

COMMUNIQUÉ DE PRESSE

DNA Drive, la Révolution de l'ADN

Première mondiale : dépôt d'archives numériques encodées sur ADN aux Archives Nationales

Paris, le 23 novembre 2021

DNA Drive est la technologie de stockage d'information numérique sur ADN développée et brevetée par Stéphane Lemaire et Pierre Crozet, respectivement directeur de recherche CNRS et maître de conférences à Sorbonne Université. L'encodage sur ADN de deux textes à forte valeur symbolique et historique, la Déclaration des droits de l'homme et du citoyen de 1789 et la Déclaration des droits de la femme et de la citoyenne, rédigée par Olympe de Gouges en 1791, a permis d'obtenir une preuve de concept de l'efficacité et de la fonctionnalité du DNA Drive. À l'occasion d'une conférence de presse exceptionnelle, organisée mardi 23 novembre au musée des Archives nationales, l'équipe de recherche a présenté son projet « La Révolution de l'ADN » et les Archives nationales ont officiellement enregistré le dépôt de ces archives numériques encodées sur ADN, une première mondiale pour une institution publique.

Le stockage et l'archivage des données numériques sont des enjeux stratégiques critiques pour l'économie, la pérennité et la sécurité de nos sociétés. Cependant, ils se heurtent aujourd'hui à la concomitance de trois facteurs de limitation majeurs : la faible durée de vie des supports de stockage ; la quantité d'énergie gigantesque requise pour ce stockage, induisant un coût économique et un impact environnemental considérables ; l'expansion vertigineuse des données qui entraîne une demande de stockage largement supérieure à nos capacités.

En particulier, les supports actuels de stockage ne sont plus suffisants pour archiver nos données : depuis 2010, la demande est supérieure à l'offre de stockage. A terme, les systèmes classiques de stockage ne sont pas soutenables¹. Ainsi, la transformation numérique de nos sociétés (intelligence artificielle, *Big Data*, objets connectés, voitures autonomes, informatique quantique...) nécessitera une évolution technologique majeure de nos systèmes de stockage de données.

DNA Drive : une stratégie bio-inspirée de stockage sur ADN

Afin de lever ces limitations et de développer des systèmes de stockage de données plus performants, l'équipe de Stéphane Lemaire, directeur de recherche CNRS au Laboratoire de biologie computationnelle et quantitative (CNRS/Sorbonne Université) et Pierre Crozet, maître de conférences à Sorbonne Université a cherché à tirer profit du potentiel offert par le vivant. Utilisant la biologie, ils ont exploité une forme de stockage de données qui n'a pas été inventée par l'Homme, mais qui se perfectionne depuis près de 4 milliards d'années : l'ADN, le support de l'information génétique.

Le stockage d'information numérique sur ADN est une technologie émergente pour laquelle toutes les approches publiées sont uniquement basées sur des méthodes chimiques, physiques, mathématiques et informatiques. En effet, la piste biologique n'avait pas encore été explorée. Des technologies très précises de manipulation de l'ADN existent dans le vivant (lecture, copie, édition, correction d'erreurs, accès direct, amplification du signal, etc.) et peuvent être domestiquées et adaptées *via* des approches de biologie synthétique pour lever les limitations actuelles du stockage sur ADN. C'est cette stratégie qui a permis à l'équipe de développer le DNA Drive, une nouvelle technologie de stockage sur ADN bio-inspirée.

Le stockage d'information numérique sur ADN constitue une solution :

¹ « Si nous sommes aujourd'hui capables de stocker environ 30% des informations que nous générons, dans seulement 10 ou 20 ans, nous ne pourrions plus en stocker qu'environ 3% », déclarait Dr. Karin Strauss de Microsoft Research, en octobre 2018.

- **durable** : la stabilité de l'ADN se compte en dizaines, voire en centaines de milliers d'années, ce qui est incomparable avec celle des supports de stockage actuels. Un génome complet a par exemple été obtenu à partir d'une défense de mammoth âgée de plus d'un million d'années.
- **non énergivore** : l'ADN est stable à température ambiante sans aucun apport d'énergie s'il est conservé dans des conditions adéquates (sans eau, ni air, ni lumière).
- **extrêmement compacte** : avec une densité maximale de 450 millions de Teraoctets par gramme d'ADN (0,45 Zettaoctets/g), l'intégralité des données mondiales pourrait tenir dans 100 g d'ADN, soit le volume d'une tablette de chocolat.

Le DNA Drive présente l'avantage d'être compatible avec tout type d'information numérique et permet d'encoder tout type de systèmes de fichiers avec des tables d'allocations, des répertoires, des métadonnées, etc. De plus, il offre un accès direct (*random access*) aux données car les secteurs ADN sont bordés de blocs d'index, permettant de relire spécifiquement chaque secteur sans avoir besoin de relire l'intégralité des données. **L'ensemble de ces propriétés permettent de définir un nouveau standard pour le stockage d'information numérique sur ADN. Un brevet protégeant la technologie DNA Drive a été déposé à l'Office européen des brevets et est en cours d'extension mondiale².**

« La Révolution de l'ADN » : un projet pluridisciplinaire

« Tout a commencé par un article publié en 2018 dans le journal de l'association étudiante Alma Mater. Dans ce journal, il y avait un article sur la technologie de stockage sur ADN. En le lisant, j'ai indiqué aux étudiants de cette association que mon équipe maîtrisait les technologies de manipulation de l'ADN et pourrait donc encoder de l'information sur ADN. Ils m'ont alors mis au défi de le prouver. Je leur ai répondu que j'étais prêt à le faire mais uniquement pour encoder des données importantes et ayant du sens. Ils ont donc réfléchi et m'ont proposé d'encoder la Déclaration des droits de l'homme et du citoyen de 1789. Je leur ai répondu que c'était une excellente idée mais qu'il fallait aussi encoder la Déclaration des droits de la femme et de la citoyenne de 1791 », Stéphane Lemaire, directeur de recherche CNRS au Laboratoire de biologie computationnelle et quantitative (CNRS/Sorbonne Université).

Ainsi est né le projet « La Révolution de l'ADN ». Dans un premier temps, il a permis de rassembler une petite équipe de biologistes et d'informaticiens qui a conçu et breveté la technologie DNA Drive avec des financements d'amorçage de Sorbonne Université et du CNRS puis de la SATT Lutech.

L'équipe a ensuite poursuivi un double objectif :

- obtenir une preuve de concept de l'efficacité de cette technologie de stockage, le DNA Drive ;
- encoder sur ADN des textes fondateurs français à forte valeur symbolique et historique.

Profondément pluridisciplinaire, à l'image de Sorbonne Université et du CNRS, le projet « La Révolution de l'ADN » a impliqué des biologistes, informaticiens, historiens, philosophes et archivistes. La Déclaration des droits de l'homme et du citoyen étant inscrite au programme Mémoire du Monde de l'UNESCO, il a semblé essentiel aux chercheuses et chercheurs de l'accompagner de la Déclaration des droits de la femme et de la citoyenne, texte non institutionnel rédigé par Olympe de Gouges et publié en 1791. L'association de ces deux déclarations véhiculant des valeurs humanistes auxquelles l'équipe est attachée.

...et la preuve de concept du DNA Drive !

Dans le cadre de ce projet, deux DNA Drive différents ont été créés : un pour stocker la Déclaration des droits de l'homme et du citoyen de 1789 et l'autre pour la Déclaration des droits de la femme et de la citoyenne de 1791. Chaque DNA Drive a été assemblé à partir de fragments d'ADN synthétisés par Twist Bioscience et stocké dans des capsules métalliques DNAShell© par Imagene³. Chaque capsule contient plus de 100 milliards de copies du fichier sous forme d'ADN. La lecture de l'information peut être effectuée avec des séquenceurs d'ADN, similaires à ceux utilisés en biologie et en médecine pour séquencer les génomes des êtres vivants. À plusieurs reprises, l'équipe a pu ouvrir des capsules, séquencer l'ADN contenu et vérifier qu'il était possible de récupérer le fichier avec une fidélité de 100%. Ces deux textes historiques sont donc désormais conservés pour des milliers d'années.

² EP 193062478, 01/10/2019; PCT EP2020/077497, 01/10/2020. CNRS, Sorbonne Université.

³ Technologie d'Imagene permettant un stockage à très long terme (> 50 000 ans) de l'ADN.

Ainsi, « La Révolution de l'ADN » a permis d'obtenir une preuve de concept de la fonctionnalité de la technologie DNA Drive. Les capsules contenant les deux DNA Drive ont été exposées lors de la conférence de presse du 23 novembre 2021, et officiellement été enregistrées par les Archives nationales comme premières archives sous forme ADN. Les capsules contenant les deux DNA Drive ont été exposées lors de la conférence de presse du 23 novembre 2021 et ont été officiellement accueillies par les Archives nationales pour garder mémoire de cette expérimentation de stockage prometteuse.

Elles sont désormais conservées dans l'Armoire de Fer qui contient les plus précieux documents des Archives nationales.

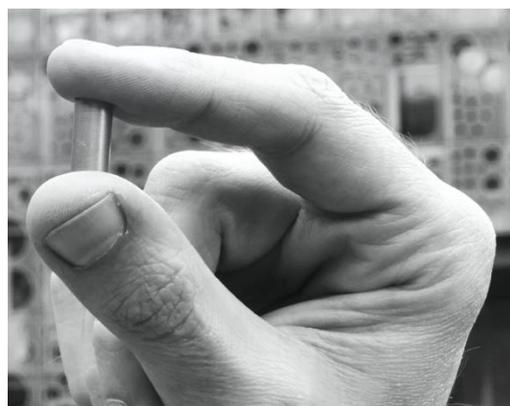
Un projet partenarial

Le projet « La Révolution de l'ADN » a été mené dans le cadre d'un partenariat étroit entre l'équipe de recherche et les Archives nationales, qui conservent l'original de la Déclaration des droits de l'homme et du citoyen. L'objectif était de permettre aux Archives nationales d'enregistrer officiellement ces deux textes encodés sur ADN, **une première mondiale pour une institution publique.**

Le projet a été également mené en partenariat avec Twist Bioscience, entreprise américaine spécialiste de la synthèse d'ADN, et Imagene, entreprise française issue de l'Institut de chimie de la matière condensée (CNRS/Université de Bordeaux/Bordeaux INP) et spécialiste de la conservation à long terme de l'ADN.



Capsules contenant les deux textes encodés sur ADN –
© Stéphane Lemaire / CNRS – Sorbonne Université



Chaque capsule peut contenir une quantité d'ADN
correspondant à 5000 To de données numériques
© Philippe Tran

Contacts presse

Marion Valzy - marion.valzy@sorbonne-universite.fr
01 44 27 37 13 - 06 14 02 20 51

Claire de Thoisy-Méchin - claire.de_thoisy-mechin@sorbonne-universite.fr
01 44 27 23 34 - 06 74 03 40 19

Catherine Vergriete - catherine.vergriete@culture.gouv.fr
01 75 47 21 34

Service de presse du CNRS - presse@cnrs.fr
01 44 96 51 51

Contacts chercheurs

Pour faciliter les échanges, merci d'écrire aux deux contacts en simultané.

Stéphane Lemaire, directeur de recherche CNRS, Laboratoire de biologie computationnelle et quantitative
stephane.lemaire@sorbonne-universite.fr

Pierre Crozet, maître de conférences à Sorbonne Université, Laboratoire de biologie computationnelle et quantitative
pierre.crozet@sorbonne-universite.fr