



QU'EST-CE QUE LE SÉQUENÇAGE DE NOUVELLE GÉNÉRATION ET POURQUOI EN AVONS-NOUS BESOIN ?

Subhrajit Barua^{1*}, Shinjini Bandopadhyay², Soham Biswas³ et Prabuddha Gupta^{2†}

¹Laboratoire de Chimie des Solutions des Matériaux et Technologies Avancées, Université ITMO, St. Petersburg, Russie

²Institut Amity de Biotechnologie, Université Amity de Kolkata, Kolkata, WB, Inde

³École des Sciences du vivant, Université de Hyderabad, Hyderabad, Inde

Savais-tu que ton ADN peut être lu ? Nous savons que l'ADN sert à stocker toute l'information qui fait de nous ce que nous sommes. Mais pour comprendre cette information, il faut d'abord apprendre à la lire. Les gènes sont des segments d'ADN qui, ensemble, agissent comme un manuel d'instructions dirigeant nos cellules, tout comme les codes qui indiquent à un ordinateur ce qu'il doit faire. Pendant de nombreuses années, nous n'avions aucun moyen de lire ces codes génétiques. Dans les années 1970, l'invention d'une technologie, connue sous le nom de séquençage de l'ADN, a rendu possible leur lecture. Le séquençage de l'ADN a permis aux scientifiques de lire, de comprendre et de comparer les informations génétiques, ce qui a constitué une avancée majeure dans notre compréhension de la biologie. Aujourd'hui, les technologies de séquençage jouent un rôle essentiel dans tous les domaines, de la médecine à l'agriculture. D'ailleurs, le séquençage de l'ADN a joué un rôle crucial lors de la pandémie de COVID-19, en nous aidant à étudier le coronavirus et à mettre rapidement au point des vaccins.

LA DÉCOUVERTE DE A, T, G, C

L'acide désoxyribonucléique (ADN), qui constitue le matériel génétique

de tout organisme vivant, a été découvert en 1869 par le chercheur suisse Friedrich Miescher qui étudiait les globules blancs. Au cours de son travail, il a isolé à partir des noyaux des cellules une nouvelle molécule qu'il a appelée nucléine. La structure de la molécule d'ADN a été identifiée par James Watson et Francis Crick en 1953, avec des contributions majeures de Rosalind Franklin [1]. La structure de l'ADN ressemble à une échelle en colimaçon (Figure 1).

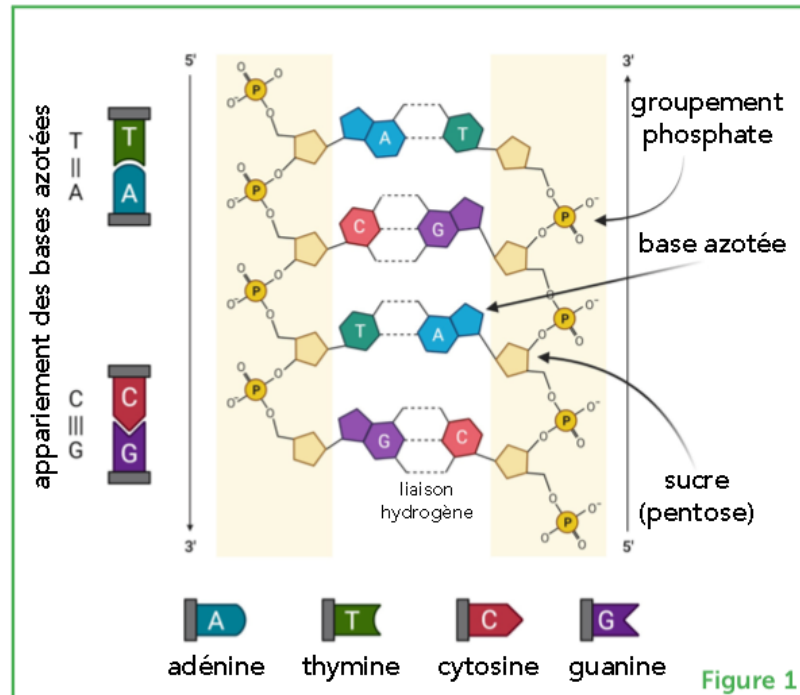


Figure 1. La structure en double hélice de l'ADN, ressemblant à une échelle en colimaçon, a été identifiée par Watson et Crick. Les montants de l'échelle sont constitués de sucres à cinq carbones reliés à des groupes phosphates. Chaque marche de l'échelle est formée par l'association de deux bases azotées, l'adénine (A) étant associée à la thymine (T), la cytosine (C) à la guanine (G). Cette association des bases azotées se fait par une liaison moléculaire, appelée liaison hydrogène. Les deux brins de la double hélice sont complémentaires. (Figure produite par <https://biorender.com/>)

BASE AZOTÉE. Molécule contenant de l'azote et composant important de l'ADN.

NUCLÉOTIDE. Composant de l'ADN qui contient une molécule de phosphate, une molécule de sucre et une base azotée.

GÈNE. Partie d'une séquence d'ADN qui contient l'information de fabrication d'une protéine.

GÉNOME. Ensemble complet des gènes d'un organisme vivant.

Les montants latéraux sont constitués d'un groupe phosphate (PO_4^-) et d'une molécule de pentose (sucre à cinq carbones), tandis que les barreaux sont formés par quatre substances appelées **bases azotées** : l'adénine (A), la guanine (G), la cytosine (C) et la thymine (T). L'association de ces bases est expliquée par la règle de Chargaff, qui stipule que A s'apparie toujours avec T, et G avec C. Ensemble, une molécule de phosphate, un sucre pentose et une base azotée constituent ce que l'on appelle un **nucléotide**. L'ordre des nucléotides dans l'ADN est très important car il représente le code génétique, tout comme l'ordre correct des lettres de l'alphabet crée des mots ayant un sens. Ces "mots" ou blocs d'information sont les **gènes**, et les gènes fournissent des instructions pour créer diverses caractéristiques, de la couleur des yeux au groupe sanguin. Le terme collectif pour tous les gènes présents dans le corps est le **génome**.

SÉQUENÇAGE DE L'ADN.

Procédé consistant à déterminer la séquence des bases azotées dans un fragment d'ADN.

Tous les membres d'une espèce ont des génomes similaires (mais pas identiques). Le génome humain possède des propriétés qui sont propres à l'espèce humaine, mais de petites variations dans les gènes font que chaque individu est légèrement différent.

PREMIÈRE MÉTHODE POUR DÉCHIFFRER LE CODE GÉNÉTIQUE : SÉQUENÇAGE DE SANGER

Après la découverte de Watson et Crick, les scientifiques avaient une compréhension claire de la structure d'une molécule d'ADN et de ses bases. Il s'agissait maintenant de déterminer l'ordre, ou la séquence, dans lequel les bases apparaissent le long de la molécule d'ADN. Cette procédure est appelée **séquençage de l'ADN**.

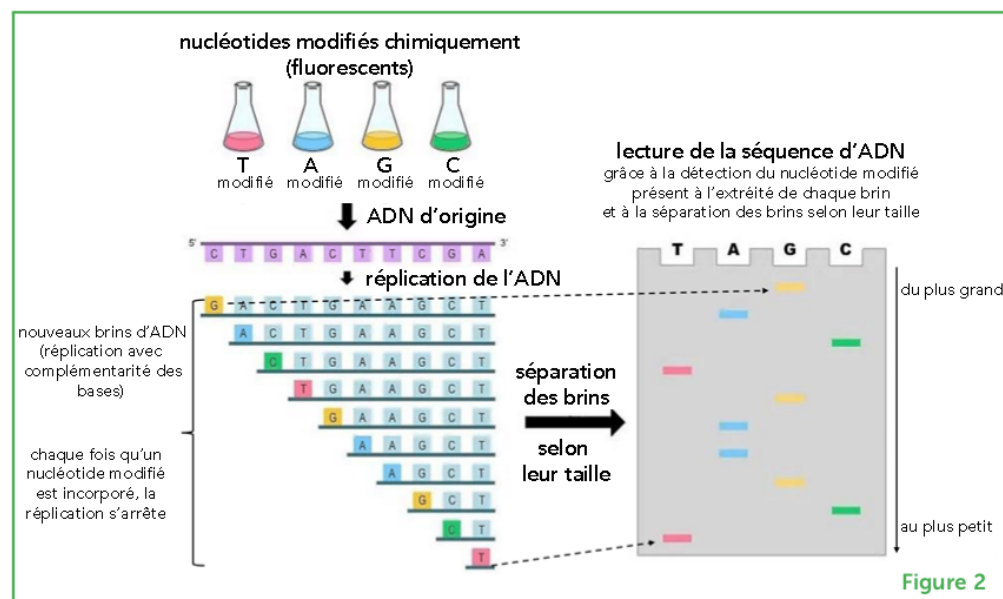


Figure 2. Méthode de séquençage de Sanger. Quatre mélanges réactionnels sont préparés, contenant l'ADN à séquencer simple brin, une polymérase et les quatre nucléotides A, T, G et C ; dans des cadres bleus sur la figure) ; pour chaque mélange, est ajouté un nucléotide (A, T, G ou C ; dans des cadres de couleur) marqué et modifié chimiquement de telle façon que la réplication de l'ADN s'arrête quand ce nucléotide modifié est incorporé. On obtient ainsi des fragments d'ADN nouvellement synthétisés complémentaires du brin original et de différentes longueurs, ce qui permet de les séparer par une technique appelée électrophorèse. Chaque fragment ne contient qu'un seul nucléotide modifié que l'on peut identifier par son marquage coloré ; on peut donc identifier quel nucléotide est présent à la fin de chaque fragment et reconstituer ainsi la séquence complémentaire du brin d'ADN séquençé. Pour déterminer la séquence de l'ADN original, il suffit d'utiliser la règle d'appariement des bases complémentaires. Ce séquençage manuel est un procédé fastidieux qui comporte de nombreuses possibilités d'erreur.

En 1975, Frederick Sanger a fait une immense percée dans le domaine du séquençage de l'ADN en mettant au point une méthode appelée « terminaison de chaîne » [2]. Dans nos cellules, l'ADN se réplique pour faire des copies identiques à lui-même en utilisant des nucléotides (A, T, G, C) comme éléments constitutifs. La méthode de séquençage de Sanger est basée sur le détournement du processus naturel de réplication de l'ADN, en ajoutant des nucléotides chimiquement

PROJET DU GÉNOME HUMAIN (PGH). Effort de recherche international visant à déterminer la séquence de tous les gènes humains.

SÉQUENÇAGE NOUVELLE GÉNÉRATION (NGS). Méthode moderne et rapide de séquençage de l'ADN qui permet de lire simultanément de très nombreuses séquences d'ADN.

modifiés, portant chacun une étiquette radioactive (plus tard, les étiquettes radioactives ont été remplacées par des étiquettes fluorescentes qui sont moins dangereuses). Ces nucléotides modifiés arrêtent la réplication de l'ADN à des endroits aléatoires, chaque fois que l'un d'eux est intégré dans la chaîne d'ADN en cours de réplication. Comme la réplication se termine de manière aléatoire, on obtient des fragments d'ADN répliqué de longueurs différentes, chaque fragment se terminant par un nucléotide modifié, radioactif ou fluorescent.

En examinant ces fragments, Sanger a pu reconstituer la séquence d'un brin d'ADN original en classant l'ensemble des fragments d'ADN répliqué selon leur taille, du plus petit au plus grand. Ainsi, le plus petit fragment, d'une longueur d'un nucléotide seulement, avait la version modifiée de A, T, G ou C, et lui indiquait la première base de l'ADN original en cours de séquençage (Figure 2). La méthode de séquençage de Sanger a ouvert la voie à des projets de séquençage d'ADN à grande échelle tels que le [Projet du génome humain](#).

LE PROJET DU GÉNOME HUMAIN

Le Projet du génome humain (PGH) était un projet international de grande envergure mené par des scientifiques du monde entier, qui ont commencé leurs recherches en 1990 et ont travaillé d'arrache-pied pendant 13 longues années pour séquencer le génome humain. Ils ont achevé ce séquençage (contenant 3,2 milliards de paires de bases) en 2003, en avance sur le calendrier prévu ! Le PGH a été un événement marquant. Pour la première fois, les scientifiques disposaient de l'intégralité de l'information génétique des humains, ouvrant ainsi la voie à de nouvelles découvertes passionnantes. Le séquençage du génome humain complet a constitué une étape importante dans l'amélioration des soins de santé [3].

Cependant, les techniques de séquençage de l'ADN utilisées pour obtenir ces résultats étaient longues et inefficaces. Malgré les résultats révolutionnaires du PGH, la naissance de technologies plus avancées, telles que le [séquençage de nouvelle génération \(NGS – New Generation Sequencing\)](#), a considérablement amélioré la précision du séquençage de l'ADN en minimisant les erreurs. Le NGS a révolutionné le secteur des soins de santé en réduisant le coût du séquençage et le temps nécessaire au diagnostic. Aujourd'hui, un génome humain peut être séquencé en seulement 6 heures !

QU'EST-CE QUE LE SÉQUENÇAGE DE NOUVELLE GÉNÉRATION ?

En 2006, une société du nom d'Illumina a annoncé une avancée majeure dans le séquençage de l'ADN, capable de séquencer simultanément un grand nombre de brins d'ADN. Cette méthode à haut débit permet de séquencer l'intégralité du génome humain en moins d'une journée. Contrairement à la méthode Sanger, le NGS ne dépend pas de la

terminaison de chaîne. Dans cette nouvelle méthode de séquençage, chaque fois qu'une nouvelle base est ajoutée au brin d'ADN en cours de réplification, un signal fluorescent est émis et détecté en temps réel (Figure 3).

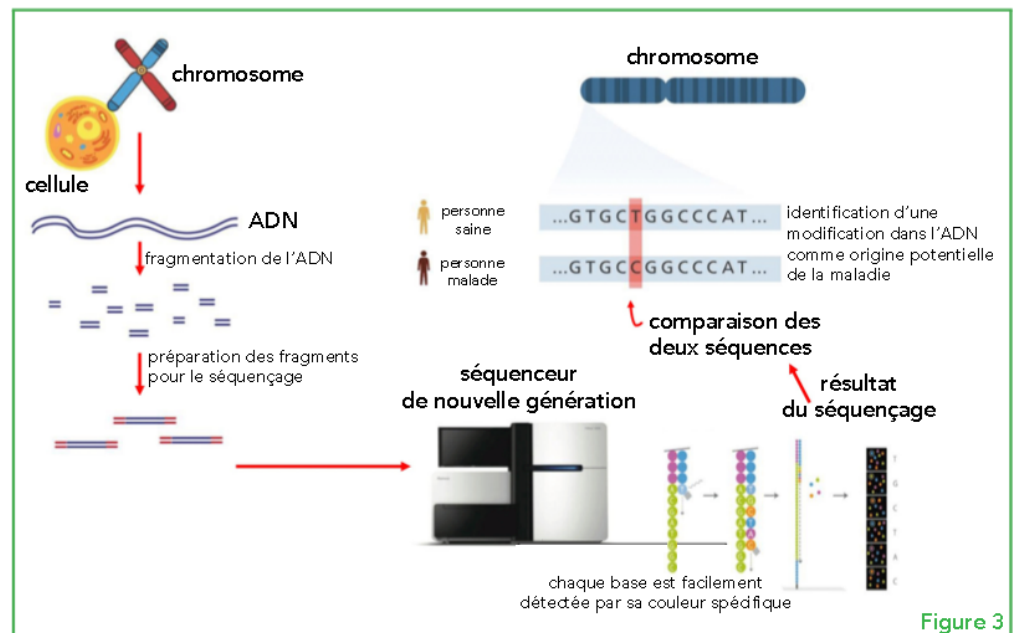


Figure 3. Dans le NGS, l'ADN est extrait des cellules et coupé en fragments plus petits. Ces fragments sont ensuite préparés pour le séquençage en les plaçant sur une plaque de verre appelée cellule de flux, qui contient le mélange réactionnel dans lequel chaque nucléotide a un marquage fluorescent qui lui est spécifique. Chaque fois qu'un nucléotide est ajouté, la cellule de flux l'identifie par sa fluorescence spécifique et le séquenceur l'enregistre. Ces signaux peuvent être interprétés par un logiciel pour construire une copie numérique de la séquence d'ADN originale. Les scientifiques peuvent alors utiliser ces données pour rechercher dans les gènes des changements qui pourraient être à l'origine de maladies comme le cancer.

COMMENT LE SÉQUENÇAGE DE L'ADN EST-IL UTILISÉ AUJOURD'HUI ?

Traitement du cancer

Le cancer a toujours posé un défi au monde médical car chaque type de cancer est différent et un même type de cancer peut agir différemment selon les patients. Le cancer peut être causé par des erreurs particulières dans l'ADN, et ces erreurs peuvent être identifiées par séquençage de l'ADN. Les scientifiques utilisent désormais le NGS pour analyser le génome des patients et identifier les erreurs d'ADN spécifiques à chaque patient. Cela aide les médecins à concevoir des traitements adaptés à chaque individu. Par exemple, deux patients atteints du même type de cancer peuvent réagir différemment au même médicament (un traitement qui peut guérir l'un échouera sur l'autre). De légers changements dans le génome des deux individus peuvent être détectés à l'aide du NGS, aidant les médecins à comprendre comment personnaliser la thérapie de différents individus.

MUTATION.

Tout changement ou altération de la séquence d'ADN qui peut ou non modifier la fonction d'un gène.

en fonction de leurs différences génétiques [4]. Ce type de traitement est appelé médecine personnalisée.

Évolution

La vie sur Terre a commencé par des organismes unicellulaires, qui sont devenus de plus en plus complexes. Certains animaux sont parmi les formes de vie les plus complexes sur Terre. Le NGS a été utilisé pour comparer le génome humain avec les génomes d'autres animaux, ce qui peut nous donner des détails sur la proximité ou la distance qui nous sépare d'eux. Par ailleurs, une part importante de l'évolution est liée à des **mutations** (modifications de l'ADN) chez certains organismes d'une espèce donnée, car chacune peut modifier de façon permanente une partie du génome de cette espèce. Le NGS permet aux scientifiques d'étudier comment les mutations se sont accumulées au cours de l'évolution.

Vaccins

De l'étude de l'origine du virus SRAS-CoV2 à la mise au point rapide de vaccins, le séquençage a été d'une aide précieuse lors de la pandémie de COVID-19. Les vaccins apprennent au système immunitaire à reconnaître les menaces et à les éradiquer. Le NGS a aidé les scientifiques à séquencer le virus SRAS-CoV2, et à mieux identifier les caractéristiques génétiques du virus qui menacent le système immunitaire, comme la protéine Spike. Le NGS a également été utile pour repérer de nouveaux variants viraux ; les virus évoluent eux aussi !

Allergies

Tu as probablement entendu parler de l'allergie au gluten ou de l'intolérance au lactose. Le NGS aide les scientifiques à déterminer si une personne est susceptible d'être allergique au gluten ou intolérante au lactose, en analysant son génome. Certaines séquences d'ADN peuvent rendre une personne plus susceptible de développer ces allergies.

Les microbes utiles

Tous les microbes vivant dans ton corps ne sont pas nocifs. Les bons microbes nous aident à combattre les infections. Le microbiote intestinal est une communauté de microbes qui vit dans le tube digestif humain et peut être considéré comme notre "second génome" car il influence notre système immunitaire et notre santé [5]. Le microbiote intestinal n'est pas facile à étudier en laboratoire, mais le NGS peut être utilisé pour analyser les génomes de ces organismes et comprendre leurs fonctions. L'American Gut Project a été une des plus grandes initiatives de science citoyenne au monde, en analysant les microbes vivant dans les intestins d'environ 11 000 personnes. Il a été rendu possible grâce au NGS (<https://www.facebook.com/AmericanGut>).

CONCLUSION

En résumé, le séquençage de l'ADN a révolutionné la recherche en sciences de la vie et ouvert la voie à des soins de santé plus précis et plus efficaces. Le séquençage de nouvelle génération (NGS) a rendu le séquençage de l'ADN plus rapide et plus abordable. Ces dernières années, des techniques de séquençage de l'ADN plus avancées, comme Oxford Nanopore, ont été mises au point. Ces méthodes sont encore plus rapides et plus précises ! Cependant, bien que nous puissions séquencer l'ADN, aucune méthode n'a encore été mise au point qui permette de séquencer entièrement une protéine quelle qu'elle soit. C'est l'un des développements les plus attendus en biologie. Espérons que l'un d'entre vous, jeunes lecteurs, sera en mesure de réaliser cet exploit !

RÉFÉRENCES

- [1] Watson, J., and Crick, F. 1953. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171:737–38. doi: 10.1038/171737a0
- [2] Heather, J. M., and Chain, B. 2016. The sequence of sequencers: the history of sequencing DNA. *Genomics* 107:1–8. doi: 10.1016/j.ygeno.2015.11.003
- [3] Green, E. D., Watson, J. D., and Collins, F. S. 2015. Human genome project: twenty-five years of big biology. *Nature* 526:29–31. doi: 10.1038/526029a
- [4] Gupta, A. K., and Gupta, U. D. 2014. "Next generation sequencing and its applications," in *Animal Biotechnology* (Elsevier). p. 345–67. doi: 10.1016/b978-0-12-416002-6.00019-5
- [5] Cao, Y., Fanning, S., Pross, S., Jordan, K., and Srikumar, S. 2017. A review on the applications of next generation sequencing technologies as applied to food-related microbiome studies. *Front. Microbiol.* 8:1829. doi: 10.3389/fmicb.2017.01829

VERSION FRANÇAISE

Cet article d'accès libre est une traduction avec modifications d'un article publié par Frontiers for Young Minds (doi : 10.3389/frym.2022.746502 ; Barua S, Bandopadhyay S, Biswas S and Gupta P (2022) What Is Next-Generation Sequencing and Why do we Need it? *Front. Young Minds.* 10:746502).

TRADUCTION : Nicole Pasteur, Association Jeunes Francophones et la Science

ÉDITION : Catherine Braun-Breton, Association Jeunes Francophones et la Science

MENTORS SCIENTIFIQUES : Charlotte André et Frédéric Lemoigno

REMERCIEMENTS : Merci à Laurent Laval pour son accueil et son implication dans l'édition de cet article par ses élèves.

JEUNES ÉDITEURS :

ZACKARIA, MATEO, GEORGIOS, YASSINE, MARGOT, AZIZA, 16-17 ANS

Nous sommes des adolescents en classe de première à Montpellier, dans le Sud de la France. Nous aimons les sciences et avons pour cela choisi les spécialités SVT, Mathématiques et Physique-Chimie.

ROSE, ELENA, SIBYLLE, LILY ET LOÏS, 16-17 ANS

Nous sommes en 1^{ère} générale à Montpellier, dans le Sud de la France et surtout, nous aimons le sport, danser et faire la fête ensemble.

ARTICLE ORIGINAL (VERSION ANGLAISE)

SOU MIS le 23 juillet 2021 ; **ACCEPTÉ** le 11 août 2022.

PUBLIÉ EN LIGNE le 31 août 2022.

ÉDITION : Michal Letek

MENTORS SCIENTIFIQUES : Ricardo Alberca , Demetrios Arvanitis , Andres Contreras, Thomas Li

CITATION : Barua S, Bandopadhyay S, Biswas S and Gupta P (2022) What Is Next-Generation Sequencing and Why do we Need it? *Front. Young Minds*. 10:746502. doi: 10.3389/frym.2022.746502

DÉCLARATION DE CONFLIT D'INTÉRÊTS : Les auteurs déclarent que les travaux de recherche ont été menés en l'absence de toute relation commerciale ou financière pouvant être interprétée comme un conflit d'intérêt potentiel.

DROITS D'AUTEURS

Copyright © 2022 Barua, Bandopadhyay, Biswas and Gupta

Cet article en libre accès est distribué conformément aux conditions de la licence Creative Commons Attribution (CC BY). Son utilisation, distribution ou reproduction sont autorisées, à condition que les auteurs d'origine et les détenteurs du droit d'auteur soient crédités et que la publication originale dans cette revue soit citée conformément aux pratiques académiques courantes. Toute utilisation, distribution ou reproduction non conforme à ces conditions est interdite.

JEUNES EXAMINATEURS

AYDEN, 14 ANS

Je m'appelle Ayden, et mes matières préférées sont l'histoire et l'anglais parce qu'on y apprend des faits fascinants et qu'on y découvre des livres intéressants. Je cours pour l'équipe d'athlétisme de mon école et j'adore le golf. J'aime passer du temps avec mes amis et jouer

à des jeux vidéo.

KALOMOIRA MARIA, 11 ANS

J'aime la nature (les forêts de conifères et les lacs), les jeux, les livres et la musique (surtout le piano). J'aime l'anglais et les mathématiques. Les sciences sont le dernier ajout à mes centres d'intérêt. Ma série de livres préférée est Pages & Co. Et mes héroïnes préférées sont Tilly des Bookwanderers, Anne aux pignons verts et la célèbre Hermione Granger de la série Harry Potter.

SANTIAGO, 12 ANS

Bonjour, je m'appelle Santiago. J'ai 12 ans, presque 13, vous pouvez m'appeler Santi. Mon sport préféré est le football. Je joue dans une équipe et je suis milieu de terrain. J'aime jouer avec mes amis. Je suis en 7^{ème} année. J'aime l'histoire et les sciences, surtout la chimie et les expériences en laboratoire.

SOPHIA, 13 ANS

Je m'appelle Sophia et j'ai 13 ans. Je suis brésilienne et j'aime jouer avec mes amis et regarder des vidéos sur Internet.

ZOE, 10 ANS

Je m'appelle Zoé et je vais entrer en 4^{ème}. L'un de mes passe-temps favoris est la danse, et mon meilleur souvenir est d'avoir dansé au Joyce Theatre à New York. J'adore faire du camping avec ma famille. Ce que je préfère, c'est faire griller des marshmallows ! Après une longue journée d'école, j'aime passer du temps avec mon Welsh Terrier, Duke.

AUTHORS

SUBHRAJIT BARUA

Je suis doctorante à l'Université ITMO de Saint-Pétersbourg, en Russie. J'étudie l'effet des biomolécules sur le système nerveux. Mes recherches font largement appel au NGS pour analyser le microbiome intestinal humain et celui de la souris. Je suis également astrobiologiste et j'étudie l'effet des voyages spatiaux sur les génomes des souris et d'autres organismes, toujours à l'aide du NGS. Restez en contact avec moi sur LinkedIn (<https://www.linkedin.com/in/subhrajit-barua>) et sur Twitter (<https://twitter.com/AstroSubhrajit>).
*subhrajit.barua11@gmail.com

SHINJINI BANDOPADHYAY

Je suis étudiante en troisième année de licence et j'achève actuellement ma licence en biotechnologie à l'Université Amity de Kolkata, en Inde. Mes recherches portent sur la biologie cellulaire et moléculaire du cancer, et j'ai récemment travaillé sur un projet de recherche visant à étudier la résistance immunothérapeutique chez les patients atteints de lymphome cutané à cellules T à l'aide de techniques NGS. Restez en contact avec moi sur LinkedIn : <https://www.linkedin.com/in/shinjinibandopadhyay>.

SOHAM BISWAS

Je suis étudiante en master de technologie en bioinformatique à l'Université d'Hyderabad (UoH), en Inde, où j'ai obtenu une bourse du gouvernement pour poursuivre mes études. Je travaille également comme chercheur diplômé au laboratoire de biologie moléculaire et structurale de l'Ecole des sciences de la vie de l'Université d'Hyderabad. Je m'intéresse à la recherche de protéines qui jouent un rôle crucial dans la survie des pathogènes leishmaniens, c'est-à-dire *Leishmania* sp. Contactez-moi sur LinkedIn à l'adresse <https://www.linkedin.com/in/soham-biswas>.

PRABUDDHA GUPTA

Prabuddha Gupta a effectué des recherches sur la division cellulaire des bactéries, la biologie cellulaire animale et le développement du poisson zèbre. Il enseigne à l'Université Amity de Kolkata. Il s'intéresse à de meilleurs outils d'enseignement pour la biotechnologie. Il est également intéressé par l'apprentissage et le partage de nouvelles approches de l'enseignement en ligne. orcid.org/0000-0003-1009-2781