

UN CHEVAL AU GALOP DANS UNE BACTÉRIE

Dans la plupart des recherches visant à coder des informations (images, textes, sons ou films) sur de l'ADN, les données sont archivées sur des brins d'ADN de synthèse. L'ADN est ensuite encapsulé dans des nanobilles de verre ou séché à l'air, ce qui permet de le conserver à très long terme. En 2017, Seth Shipman, alors à l'université Harvard, aux États-Unis, et ses

collègues ont testé une autre approche : stocker des images animées dans l'ADN de bactéries vivantes.

Leur image de départ était une photographie de main de 56 x 56 pixels de quatre couleurs de pixels différentes (quatre niveaux de gris, voir ci-dessous). En langage binaire, chaque couleur de pixel peut être codée par un duo de 0 et/ou de 1. Pour encoder l'image dans de l'ADN, les chercheurs ont associé chaque couleur de pixel à un nucléotide. Puis ils ont scindé les 3 136 pixels de l'image en 112 ensembles de 28 pixels, disséminés dans toute l'image. Chaque ensemble de pixels était ainsi codé par une succession de 28 nucléotides, à laquelle était ajouté un code-barres de 4 nucléotides permettant de reconnaître la zone et de la situer dans l'image. Les biologistes ont ensuite synthétisé les 112 fragments d'ADN obtenus en de nombreux exemplaires.

Pour intégrer ces fragments au sein du génome de bactéries *Escherichia coli*, ils ont détourné l'arsenal CRISPR de ces organismes – une sorte de système immunitaire qui enregistre, dans l'ADN bactérien, une trace des infections virales subies. Par électroporation, une technique qui consiste à appliquer un champ électrique sur les membranes cellulaires afin de les perméabiliser, les chercheurs ont introduit les fragments d'ADN dans les bactéries. Grâce à la machinerie CRISPR, dont les biologistes avaient forcé l'activation, les fragments d'ADN ont été copiés au sein du génome des bactéries. Le jour suivant, les chercheurs ont séquencé l'ADN de la population bactérienne,

puis, grâce à un programme, ils ont converti les séquences nucléotidiques en langage binaire et ont ouvert l'image : 88 % des ensembles de pixels étaient correctement restitués. Les chercheurs ont ensuite réalisé la même expérience avec une image de main de 33 x 33 pixels comportant cette fois 21 couleurs de pixels. Les couleurs de pixels étaient codées non plus sur un nucléotide, mais sur un triplet. Avec cette méthode, 96 % des ensembles de pixels étaient fidèlement restitués 2 heures 40 après l'électroporation. Ensuite, le pourcentage de reconstitution de l'image diminuait légèrement dans les 24 heures, puis restait stable pendant les 6 jours suivants, au cours desquels 48 générations de bactéries se succèdent.

Les scientifiques ne se sont pas arrêtés à ce premier succès : ils ont souhaité enregistrer une séquence de cinq images issues du premier film d'Eadweard Muybridge, mettant en scène un cheval au galop et sa cavalière. Pour cela, ils ont exploité une caractéristique de l'arsenal CRISPR : lorsque les bactéries copient de l'ADN étranger au sein de leur propre ADN, l'ADN est copié dans un ordre précis selon la chronologie d'apparition. Ainsi, sur le même principe que pour l'image de main, ils ont d'abord introduit les brins d'ADN constituant la première image. Puis, le lendemain, ils ont refait de même avec ceux de la deuxième image, et ainsi de suite pendant cinq jours, réussissant de cette façon à coder une image animée dans de l'ADN bactérien.

En réalisant ces travaux, les chercheurs imaginaient-ils un futur où nos films favoris seraient stockés dans le génome de populations bactériennes ? Non, car les brins d'ADN de synthèse sont plus appropriés à cette fin. « Mais ce qui est vraiment intéressant dans cette étude, commente Christophe Antoniewski, directeur de la plateforme de bio-informatique ARTbio de l'institut de biologie Paris-Seine, c'est la perspective d'utiliser l'ADN comme enregistreur moléculaire *in vivo* de l'activité des cellules, afin d'étudier, par exemple, leur développement, ou leur dérèglement dans les cancers. »

HÉLÈNE GÉLOT
Journaliste scientifique

AVANT / APRÈS

Image originale



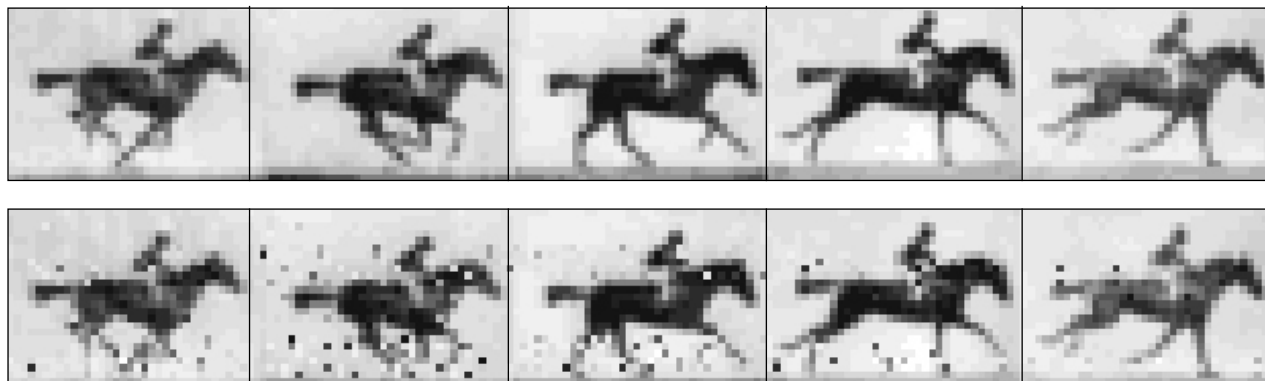
00 = C
01 = T
10 = A
11 = G

C → ■
T → ■
A → ■
G → ■

Image reconstituée



Les points blancs correspondent à des informations non restituées.



La séquence originale (en haut) et celle reconstituée à partir des bactéries (en bas) ne diffèrent que de quelques pixels.

➤ résistance, elle se divisait rapidement. Par conséquent, au fil du temps, la quantité de codes-barres A, par exemple, augmentait si la cellule de type A devenait résistante au médicament ou diminuait si elle était neutralisée. En séquençant tous les codes-barres des cellules survivantes au cours du temps, le laboratoire a quantifié les réponses de chaque type cellulaire au médicament.

Quelques mois plus tard, ce fut le tour de l'équipe de Monte Winslow, de l'université de Stanford, de mener des recherches à l'aide de codes-barres ADN, cette fois pour trouver des inhibiteurs de métastases dans le cancer du pancréas. Les biologistes ont d'abord produit des lignées de cellules cancéreuses à partir de

cardiovasculaires. Dans les exemples cités plus haut, remplacez le mot « cancer » par une autre maladie et le mot « résistance » par n'importe quelle réponse à un médicament et vous percevrez l'étendue de leur potentiel.

ENCODER UN FILM DANS L'ADN D'UNE BACTÉRIE

Ce n'est pas tout. Les approches utilisant des codes-barres ADN reposent sur la lecture de séquences connues d'ADN. Mais cela suppose de préalablement « écrire » ces séquences. Or, jusqu'à récemment, la synthèse de séquences d'ADN restait compliquée dès qu'elles étaient un peu trop longues. Il était donc inenvisageable de convertir d'autres formes d'informations – images, films, états biologiques... – en séquences de nucléotides que l'on pourrait stocker et lire plus tard. Mais de nouvelles techniques d'écriture offrent désormais cette possibilité. Elles reposent sur des outils d'édition du génome dérivés du système CRISPR.

Ce système est un moyen de défense des bactéries contre les attaques de virus. Lorsqu'un virus infecte une bactérie, il se fixe à sa surface et introduit son ADN (ou ARN) à l'intérieur. Mais la bactérie garde alors une trace de l'attaque en insérant des extraits de cet ADN (ou ARN) étranger dans une région de son propre génome – une sorte de bibliothèque nommée CRISPR. En cas de nouvelle attaque du même virus, la bactérie le reconnaît alors grâce à sa bibliothèque CRISPR et détruit son matériel génétique avant qu'il n'ait le temps de se répliquer.

En 2017, dans le laboratoire du généticien George Church à l'université Harvard, Seth Shipman, aujourd'hui à l'université de Californie à San Francisco, a exploité ce mécanisme de défense pour enregistrer une image de main dans le génome d'une population de bactéries *Escherichia coli* (voir l'encadré page ci-contre). Dans de précédents travaux, l'équipe avait montré qu'il suffisait d'induire l'expression de deux protéines de ces bactéries – Cas1 et Cas2 –, associées au fonctionnement du système de défense CRISPR, pour forcer les cellules à insérer de courtes séquences d'ADN de synthèse dans leur bibliothèque CRISPR. Seth Shipman et ses collègues ont donc recommencé, cette fois avec des fragments d'ADN qui encodent les pixels de l'image.

En séquençant l'ADN d'*E. coli*, ils ont retrouvé l'image initiale avec une précision de plus de 90%. Ils ont ensuite reproduit l'expérience, mais avec une difficulté supplémentaire : ils ont ajouté les séquences d'ADN à des temps différents et mis au point une méthode pour enregistrer la position de ces séquences les unes par rapport aux autres. En retrouvant l'ordre d'ajout des séquences au sein du ➤

Ils ont encodé une série d'images dans l'ADN et l'ont restituée

tumeurs primaires et de métastases provenant de modèles murins de la maladie. Parmi elles, ils n'ont gardé que celle qui, injectée chez des souris, formait le plus facilement des métastases. Puis, à l'aide d'un virus, ils ont introduit une étiquette d'ADN dans 96 lignées cellulaires issues de la lignée sélectionnée, avant de disposer chacune d'elles dans un puits distinct. Ils ont ensuite ajouté un inhibiteur potentiel dans chaque puits. Ainsi, chaque molécule était associée à un code-barres différent. Les scientifiques ont alors immédiatement injecté les cellules dans les vaisseaux sanguins de souris et suivi, grâce aux codes-barres, quelles cellules se sont développées dans les poumons. En identifiant les codes-barres présents ou absents des poumons, ils ont ainsi mis en évidence plusieurs molécules qui favorisaient ou empêchaient les métastases. Cette méthode leur a permis de tester plus de 700 inhibiteurs d'enzymes sur 20 fois moins de souris.

Dans ces approches, l'ADN est à la fois un enregistreur moléculaire – une étiquette qui permet de suivre des expériences menées simultanément et les repérer les unes par rapport aux autres – et un générateur de données, puisqu'il offre la possibilité de réaliser de nombreuses expériences en même temps. Les implications sont stupéfiantes : les mêmes techniques peuvent être appliquées contre les maladies auto-immunes, neurologiques ou