

## L'ADN de synthèse pour stocker des données

### Description

L'acide désoxyribonucléique, ou ADN, est une macromolécule biologique présente dans toutes les cellules du vivant qui contient l'information génétique, appelée génome, permettant le développement, le fonctionnement et la reproduction des organismes. Partant du constat que l'ADN est naturellement utilisé pour stocker de l'information, des équipes de chercheurs travaillent depuis plusieurs années afin de l'utiliser comme support destiné au stockage d'informations numériques.

L'idée de stocker des informations numériques dans l'ADN a été proposée par le physicien Richard P. Feynman dès 1959. Il aura cependant fallu attendre 2013 pour que des chercheurs du LEBM-Institut européen de bio-informatique montrent qu'il était possible d'enregistrer des données dans l'ADN et de les récupérer. À l'époque, le coût par mégaoctet était estimé à 12 400 dollars pour l'encodage des données et 220 dollars pour leur récupération. La société française de biotechnologie DNA Script a annoncé avoir reçu, le 21 janvier 2020, un financement de l'Intelligence Advanced Research Projects Activity (IARPA), l'agence chargée des projets de recherche avancée en renseignement pour le gouvernement américain, avec pour objectif de *« développer une machine capable d'encoder dans une molécule un téraoctet de données à l'équivalent de 250 films en vingt-quatre heures, pour un coût maximal de 1 000 dollars (902 euros) »*.

En collaboration avec l'Institut Pasteur et l'École supérieure de physique et de chimie industrielles (ESPCI), la société, créée en 2014 par Thomas Ybert, Sylvain Gariel et Xavier Godron, a déjà levé 27 millions de dollars auprès d'investisseurs publics et privés et développe une technologie de synthèse d'ADN par voie enzymatique. Le procédé consiste à convertir des données numériques, binaires, au format quaternaire des quatre nucléotides qui composent la molécule d'ADN : G, C et A, T. L'écriture de données est appelée *« la synthèse »* et sa lecture *« le séquençage »*. *« La synthèse chimique de l'ADN a été optimisée au maximum ces dernières décennies, tout a été creusé en recherche fondamentale et elle ne s'améliorera plus »*, déclare Jean-François Lutz, directeur de recherche au laboratoire de chimie macromoléculaire de précision au CNRS. C'est pourquoi la société française innove en utilisant une enzyme pour réaliser l'écriture des données dans l'ADN.

En signant avec l'IARPA, DNA Script va collaborer pendant quatre ans au sein d'un consortium doté d'une première enveloppe de 23 millions de dollars avec des chercheurs du Massachusetts Institute of Technology (MIT), Harvard, et Illumina le géant américain du

s'acquiesce génétiquement. L'IARPA a constitué un second consortium, réunissant Microsoft et la société de biotechnologie californienne Twist Bioscience, qui dispose d'un budget de 25 millions de dollars, portant ainsi le financement total à 48 millions de dollars.

L'intérêt de stocker des données dans l'ADN est double. D'une part, la densité de stockage est bien plus grande qu'avec n'importe quel support traditionnel et, d'autre part, sa durée de vie est nettement supérieure. Actuellement, la capacité de stockage d'un centre de données numériques est de l'ordre d'un exabyte, soit un milliard de gigabytes. « Par comparaison l'ADN d'un corps humain, environ 150 grammes, stocke 150 zétabytes à 150 000 exabytes de données » 150 data centers, et n'a pas besoin d'énergie. Il est admis qu'en 2040, la quantité mondiale de données générées par le Big Data depuis l'origine par l'ensemble de la planète sera de 3 000 zétabytes soit l'équivalent de 3 kg d'ADN », rapporte Nicolas Viard pour Innocherche. En outre, stocker des données dans l'ADN pourrait permettre de résoudre l'obsolescence des supports traditionnels, notamment les disques durs dont la bande magnétique ne dure pas plus de dix ans. L'information codée dans l'ADN a une durée de vie de plusieurs milliers, voire de centaines de milliers d'années.

Des applications tournées vers le grand public pourraient même potentiellement voir le jour d'ici à une vingtaine d'années. « Cela permettrait aux grands acteurs du cloud et de la tech de conserver les données à proximité de leurs clients, avec un coût bien moindre », explique David Markowitz, chargé du programme de stockage moléculaire des données (MIST) à l'IARPA.

Sources :

- « Synthetic double-helix faithfully stores Shakespeare's sonnets », Ed Yong, Nature, January 23, 2013.
- « L'ADN, une piste sérieuse pour stocker nos milliards de données », Stéphane Desmichelle, sciencesetavenir.fr, 30 juin 2017.
- « L'impression d'ADN, c'est facile à San Francisco », Nicolas Viard, Innocherche, 7 février 2018.
- Thomas Ybert (DNA Script) : « Grâce à nos enzymes, plus besoin d'être un chimiste pour fabriquer de l'ADN », Cécile Deluzarche, futura-sciences.com, 5 février 2019.
- « Le gouvernement américain investit dans le stockage de données dans l'ADN », Chloé Hecketsweiler, Le Monde, 22 janvier 2020.

## Categorie

### 1. Techniques

**date création**

7 mai 2020

**Auteur**

jacquesandrefines