

Le chemin hamiltonien codé par des séquences d'ADN

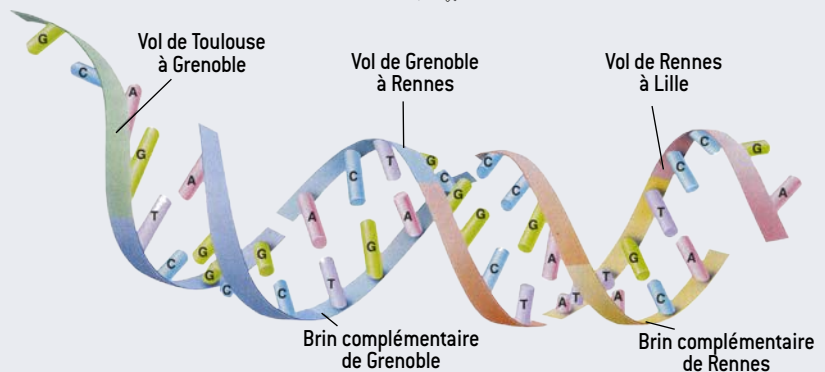
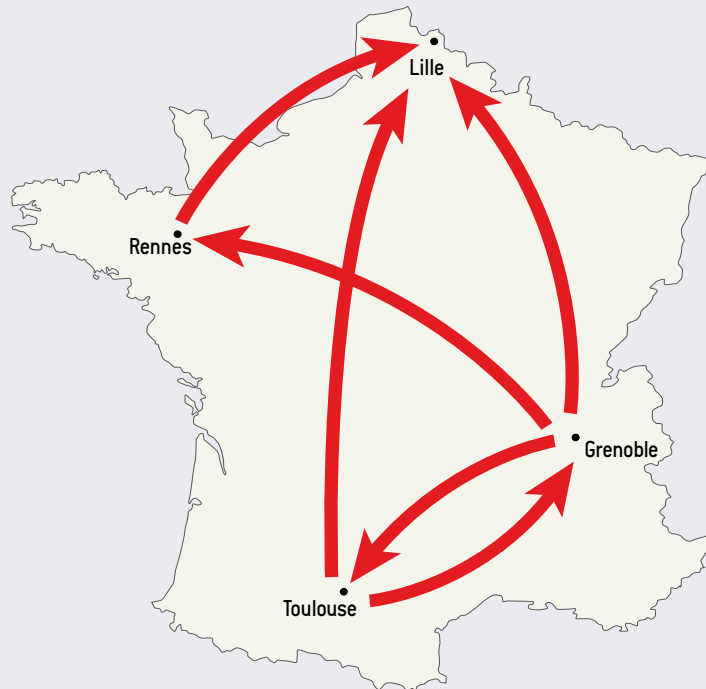
La méthode conçue par l'informaticien américain Leonard Adleman pour résoudre le problème des chemins hamiltoniens est expliquée à partir d'un exemple (emprunté à son article « Calculer avec l'ADN », *Pour la Science*, n° 252, pp. 56-63, octobre 1998).

Imaginons que les vols reliant quatre villes soient ceux dessinés sur le schéma ci-contre. On peut par exemple aller de Grenoble à Lille, mais pas l'inverse. On cherche un chemin partant de Toulouse et arrivant à Lille, qui passe par chacune des quatre villes une fois et une seule.

On choisit pour chaque ville une séquence d'ADN. Pour Toulouse, par exemple, c'est la séquence ACTTGACG. On nomme « début » les quatre premiers nucléotides, ACTT, et « fin » les quatre derniers, GACG. On peut alors coder chaque vol en mettant bout à bout la fin de la séquence associée à la ville de départ et le début de la séquence associée à la ville d'arrivée (voir le tableau) et chaque chemin combinant plusieurs vols sera représenté par l'enchaînement des séquences d'ADN correspondantes.

À chaque ville correspond aussi une séquence complémentaire, obtenue en échangeant C et G, et A et T. L'utilisation de ces séquences complémentaires permet de filtrer, dans une soupe contenant au départ les molécules codant tous les chemins possibles jusqu'à la longueur 3 au moins (correspondant au trajet reliant les quatre villes), les molécules qui correspondent aux chemins passant par toutes les villes.

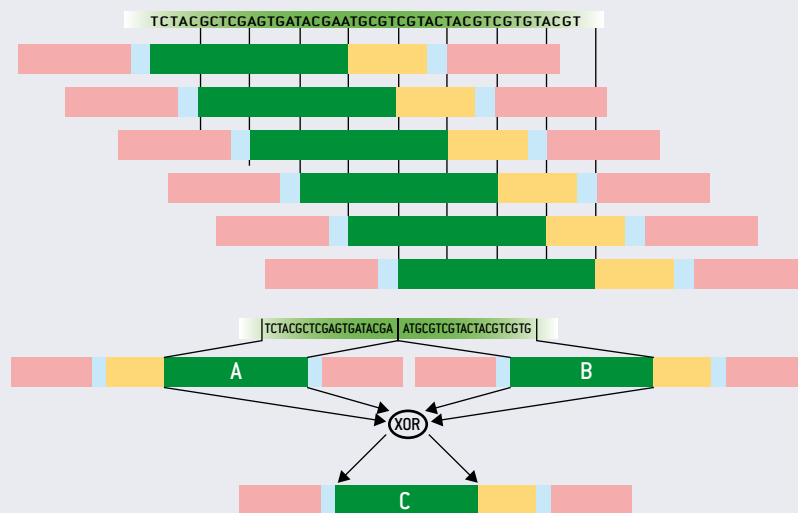
Dans notre problème, il y a un chemin hamiltonien entre Toulouse et Lille, qui sera le seul retenu ayant la longueur voulue et passant par chaque ville. Il passe dans l'ordre par Toulouse, Grenoble, Rennes et Lille. Il est représenté par la séquence GCAGTCGGACTGGGCTA-TGTCCGA, obtenue à l'issue des manipulations de la soupe initiale opérant les filtrages.



Ville	ADN	Complémentaire
TOULOUSE	ACTTGACG	TGAACGTC
GRENOBLE	TCGGACTG	AGCCTGAC
RENNES	GGCTATGT	CCGATACA
LILLE	CCGAGCAA	GGCTCATT
Vol	Code ADN du vol	
TOULOUSE-GRENOBLE	GCA ^G TC ^G GG	
TOULOUSE-LILLE	GCA ^G CC ^G GA	
GRENOBLE-RENNES	ACT ^G GG ^G CT	
GRENOBLE-LILLE	ACT ^G CC ^G GA	
GRENOBLE-TOULOUSE	ACT ^G ACT ^T	
RENNES-LILLE	AT ^G TCC ^G A	

Mémorisation dans l'ADN

James Bornholt et ses collègues ont conçu et mis en application un procédé de mémorisation massive d'informations exploitant les séquences d'ADN. Le procédé se fonde sur l'inscription de couples clé-données, ce qui permet un accès direct aux informations enregistrées analogue à ce qui existe dans la mémoire d'un ordinateur utilisant un système d'adresses : pas besoin, contrairement au cas d'une bande magnétique, de tout parcourir pour retrouver une information.



Le défi relevé était particulièrement difficile pour plusieurs raisons. D'abord, les méthodes de synthèse et de lecture de l'ADN sont imparfaites, avec des taux d'erreurs de l'ordre de 1 % par nucléotide. Les séquences se dégradent aussi progressivement. Il fallait donc concevoir un système fortement redondant. Selon le type de données (image, son, texte), la présence d'une erreur dans les données est plus ou moins grave ; il fallait alors prévoir un système ajustable de redondances.

C'est ce que propose leur procédé. Le point le plus important de leur travail est cependant la mise en place d'un codage qui autorise l'accès à certaines données sans avoir à tout lire. Le système d'adressage conçu utilise la réaction en chaîne par polymérase (PCR, *polymerase chain reaction*) qui amplifie, et donc donne accès, uniquement aux données souhaitées. Bien évidemment, cette lecture choisie accélère l'accès aux données puisqu'elle évite de parcourir tout l'ADN

d'une base de données moléculaire. Notons toutefois que les temps d'accès restent bien supérieurs à ceux de l'électronique et se comptent en heures.

L'information sur les séquences d'ADN doit être inscrite avec des redondances (car la transcription et la lecture amènent des erreurs), ce que plusieurs procédés permettent. On inscrit la longue séquence en s'arrangeant pour que chaque brin d'ADN (plus court que la séquence qu'on veut inscrire) contienne une portion de la séquence à mémoriser et qu'il y ait des chevauchements entre les brins (*schéma du haut*). L'idée utilisée par l'équipe de James Bornholt consiste, pour inscrire la séquence A et la séquence B, à inscrire A sur un brin, B sur un autre et le OU exclusif (XOR) des séquences A et B (après les avoir traduites en binaire), noté C, sur un troisième brin (*schéma du bas*). Les trois séquences A, B, C ont alors pour propriété que la connaissance de deux quelconques d'entre elles (A et B, A et C, B et C) suffit pour retrouver A et B.

épouse : "Bon sang ! Ces choses peuvent calculer." Je ne dormis pas de la nuit, cherchant un moyen d'utiliser l'ADN pour résoudre des problèmes. » La mise au point de mécanismes de calcul fut cependant délicate, car ce que fait la cellule se décompose en opérations élémentaires qui ne permettent pas facilement d'effectuer des opérations complexes. En plus de la copie de brins d'ADN par l'enzyme ADN polymérase, on dispose d'enzymes nommées ligases qui raboutent des molécules d'ADN et d'« enzymes de restriction » qui coupent les molécules d'ADN à des endroits précis. L'enzyme EcoRI, par exemple, coupe l'ADN de la bactérie *Escherichia coli* après le G de la séquence de nucléotides GAATTC.

Des soupes pour mémoriser et calculer

Poursuivant son travail nocturne, dans les jours qui suivirent son eureka, Leonard Adleman conçut et réussit à faire fonctionner un procédé de calcul à base d'ADN et d'enzymes qui résout le problème combinatoire de la recherche des chemins hamiltoniens dans un graphe. Il s'agit de trouver un chemin dans un graphe qui passe par chaque nœud du graphe une fois et une seule. Organiser une tournée visitant des villes fixées à l'avance sur un réseau routier ou aérien revient à résoudre un tel problème.

Ce problème appartient à la classe des problèmes NP-complets. Cela signifie deux choses : d'une part, qu'on ne connaît aucun algorithme rapide qui le résout dès que le nombre de nœuds du graphe devient important (« rapide » signifie que le temps d'exécution est une fonction polynomiale du nombre de nœuds du graphe) ; d'autre part, si l'on connaissait un moyen de résoudre rapidement ce problème, on pourrait aussi résoudre rapidement, c'est-à-dire en temps polynomial, tous les problèmes de la classe NP (problèmes dont on peut vérifier qu'une solution est correcte en temps polynomial).

L'algorithme proposé et mis en œuvre par Leonard Adleman en 1994 avec des soupes moléculaires réelles est le suivant :

- On engendre toutes les séquences possibles correspondant à des chemins

Le potentiel de la mémoire à ADN

La quantité totale d'informations numériques sur Terre est aujourd'hui de l'ordre de 10^{20} (1 Zo = 1 zettaoctet = 10^{21} caractères), soit 10^{22} caractères, ou 5 milliards de milliards de pages imprimées de 2 000 caractères.

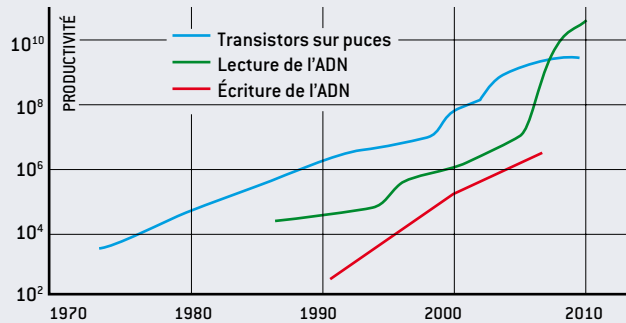
Si l'augmentation se poursuit au rythme des dernières décennies, nous serons incapables, par les moyens utilisés aujourd'hui, de répondre à la demande d'archivage de ces données.

Aussi étudie-t-on des moyens de stockage autres que magnétiques ou optiques. Les méthodes exploitant l'ADN pourraient être la solution pour le stockage massif à long terme, d'autant que la densité d'information envisageable avec de l'ADN est bien supérieure à celle des meilleures mémoires magnétiques ou optiques. Ce n'est peut-être pas étonnant : la mémorisation par l'ADN se fait à l'échelle de la molécule, ce qui n'est pas le cas en électronique ou en optique.

Les trois courbes calculées par Rob Carlson (un investisseur américain dans

la bioéconomie) montrent que le rythme de croissance de la puissance informatique, qui suit la loi de Moore (doublement tous les 18 mois à peu près depuis plus de quarante ans), reste plus faible que celui de l'efficacité des dispositifs d'écriture et de lecture de l'ADN. Cependant, si un moyen de stockage permet d'écrire et de lire vite, alors il ne conserve pas l'information longtemps.

Comparons les performances des quatre techniques que sont les mémoires



flash (clés USB, disques SSD), les disques durs des ordinateurs, les bandes magnétiques et l'ADN. Les temps d'accès se mesurent respectivement en millisecondes, en dizaines de millisecondes, en minutes et en dizaines d'heures.

Les durées de conservation sont respectivement de l'ordre de 5 ans, 5 ans, 15-30 ans et plusieurs siècles. Le stockage sur de l'ADN, bien que lent, apparaît donc beaucoup plus durable.

aléatoires selon un codage [expliqué dans l'encadré page 79], et ce jusqu'à la taille n , le nombre de nœuds du graphe.

- On filtre les séquences selon leur masse pour ne garder que les chemins de longueur n .

- On filtre en plusieurs étapes les séquences restantes en utilisant les mécanismes moléculaires connus, pour ne garder dans la soupe que les chemins passant une fois et une seule par chaque nœud.

- Les séquences qui restent indiquent des chemins hamiltoniens.

Cette méthode exploite le parallélisme des réactions chimiques qui, d'un endroit à l'autre de la soupe moléculaire, « calculent » simultanément, c'est-à-dire essaient des chemins différents en éliminant les chemins insatisfaisants. En théorie, le problème est résolu en un nombre d'étapes de réactions chimiques qui croît linéairement en fonction du nombre de nœuds du graphe. On disposerait donc d'une méthode de résolution en temps polynomial d'un problème NP-complet, et, en conséquence, d'une méthode polynomiale pour résoudre

n'importe quel problème de la classe NP, ce qui serait une avancée considérable.

Il faut cependant tempérer notre enthousiasme, car la quantité d'ADN nécessaire augmente exponentiellement en fonction de la taille du problème, à cause de la première étape de l'algorithme.

Plus que la masse de la Terre !

Une machine à ADN qui résoudrait le problème pour un graphe de 200 sommets pèserait environ 10^{25} kilogrammes, soit plus que la masse de la Terre ! L'algorithme impose la mise à disposition d'un très grand nombre de types distincts de la molécule d'ADN qui participeront à la réaction, car il faut que chaque chemin possible de longueur n soit présent au départ. Or les préparer n'est pas simple et prend du temps.

De plus, avec les méthodes utilisées, il existe une probabilité non nulle que certains chemins ne soient pas représentés et donc que l'on passe à côté d'une solution du problème. Ces obstacles n'ont pas permis de

tirer du procédé de Leonard Adleman une méthode plus efficace que celles des machines électroniques pour le problème des chemins hamiltoniens.

Depuis la création de ce premier algorithme, tout un domaine scientifique nouveau s'est développé. Ses chercheurs se retrouvent périodiquement au colloque *DNA computing and molecular programming*, dont la 22^e édition s'est réunie en septembre 2016 à Munich. Toutes sortes d'algorithmes à ADN sont aujourd'hui maîtrisés et ont donné lieu à des mises en œuvre pratiques à petites échelles.

On a en particulier proposé des systèmes assimilables à des ordinateurs où l'on entre des données sous la forme de soupes moléculaires et où, après le calcul, on lit le résultat à nouveau en analysant les molécules fabriquées pendant les phases de réaction et les longues étapes de filtrage [voir l'article de Y. Benenson et al. dans la bibliographie]. On a aussi conçu un système qui joue au jeu tic-tac-toe, une sorte de morpion sur une grille 3×3 [voir « Des assemblages d'ADN rompus au jeu et au travail », *Pour la Science*, janvier 2009].

Aujourd'hui pour certains problèmes combinatoires particuliers tels que celui de l'affectation des tâches, les chercheurs, notamment Jian-Jun Shu et ses collègues à Singapour, espèrent que les ordinateurs à ADN seront plus efficaces que les ordinateurs classiques. Dans ce problème, on se donne une série de travaux à exécuter et d'agents aux compétences variables selon les tâches auxquels on les affecte (confier la tâche T à l'agent A à un coût C) et l'on cherche à savoir comment assigner une tâche à chaque agent en minimisant le coût total.

Plus surprenant, des algorithmes de multiplication rapide fondés sur des méthodes moléculaires ont aussi été décrits, notamment par Aran Nayebe, de l'université Stanford.

Théorie... et pratique?

En théorie, la puissance des machines à ADN est énorme. Avec un litre de liquide contenant 5 grammes d'ADN, on disposerait d'une mémoire de plus de 10^{21} octets et d'une puissance de calcul évaluée à environ 10^{17} opérations par seconde, soit 100 pétaflops, ce qui est supérieur aux performances du plus puissant ordinateur actuel, le chinois Tianhe-2, qui ne réalise que 33 pétaflops.

Ce potentiel est cependant délicat à exploiter et aujourd'hui, malgré tous les efforts faits, les ordinateurs à base d'électronique sont bien plus faciles à programmer, calculent mieux et plus vite. On pourra peut-être les surpasser en utilisant des soupes à ADN et il faut continuer à explorer cette voie. Mais ce n'est pas facile, car du côté électronique les progrès sont réguliers et donc les systèmes à ADN luttent contre un concurrent de plus en plus difficile à rattraper !

Notons aussi qu'il n'est pas absurde de penser que, dans le cadre de la physique classique, un modèle de calcul qui résoudrait en temps polynomial les problèmes NP-complets ne peut exister matériellement. Les gains envisageables avec les calculateurs à ADN risquent donc d'être réservés à des problèmes particuliers et ne conduire qu'à des progrès limités, contrairement aux espoirs initiaux. Même les ordinateurs quantiques, d'après le spécialiste Scott Aaronson,

ne pourront pas résoudre en temps polynomial les problèmes NP-complets, qui semblent condamnés à rester hors de portée de tout ce que nous pouvons concevoir comme calculateurs. Ne rêvons pas trop !

Mémoires massives et durables

Récemment, en perfectionnant les méthodes de stockage d'informations par l'ADN, on a cependant réalisé des exploits qui montrent que les mécanismes de stockage de données mis au point par l'évolution du vivant pourraient être utiles et que la manipulation de l'ADN nous donnera sans doute les clés permettant la mémorisation, sur le très long terme, de quantités massives de données.

En effet, lorsqu'on souhaite garder pour des siècles des quantités importantes d'informations, on sait que ni les disques optiques (*Blu-ray* ou autres) ni les mémoires magnétiques (bandes, clés USB à mémoire flash, disques souples ou durs à plateau rotatif, ou disque SSD – *Solid-state drive*) ne sont vraiment fiables. À l'opposé, des expériences réalisées avec de l'information stockée sous forme de séquences d'ADN indiquent que l'on peut s'y fier.

En 2015, Robert Grass et son équipe, de l'École fédérale polytechnique de Zurich, ont inscrit des informations sur de l'ADN placé dans des matériaux inorganiques. Ils ont utilisé des codes correcteurs d'erreurs pour introduire de la redondance dans l'information inscrite et corriger à la lecture les petites dégradations susceptibles de se produire.

Leur expérience a consisté précisément à traduire 83 kilooctets (83 000 caractères) d'informations sur 4 991 segments d'ADN, chacun long de 158 nucléotides, qui ont été encapsulés dans de la silice. Des traitements de vieillissement accéléré ont été appliqués pour mesurer la vitesse de dégradation de l'ADN. L'information originale a été lue sans erreur, même après avoir chauffé la silice à 70° C pendant une semaine, ce qui est thermiquement équivalent au stockage d'informations sur l'ADN pendant 2 000 ans à température ambiante. On en a déduit que des données peuvent être archivées dans l'ADN pendant des millénaires.

D'autres expériences ont été effectuées et l'on s'est amusé à écrire dans de l'ADN des sonnets de Shakespeare, un passage de l'enregistrement du discours « I have a dream » de Martin Luther King et l'article scientifique célèbre de James Watson et Francis Crick sur la structure de l'ADN.

Le coût estimé de ces enregistrements est aujourd'hui de l'ordre de 10 000 euros par mégaoctet (1 million de caractères) pour l'écriture, et de 200 euros par mégaoctet pour la lecture. C'est assez cher, mais cela baisse rapidement avec les progrès des méthodes de synthèse et de séquençage. Nous avons donc de bonnes raisons de croire qu'à court terme, dans une dizaine d'années, nous disposerons de méthodes de stockage de ce type, compétitives et adaptées aux longues durées, qui exploiteront l'ADN ou peut-être d'autres molécules biologiques. L'équipe de Zurich imagine ainsi bientôt stocker le contenu de l'encyclopédie Wikipédia dans une archive à base d'ADN.

Des codes-barres moléculaires

On envisage aussi d'utiliser ces méthodes pour réaliser des codes-barres moléculaires marquant des objets ou des solutions, ce qui aiderait à protéger de la contrefaçon des parfums ou des molécules pharmaceutiques. On mélangerait de très petites quantités de molécules de synthèse où une sorte de signature aurait été inscrite. Un appareillage de lecture donnerait rapidement l'information qu'un produit est authentique ou, à l'inverse, qu'il est contrefait. On pourrait également tracer ainsi la provenance d'armes à feu et pister les trafiquants.

Tout récemment, en juillet 2016, James Bornholt, de l'université de Washington, et des chercheurs de Microsoft ont présenté une méthode d'archivage massif fondée sur l'ADN permettant des accès en lecture relativement rapide pour l'information qui serait ainsi stockée. Une méthode qui atteint un record de 200 mégaoctets d'informations enregistrées.

Les calculs de la capacité de ces technologies montrent qu'elles ont un potentiel supérieur, à moyen terme, à celui des technologies à base de silicium.

Si les difficultés du calcul avec l'ADN se confirment en même temps que les succès du stockage moléculaire de données, nous serons en mesure d'affirmer que l'électronique est le meilleur moyen de faire rapidement du calcul (qu'on pourra encore pousser plus loin avec des méthodes quantiques), mais que pour le stockage, ce qu'a mis au point la nature est plus économique, plus robuste et capable de stocker plus massivement l'information que nous souhaitons préserver pour de longues périodes.

Des cellules vivantes programmées

Mentionnons un autre type d'exploit dans le domaine du calcul biologique. Il vient d'être réalisé par une équipe de chercheurs réunis autour de Nathaniel Roquet, du MIT, et ouvre la perspective de programmer à volonté des cellules vivantes (et non plus de soupes moléculaires en bouteille) qui deviendraient des automates à notre service et seraient aussi complexes qu'on le souhaite.

Les cellules vivantes exécutent des sortes de programmes qui, en fonction de leur identité et des signaux chimiques qui leur parviennent, les font réagir de manière spécifique, par exemple en envoyant d'autres signaux chimiques dans la cellule ou à l'extérieur. Jusqu'à présent, on ne savait ni simuler ni reproduire ces réseaux de régulation avec des programmes génétiques artificiels qui, à l'opposé de ceux envisagés par Leonard Adleman, n'ont pas pour objectif un calcul combinatoire ou abstrait, mais une réaction particulière programmée, qui produit une interaction déterminée au sein de la cellule ou avec d'autres cellules. La compréhension et la mise au point de circuits artificiels internes aux cellules qui reconnaissent des entrées successives arrivant dans un ordre fixé et qui répondent spécifiquement semblaient difficilement atteignables. On sait maintenant que c'est possible.

La méthode conçue a pour effet de créer biologiquement des machines à états : la cellule programmée se met dans un état qui est déterminé par la série d'informations qu'elle reçoit et dont l'ordre importe. Ces machines à états biologiques s'appuient sur

des enzymes nommées recombinaisons. Quand une recombinaison est activée par un signal chimique, elle enlève ou inverse un segment particulier d'ADN situé entre deux portions de séquences qui délimitent le segment et qui sont spécifiques à la recombinaison. La séquence (retirée ou inversée) peut contenir des sites de reconnaissance pour d'autres recombinaisons qui pourront alors répondre à d'autres signaux chimiques.

Après avoir créé de tels circuits qui enregistrent des événements, les chercheurs ont incorporé des gènes dans le réseau des séquences manipulées. Dans ces circuits, l'ADN se réorganise, ce qui non seulement enregistre des informations, mais détermine l'activité des gènes, ces derniers étant activés ou désactivés selon les signaux reçus. Les chercheurs ont testé cette approche avec des gènes qui codent pour différentes protéines fluorescentes. Les cellules qui contiennent le circuit artificiel auront par exemple la propriété de présenter une fluorescence rouge si elles reçoivent le signal A puis le signal B, alors que si elles reçoivent le signal B puis le signal A, elles présenteront une fluorescence bleue.

Un ordinateur-univers en perpétuelle reprogrammation

La biologie est une science où le calcul et l'information jouent des rôles centraux. Aujourd'hui, on progresse encore et au lieu d'être simplement les spectateurs de ces calculs et mémorisations d'informations que la vie opère pour fonctionner et se perpétuer, nous apprenons à utiliser et à programmer nous-mêmes des mécanismes moléculaires de traitement de l'information pour en faire ce qui nous arrange.

C'est merveilleux, et cela donne de nouveaux arguments à ceux qui voient dans l'Univers un grand ordinateur : les organismes vivants constitueraient des lieux où s'effectueraient des calculs particulièrement intenses, et leur évolution s'interpréterait comme un perfectionnement de l'ordinateur-univers, qui se reprogrammerait constamment.

L'AUTEUR



J.-P. DELAHAYE
est professeur
émérite
à l'université
de Lille
et chercheur

au Centre de recherche
en informatique, signal
et automatique de Lille (CRISTAL).

BIBLIOGRAPHIE

N. Roquet *et al.*, **Synthetic recombinaison-based state machines in living cells**, *Science*, vol. 353, 22 juillet 2016.

A. Exrance, **How DNA could store all the world's data**, *Nature*, vol. 537, pp. 22-24, 2016.

J. Bornholt *et al.*, **A DNA-based archival storage system**, *Proceedings of the 21st International Conference on Architectural Support for Programming Languages and Operating Systems*, ACM, 2016.

V. Zhirnov *et al.*, **Nucleic acid memory**, *Nature Materials*, vol. 15, pp. 366-370, 2016.

V. Zhirnov et Ralph Cavin III, **Microsystems for Bioelectronics**, William Andrew, 2^e éd., 2016.

Y. Benenson *et al.*, **An autonomous molecular computer for logical control of gene expression**, *Nature*, vol. 429, pp. 423-429, 2004.



Retrouvez la rubrique
Logique & calcul sur
www.pourlascience.fr

SCIENCE & FICTION

Glapum'tien, de quelle espèce es-tu ?

Pour répondre à cette question, les naturalistes du XIX^e siècle auraient chassé cet être imaginaire de la bande dessinée pour en prélever un spécimen et le décrire. Mais il est bien plus enrichissant de l'observer dans son environnement...

Jean-Sébastien STEYER et Roland LEHOUCQ. Illustration : Marc BOULAY

Dans les œuvres de fiction, les extraterrestres intelligents sont souvent des humanoïdes qui nous ressemblent. Cet anthropocentrisme persistant passe sous silence les nombreuses autres espèces bien réelles – cétagés, insectes, céphalopodes ou autres – qui présentent elles aussi des formes d'intelligence. Mais heureusement, il existe parfois des exceptions qui confirment la règle, comme c'est le cas des Glapum'tiens.

Les Glapum'tiens sont des extraterrestres de l'univers de « Valérian et Laureline », série de bandes dessinées de Pierre Christin et Jean-Claude Mézières dont la première publication date de 1967 (et dont l'album *L'Ambassadeur des ombres* est en cours d'adaptation au cinéma par Luc Besson). Cette saga narre, dans un futur lointain, les aventures trépidantes de deux agents spatiotemporels censés défendre les intérêts de l'« Empire Terrien » dans l'univers.

Dans l'album *Les Spectres d'Inverloch* (1984), Valérian est envoyé sur la planète Glapum't, où il a pour mission de capturer un Glapum'tien. Notons que le qualificatif de « Glapum'tien » suggère que, à l'image du mot « Terrien » que nous utilisons pour nous qualifier, l'espèce est dominante sur la planète en question. Sur Glapum't, magnifique planète bleue aux paysages côtiers et aux îles inondées par de grandes marées, Valérian fait preuve d'ingéniosité pour capturer à contrecœur un sympathique Glapum'tien : « Je ne suis ni zoologue ni collectionneur [...] et j'ai horreur de chasser



AVEC SA HAUTE STATURE ET SA PEAU BLEUE, le Glapum'tien, l'un des personnages de la bande dessinée *Les Spectres d'Inverloch*, ne passe pas inaperçu sur la couverture de l'album.

l'autochtone pour rapporter des spécimens, même en état de marche. » Cette phrase évoque très justement les naturalistes des siècles derniers qui tuaient les animaux afin de les prélever, au lieu de les observer ou de les étudier directement dans leur environnement naturel. Dans l'album suivant, intitulé *Les Foudres d'Hypsis* (1985), le Glapum'tien en question devient l'ami de Laureline et Valérian, qu'ils prénomment Ralph.

Ralph est un être sensible et intelligent, haut de plus de deux mètres, à grande tête plate en forme de dôme et aux petits yeux malins. Il est doté de quatre tentacules souples et élastiques et d'une longue queue puissante et robuste. Comme les céphalopodes, il n'a pas de cou ; ses membres sont directement situés sous la tête. Mais comme les tétrapodes (vertébrés munis de pattes), Ralph a des doigts. S'agit-il donc d'un mollusque (céphalopode) ou d'un vertébré (tétrapode) ? La réponse est donnée dans *Les Habitants du Ciel – Atlas cosmique de Valérian et Laureline* (1991), sorte d'encyclopédie illustrée, publiée par les auteurs et qui décrit avec grande précision les principales formes de vie rencontrées dans la saga. Verdict : les Glapum'tiens sont des mammifères. Certes, mais lesquels ? Observons Ralph dans son environnement pour en savoir plus.

Ralph est un excellent nageur qui se déplace aussi sur la terre ferme à l'aide de ses membres chiridiens (munis de doigts). Il est alors soit bipède, soit quadrupède. À l'arrêt, la souplesse de ses tentacules ainsi