

> 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

➤ génome de la bactérie, ils sont parvenus à encoder une série d'images dans l'ADN et à la restituer. Les chercheurs ont ainsi enregistré un extrait de la première séquence d'animation créée en 1878 par le photographe britannique Eadweard Muybridge, qui met en scène un cheval au galop.

## ENREGISTRER EN TEMPS RÉEL L'EXPRESSION DES GÈNES

Plus récemment encore, en 2018, des chercheurs du laboratoire de Randall Platt, à l'École polytechnique fédérale de Zurich, en Suisse, ont mené ces approches encore plus loin en ciblant non plus l'ADN, mais l'ARN messager. Cette molécule est un cousin moléculaire clé de l'ADN, intermédiaire indispensable entre les gènes portés par l'ADN et la synthèse des protéines qu'ils codent. Au lieu de stocker des images dans des séquences d'ADN de synthèse, les chercheurs ont utilisé le système CRISPR d'une bactérie pour enregistrer en temps réel les ARN messagers qu'elle produisait, et donc l'expression de ses gènes. L'ensemble de tous les ARN messagers présents dans une cellule à un instant donné permet de savoir quelles protéines sont produites, et ainsi de mieux comprendre le fonctionnement de la cellule étudiée.

Afin d'enregistrer dans le génome d'une cellule les séquences des ARN messagers qu'elle a produits à divers moments, l'équipe de Randall Platt a tout d'abord testé différentes protéines Cas (associées au système CRISPR) issues de nombreuses souches bactériennes distinctes. Elle a ainsi découvert que les protéines Cas1 et Cas2 de la bactérie *Fusicatenibacter saccharivorans* étaient capables de convertir l'ARN messager en ADN et de l'encoder dans sa bibliothèque CRISPR. Les chercheurs ont utilisé ce procédé afin d'enregistrer le comportement d'une cellule à la suite de différents événements: un stress oxydatif, des conditions acides et l'exposition à un herbicide.

Ces résultats sont particulièrement intéressants: ils fournissent la preuve qu'il est possible d'enregistrer dans le génome d'une cellule, à un instant donné, les gènes qu'elle exprime, en vue d'une analyse ultérieure. Le laboratoire de Randall Platt continue d'améliorer cette technique. L'enregistrement en temps réel de l'expression des gènes pourrait s'étendre à d'autres domaines et permettre, par exemple, de suivre les événements qui rendent cancéreuse une cellule ou encore d'observer comment une cellule répond à une infection au cours du temps.

Alors que l'ADN est utilisé pour produire, suivre et stocker de l'information dans des domaines de plus en plus nombreux, on peut se demander s'il a la capacité de rivaliser avec les dispositifs électroniques usuels de stockage et de prendre en charge l'ensemble des

# Conservé au sec et au frais, l'ADN résiste des dizaines de milliers d'années

données générées par les humains. Pour l'instant, la réponse est non: disques durs et cartes mémoires sont bien plus adaptés pour conserver des informations que le plus avancé des systèmes génétiques.

En même temps, comme toute technologie, les appareils électroniques conventionnels ont leurs limites. Ils prennent de la place et nécessitent des conditions environnementales particulières. Et même les dispositifs les plus robustes tiennent rarement plus de quelques décennies. À cause de ces obstacles, il deviendra de plus en plus difficile de conserver toutes les données que nous produisons.

Au contraire, l'ADN résiste des dizaines de milliers d'années s'il est conservé dans des environnements frais et secs. Il est aujourd'hui couramment stocké à  $-20^{\circ}\text{C}$ , voire à  $-80^{\circ}\text{C}$  dans les laboratoires, mais pourrait aussi être conservé sous des chaleurs extrêmes que les appareils électroniques ne supportent pas. En 2015 en effet, Robert Grass et Wendelin Stark, tous deux à l'École polytechnique fédérale de Zurich, ont montré que des molécules d'ADN encapsulées dans des billes de verre résistaient à une température de  $70^{\circ}\text{C}$  pendant une semaine sans qu'aucune erreur n'apparaisse. Et si les disques durs permettent de stocker 19 gigaoctets par centimètre carré, des estimations récentes suggèrent que l'ensemble des informations générées dans le monde tiendraient, en théorie, dans moins de 1 kilogramme d'ADN.

## UNE RECHERCHE À ENCADRER

Il reste du chemin à parcourir avant que le stockage sur ADN devienne courant. Le premier obstacle réside dans l'extraction des informations: si elle est presque instantanée à partir d'un disque dur, récupérer des données encodées dans l'ADN requiert du séquençage, ce qui peut prendre de quelques minutes à une journée entière. Et même si les séquenceurs ont connu des avancées considérables ces dernières années, ils restent imposants et chers par rapport aux disques durs.