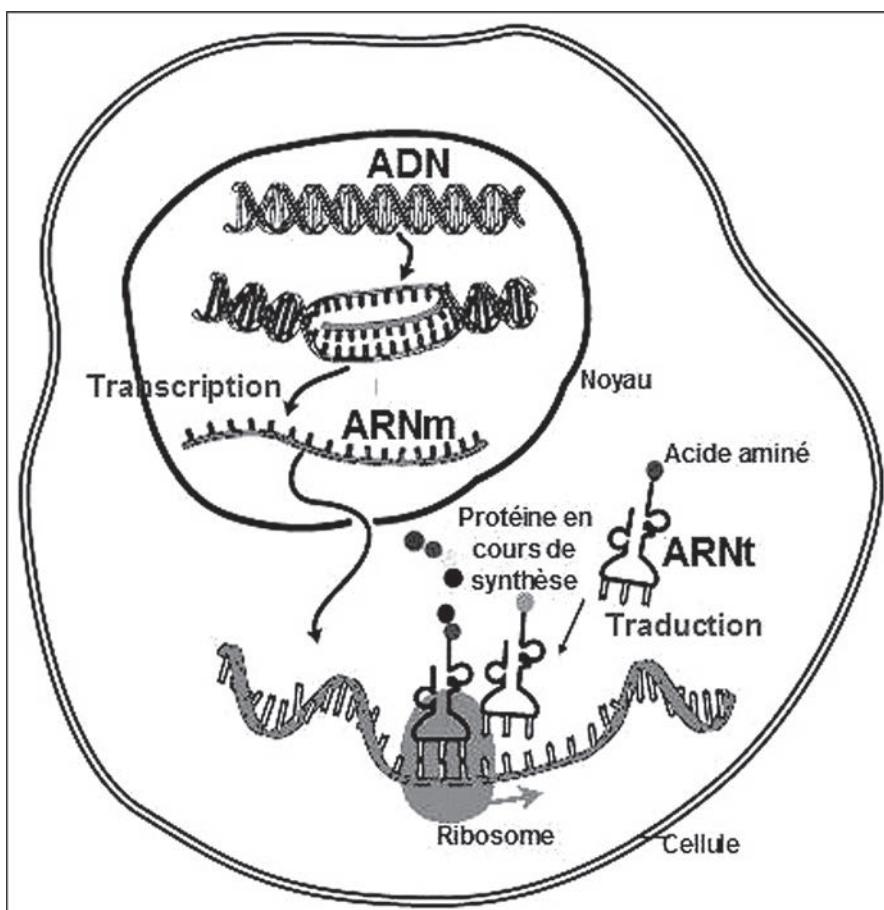


Figure 5. Expression de l'ADN. Celle-ci passe dans un premier temps par la « transcription », fabrication d'une molécule d'ARN messager (ARNm, « copie de travail » de l'ADN), puis par la « traduction » au cours de laquelle le message est lu par les ribosomes et la protéine correspondante synthétisée.

complémentaire, aboutissant à deux exemplaires de la molécule initiale ; l'expression, elle, passe par la fabrication dans le noyau de la cellule (où réside l'ADN) d'une copie de travail du gène, l'ARN messager³, qui passe ensuite dans le cytoplasme où son message est lu par des organites spécialisés (et dont nous reparlerons), les ribosomes, avec l'aide d'adaptateurs, les ARN de transfert, qui associent une séquence de trois bases avec un acide aminé. Les ribosomes, en se déplaçant le long de l'ARN messager, fabriquent la chaîne d'acides aminés qui se replie ensuite sur elle-même pour former la protéine fonctionnelle (*Figure 5*). L'ensemble constitue ce que Francis Crick a appelé le « dogme central » de la biologie moléculaire : ADN → ARN → protéine [8] – dogme qui devait subir quelques modifications une dizaine d'années plus tard.

L'établissement de ce schéma fut l'œuvre d'un ensemble de pionniers (parmi lesquels les Français François Jacob, Jacques Monod et François Gros) encore très minoritaires parmi les biologistes et biochimistes « classiques », et d'ailleurs souvent vus avec une certaine défiance par ceux-ci. J'avais par exemple constaté que l'édition parue en 1958 d'un manuel de biochimie faisant autorité (celui de Fruton et Simmons) [9] mentionnait à peine la double hélice (ourtant publiée cinq ans plus tôt) en la qualifiant d'« ingénieuse spéulation »...

L'élucidation du code génétique

L'un des tours de forces de cette période fut le déchiffrage, en 1964/1965, du « code génétique », de la correspondance entre le langage de l'ADN et celui des protéines – résultat que je devais découvrir lors de ma visite à Werner Arber, à Genève. Rappelons que l'ADN comporte quatre lettres, T, A, G et C, alors que les protéines sont formées à partir de vingt acides aminés différents, de l'arginine à la valine en passant par la méthionine et la tyrosine... Ces acides aminés sont de petites molécules présentant de légères différences, et c'est l'ordre dans lequel ils sont disposés le long de la chaîne de la protéine, sa séquence, qui va donner à cette dernière sa forme et ses propriétés. Bien sûr, on peut tout écrire avec quatre lettres, avec vingt-six comme dans notre alphabet ou avec deux (0 et 1) dans nos ordinateurs... encore faut-il savoir comment on passe d'un alphabet à l'autre. Une série d'expériences très élégantes résolut le problème en quelques mois et montra que l'ADN codait chaque acide aminé par un groupe de trois lettres, un « codon » formé de trois bases successives dans l'ADN. Comme on peut former 64 codons si l'on dispose de quatre lettres différentes, ce code est « dégénéré », c'est-à-dire qu'un même acide aminé peut correspondre à plusieurs codons différents. Par exemple, l'acide aminé leucine peut être codé par CTA, CTC, CTT ou CTG⁴ ; en revanche, la méthionine a un seul codon, ATG (*Figure 6*). Ce sont ces codons qui sont reconnus, grâce au ribosome, par les ARN de transfert qui apportent l'acide aminé correspondant à la protéine en formation.

3. L'ARN (acide ribonucléique) a une structure très proche de celle de l'ADN (acide désoxyribonucléique), la nature de la chaîne est légèrement différente et les bases dans l'ARN sont A, G, C et U (uracile) au lieu de A, G, C et T (thymine).

4. Dans la copie du gène portée par l'ARN messager, la base qui correspond à T (thymine) dans l'ADN est l'uracile (U), structure légèrement différente mais ayant les mêmes propriétés. Les codons mentionnés sont donc, dans l'ARN, CUA, CUU, CUU et CUG.

Seconde base						
	U	C	A	G		
Première base	U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G
	C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
	A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
	G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

Figure 6. Le code génétique. Dans chaque case, l'acide aminé correspondant au codon (première base à gauche, deuxième base en haut, troisième base à droite – par exemple UUU code pour l'acide aminé Phe (phénylalanine). Le codon AUG (méthionine) est aussi le signal de « début de message » ; les codons UGA, UAA et UAG signalent la fin de message (STOP). Dans l'ADN, le U (uracile) est remplacé par la thymine (T) qui a les mêmes propriétés.

Ainsi, en 1965, les éléments principaux de la nouvelle biologie étaient en place, on savait (dans les grandes lignes) où et comment était inscrit le message héréditaire reçu de nos parents et que nous transmettons à nos enfants, et de quelle manière ces informations étaient lues pour produire les différentes protéines nécessaires à la construction et au fonctionnement d'un organisme. En une dizaine d'années à peine, les mécanismes de base de la Vie avaient été dévoilés : avancées spectaculaires marquant le début d'une véritable révolution qui allait se poursuivre au cours des décennies suivantes.

Références et lectures conseillées

4. Avery OT, MacLeod CM, McCarty M. Studies of the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J Exp Med* 1944 ; 79 : 137-58.
Première démonstration que c'est l'ADN qui porte l'information génétique – largement ignorée à l'époque.

5. Hershey A, Chase M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* 1952 ; 36 : 39-56.
Nouvelle démonstration mieux acceptée que la précédente.
6. Watson JD, Crick F. A structure for deoxyribose nucleic acid. *Nature* 1953 ; 171 : 737-8.
Un bref article qui marque le vrai début de la biologie moléculaire.
7. Watson JD. *The double helix. A personal account of the discovery of the structure of DNA*. Atheneum Press : New York, 1968 (et plusieurs éditions plus récentes).
Édition française la plus récente : *La double hélice*. Paris : Éditions Robert Laffont, 2003.
Un récit passionnant et révélateur, écrit sans langue de bois, qui fit scandale à l'époque.
8. Crick F. Central dogma of molecular biology. *Nature* 1970 ; 227 : 561-3. *Discussion et historique du concept qui n'a pas été formellement publié lorsqu'il fut proposé quelques années plus tôt.*
9. Fruton JS, Simmonds S. *General biochemistry*. New York : John Wiley and Sons, 1958.
Un manuel de biochimie qualifiant, en 1958, le modèle de Watson et Crick d'« ingénieuse spéculation » : la double hélice n'a pas été immédiatement acceptée...

Thrombose : formation d'un caillot (thrombus) dans le réseau veineux, avec des conséquences souvent graves (phlébite, embolie pulmonaire, accident vasculaire cérébral...).

Transgénique : organisme dont l'**ADN** a été modifié (généralement par adjonction d'un gène provenant d'une autre espèce) au niveau de la lignée germinale, donc en quelque sorte définitivement. Concerne surtout des végétaux, aussi appelés « organismes génétiquement modifiés » ou OGM.

INDEX DES ACRONYMES

AFM : Association Française contre les Myopathies, organisatrice du Téléthon annuel et intervenante importante dans la recherche médicale française.

CERN : Centre Européen pour la Recherche Nucléaire, créé en 1952, le plus important centre de recherche mondial dans ce domaine.

CIML : Centre d'Immunologie de Marseille-Luminy, institut mixte **Inserm/CNRS** créé en 1976 sur le campus de Luminy à Marseille, un des meilleurs laboratoires d'immunologie mondiaux.

CNRS : Centre National de la Recherche Scientifique, acteur majeur intervenant dans tous les secteurs de la science en France. Son département de Biologie est d'une taille équivalente à l'**Inserm**.

DGRST : Délégation Générale à la Recherche Scientifique et Technique, organisme interministériel qui, de 1960 à 1975, a joué un grand rôle dans le développement de la biologie moléculaire en France.

DOE : *Department Of Energy*, équivalent aux États-Unis de notre Commissariat à l'Énergie Atomique (CEA) et ayant comme lui une activité notable en recherche biologique.

FDA : *Food and Drug Administration*, l'organisme qui, aux États-Unis, doit approuver médicaments et dispositifs médicaux avant leur commercialisation.

INCa : Institut National du Cancer, agence sanitaire et scientifique de l'État, qui développe l'expertise et finance des projets dans le domaine des cancers.

Inserm : Institut National de la Santé et de la Recherche Médicale, principal organisme de recherche français spécialisé pour la biologie et la recherche médicale.

MRT : à certaines époques, ministère de la Recherche et de la Technologie (souvent réduit à un secrétariat d'État rattaché au ministère de l'Éducation nationale).

NIH : *National Institutes of Health*, équivalent Nord-Américain de l'**Inserm**.

NICE : *National Institute for Health and Care Excellence*, organisme britannique chargé d'évaluer le rapport coût/efficacité des traitements et d'en recommander (ou non) la prise en charge financière.