

## ÉDITO



## Changement d'échelle

Grâce au progrès technologique, des quantités phénoménales d'informations sont aujourd'hui produites en recherche biomédicale comme en médecine: observations en microscopie, images médicales (IRM, scanners...), séquences génomiques... les chercheurs ont désormais accès à une masse d'informations stockées dans des entrepôts numériques, puis croisées, comparées, analysées à l'aide d'outils informatiques toujours plus puissants. Ce gigantesque changement d'échelle est une chance extraordinaire pour faire avancer nos connaissances sur la santé humaine. Comme vous le lirez dans ces pages, il est la clé d'une médecine « de précision », prenant en compte les caractéristiques propres à chacun d'entre nous, fait avancer la recherche sur l'autisme ou permet d'améliorer le diagnostic des cancers. Mais pour explorer ces quantités massives d'informations, nous avons besoin de moyens supplémentaires. Votre soutien est plus que jamais déterminant. Soyez-en ici vivement remerciés.

**Pr Stewart Cole,**

Directeur général de l'Institut Pasteur

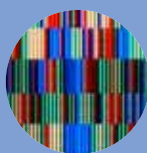
## LE DOSSIER



## Comment le *Big Data* révolutionne la recherche en santé

Identification de facteurs de risque de maladies, aide au diagnostic de cancers, au choix et au suivi de l'efficacité des traitements, surveillance des agents pathogènes... Dans de très nombreux domaines, la recherche biomédicale bénéficie aujourd'hui grâce aux progrès technologiques d'une avalanche d'informations – on parle de « données massives » ou *Big Data* – relatives aux maladies étudiées, source fantastique de nouvelles connaissances et de progrès médicaux. Des flux d'informations auxquels l'esprit humain ne peut plus faire face seul et ne peut gérer qu'à l'aide d'outils informatiques de plus en plus élaborés. De quoi s'agit-il ?

SUITE P. 2



**P. 09**

**GÉNOMIQUE**

Qu'est-ce que le  
séquençage ?



**P. 10**

**PORTRAIT**

Dr Halima Maïnassara



**P. 11**

**LE POINT SUR...**

Le prélèvement  
à la source et le don



Aujourd'hui, on peut obtenir la séquence du génome d'un individu en une journée.



## Maladies infectieuses

# Mieux traquer les agents pathogènes



L'Institut Pasteur héberge 14 « centres nationaux de référence », en lien avec Santé Publique France, chargés de la surveillance de maladies infectieuses en France et de l'alerte en cas d'épidémie : grippe, infections alimentaires (salmonelloses, listérioses...), rage, méningites bactériennes... « Depuis deux ans, au sein du programme PIBnet\*, les génomes des souches de bactéries, virus ou champignons isolés chez des patients en France et envoyés pour analyse à plusieurs de ces centres – soit actuellement environ 25 000 souches par an – sont systématiquement séquencés sur une plateforme mutualisée, et toutes les données sur chaque souche (provenance, virulence, présence ou non de gènes de résistance, etc.) sont conservées », explique **Sylvain Brisse**, responsable de l'unité Biodiversité et épidémiologie des bactéries pathogènes à l'Institut Pasteur et titulaire d'une Chaire d'excellence pour l'épidémiologie génomique en Santé globale. « Cette base de données permet de croiser nos informations avec celles obtenues dans d'autres pays du monde pour détecter la diffusion d'agents pathogènes à l'échelle internationale. »

\* Pasteur International Bioresources network.

## La génomique, reine du *Big Data*

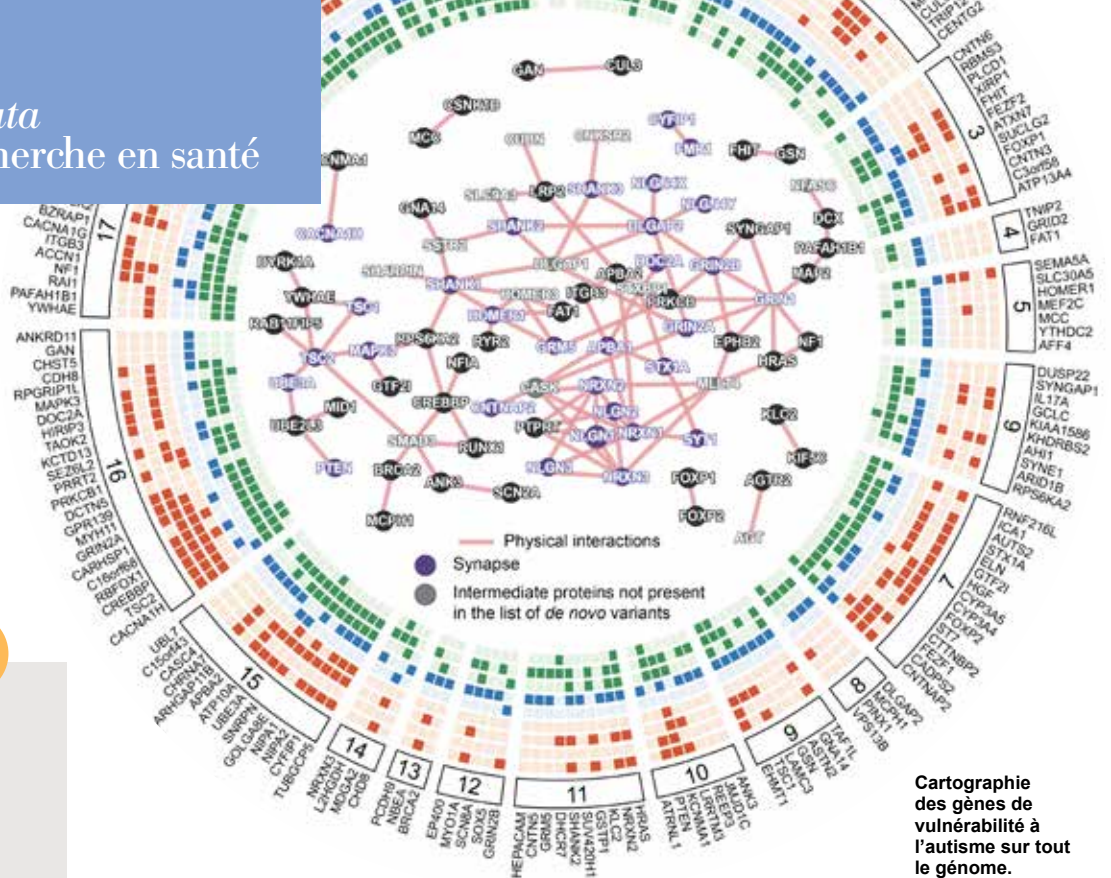
Prenons l'exemple de la génomique, considérée comme la « reine du *Big Data* ». L'obtention de la première séquence du génome humain (nos 23 000 gènes, soit un « texte » de 3 milliards de lettres), révélée en 2003, a pris près de 15 ans et coûté près de 3 milliards de dollars. Aujourd'hui, on peut obtenir la séquence du génome d'un individu en une journée, grâce aux techniques de séquençage « à haut débit », pour moins de 1 000 dollars. Ce coût réduit – qui continue de chuter... – permet désormais de séquencer le génome de patients volontaires dans nombre d'études, qu'elles concernent l'autisme, les cancers, les maladies neurodégénératives ou autres. Un généticien se basait il y a peu sur l'analyse de quelques gènes chez un nombre limité de patients. Il peut aujourd'hui accéder à l'information de l'ensemble du génome de centaines voire de milliers de personnes, associée bien sûr à d'autres données sur ces patients (âge, sexe, maladies dont ils souffrent, médicaments qu'ils prennent, etc.). Le défi majeur pour les chercheurs aujourd'hui est l'interrogation et l'analyse de ces données, « rangées » dans des « entrepôts numériques », des bases de données consultables *via* Internet.

## Donner du sens aux données

Partons d'un cas concret : on souhaite comparer des dizaines de génomes de patients répondant à un traitement avec ceux de

patients chez lesquels ce traitement est inefficace, pour voir si des différences génétiques existent entre les deux groupes afin d'identifier des marqueurs prédisant l'efficacité de la thérapie en question. Une telle étude n'est pas « humainement » réalisable : c'est l'outil informatique qui va, une fois programmé, effectuer cette vaste comparaison, passer en revue les 3 milliards de lettres de chaque génome, repérer des différences de texte entre les différents génomes, et identifier celles en relation avec la réponse au traitement. Une grande partie du travail des biologistes à l'initiative de l'étude va être d'interagir avec des bio-informaticiens, statisticiens et autres mathématiciens pour concevoir et développer des programmes et des algorithmes informatiques et statistiques dédiés à leur étude. Car comme le soulignent tous les chercheurs « *les données seules ne sont rien, il faut leur donner du sens* ».

La production exponentielle de séquences génomiques ne se limite pas au génome humain. On séquence à tout va l'ensemble de l'ADN contenu dans des échantillons de selles de patients – on parle de métagénomique – pour étudier le microbiote intestinal – ces 100 000 espèces microbiennes de notre flore intestinale dont la composition varie d'un individu à l'autre, ayant une influence majeure sur notre santé –, ou encore les génomes des virus, bactéries et autres microbes qui nous menacent, pour mieux les traquer (lire ci-contre).



Cartographie des gènes de vulnérabilité à l'autisme sur tout le génome.



Autisme

## La base de données la plus complète au monde hébergée à l'Institut Pasteur

L'Institut Pasteur va héberger la plus grande base de données au monde sur l'autisme, non pas en nombre de patients – 800 actuellement et jusqu'à 3 000 à terme – mais en quantité et richesse d'informations recueillies pour chacun : imagerie cérébrale, électro-encéphalogrammes, données génomiques, profils psychologiques, quotient intellectuel, qualité de vie, tests comportementaux, etc. « Comparée aux bases de données existantes pour l'autisme, c'est de la "haute-couture" », souligne Thomas Bourgeron (voir aussi p.7), responsable de l'unité de Génétique humaine et fonctions cognitives à l'Institut Pasteur, pionnier dans la découverte et l'étude des gènes associés à l'autisme. Ce chercheur coordonne le volet génétique du programme européen EU-AIMS\* impliquant 106 sites dans 37 pays d'Europe pour la recherche de nouveaux traitements contre l'autisme, grâce auquel ces multiples données ont été réunies. Dans son équipe, **Guillaume Dumas** précise : « L'élaboration de cette base de données est un défi technique, qui obéit aussi aux règles éthiques très strictes qui encadrent l'accès et la sécurisation des données de santé des personnes. Nous travaillons avec la Direction déléguée aux systèmes d'information (DSI) pour stocker et organiser le « rangement » de ces données, et avec des biostatisticiens du Centre de bioinformatique pour développer des outils pour leur analyse et faciliter l'accès des informations aux chercheurs. » Ce spécialiste des ondes cérébrales rappelle qu'il y a peu, il travaillait « avec de gros classeurs à l'hôpital contenant



des centaines de DVD gravés d'électro-encéphalogrammes des patients » ! La centralisation de toutes les données numérisées dans la nouvelle base EU-AIMS, disponible via Internet aux scientifiques du monde entier, va permettre aux chercheurs, depuis leurs ordinateurs, de croiser une multitude d'informations pour lancer des analyses robustes et reproductibles qui vont sans aucun doute accélérer à vitesse grand V la compréhension de l'autisme. Et, c'est l'objectif, aboutir à des traitements adaptés aux différents types d'autisme et aux caractéristiques de chaque patient. L'enjeu est de taille : rappelons que les « troubles du spectre autistique » touchent plus de 1% de la population.



\* European Autism Interventions – A Multicentre Study for Developing New Medications (EU-AIMS).



## Profession : Data manager

Après un master en bio-informatique, une thèse de doctorat en génomique des populations sur un petit poisson cavernicole du Mexique, puis un stage post-doctoral sur un poisson – cubain celui-là –, **Julien Fumey** change de voie. Particulièrement intéressé par l'autisme, il décroche, grâce à ses compétences en informatique et en analyse des génomes, un poste de « Data/project manager », ouvert en janvier dernier à l'Institut Pasteur dans le laboratoire de Thomas Bourgeron : un nouveau métier né de l'essor du *Big Data*. « Je dois coordonner la gestion des données du laboratoire et du projet EU-AIMS (lire ci-dessus), qu'il s'agisse de leur stockage ou des démarches juridiques et légales », explique Julien, nouveau lien entre les chercheurs du laboratoire, les cliniciens des hôpitaux en charge des données des patients, les informaticiens et autres biostatisticiens du campus, et surtout le service juridique et la cellule éthique. « Avec les poissons, c'était simple, il suffisait d'une autorisation pour les collecter et les transporter », remarque Julien. « Avec l'humain, j'ai découvert une réglementation extrêmement stricte, avec de multiples demandes d'autorisation auprès de comités d'éthique, de la CNIL, etc. Tout ce qui touche à la protection et à la sécurisation des données relatives aux patients prend énormément de temps, d'autant que l'Institut Pasteur a un haut niveau d'exigence dans ce domaine. » Une des missions du Data manager...

Après un master en bio-informatique, une thèse de doctorat en génomique des populations sur un petit poisson cavernicole du Mexique, puis un stage post-doctoral sur un poisson – cubain celui-là –, Julien Fumey change de voie. Particulièrement intéressé par l'autisme, il décroche, grâce à ses compétences en informatique et en analyse des génomes, un poste de « Data/project manager », ouvert en janvier dernier à l'Institut Pasteur dans le laboratoire de Thomas Bourgeron : un nouveau métier né de l'essor du *Big Data*. « Je dois coordonner la gestion des données du laboratoire et du projet EU-AIMS (lire ci-dessus), qu'il s'agisse de leur stockage ou des démarches juridiques et légales », explique Julien, nouveau lien entre les chercheurs du laboratoire, les cliniciens des hôpitaux en charge des données des patients, les informaticiens et autres biostatisticiens du campus, et surtout le service juridique et la cellule éthique. « Avec les poissons, c'était simple, il suffisait d'une autorisation pour les collecter et les transporter », remarque Julien. « Avec l'humain, j'ai découvert une réglementation extrêmement stricte, avec de multiples demandes d'autorisation auprès de comités d'éthique, de la CNIL, etc. Tout ce qui touche à la protection et à la sécurisation des données relatives aux patients prend énormément de temps, d'autant que l'Institut Pasteur a un haut niveau d'exigence dans ce domaine. » Une des missions du Data manager...



**Titan Krios™, le microscope le plus puissant du monde, peut fournir en une heure une quantité d'images équivalent en volume à 500 films vidéos en haute définition.**

### Un tsunami d'images

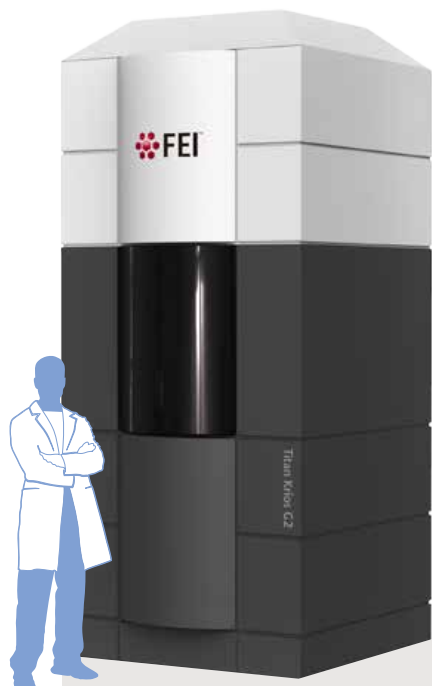
La deuxième « reine » du *Big Data*, c'est l'imagerie, qu'il s'agisse d'imagerie médicale ou de microscopie. Les microscopes de plus en plus puissants et automatisés génèrent un tsunami d'images : Titan Krios™, le microscope le plus puissant du monde, peut fournir en une heure une quantité d'images équivalent en volume à 500 films vidéos en haute définition (lire ci-dessous), impossible à analyser par des cerveaux humains en un temps raisonnable. Là encore, le développement d'outils informatiques *ad hoc* est nécessaire. En matière d'analyses d'images, les chercheurs utilisent de plus en plus l'intelligence artificielle, et notamment le *deep learning* ou apprentissage profond, né de l'essor du *Big Data* et de l'augmentation de la puissance de calcul des ordinateurs (lire L'entretien ci-contre). L'an dernier, grâce à plus de 100 000 images de grains de beauté déjà caractérisés comme bénins ou malins, une équipe américaine a ainsi pu entraîner un ordinateur à effectuer le diagnostic de mélanomes. Une étude a récemment comparé les performances de la machine à celles de 58 médecins spécialistes, de 17 pays. Résultat : « *la plupart des dermatologues ont fait moins bien* », écrivent les chercheurs, 87 % de bon diagnostic en moyenne contre 95 % pour l'ordinateur. Bien « entraîné », l'ordinateur est donc

plus performant... et n'a ni coup de fatigue, ni chute de concentration ! Ces techniques sont en passe de révolutionner le diagnostic des maladies et l'interprétation d'images médicales (radios, scanners, IRM, observation microscopiques des tissus de biopsies...) sera de plus en plus confiée à des ordinateurs, plus rapides et plus fiables, laissant plus de temps au médecin pour d'autres activités de recherche et pour sa relation au patient.

### Le *Big Data* a poussé les chercheurs à partager leurs données

Données génomiques, imagerie... le nombre et la qualité des données collectées chez un même patient ne cesse de croître, avec des centaines d'informations recueillies chez un même individu, contre une dizaine il y a quelques années. Au Royaume-Uni, le projet « UK biobank » collecte une multitude de données de santé sur pas moins de 500 000 volontaires britanniques, dans le but « *d'aider à améliorer la prévention, le diagnostic, et le traitement de nombreuses maladies sévères* » - cancers, maladies cardiovasculaires, diabète, arthrite, ostéoporose, dépression ou démences... Cette base est publique, gratuitement accessible aux chercheurs *via* Internet, sur inscription. La seule contrepartie demandée aux utilisateurs est d'informer

SUITE P. 6



Le microscope le plus puissant du monde, Titan Krios.™



## Des données titanesques !

En juillet dernier était inauguré à l'Institut Pasteur le « bâtiment Nocard » abritant le microscope le plus puissant du monde, Titan Krios™ – le seul de ce type en France dédié à l'étude des maladies, acquis avec l'aide de nombreux donateurs. Permettant l'observation à l'échelle atomique, ce cryo-microscope est automatisé, capable d'observer plusieurs échantillons à la fois et par là-même un gigantesque producteur de *Big Data*. « *Le stockage des données brutes générées par Titan est un premier défi majeur... elles pourraient remplir en peu de temps tous les disques durs du campus !* », souligne **Michael Nilges**, directeur du Département de Biologie structurale et chimie. « *De plus, il faut les analyser et les traiter pour déduire des images en trois dimensions. Quelques jours d'observations par Titan peuvent nécessiter des mois d'analyse informatique. Il nous faut augmenter la capacité de calcul de l'Institut.* ». En plus d'une équipe experte en microscopie à très haute résolution chargée de l'utilisation du microscope\*, un groupe de scientifiques spécialement dédié au traitement des données a été recruté.

\* Unité d'Études structurales de machines moléculaires in cellulo, dirigée par Dorit Hanein.





## L'ENTRETIEN

## Christophe Zimmer

Responsable de l'unité Imagerie et modélisation à l'Institut Pasteur.

## « Le *Big Data* est au cœur du renouveau actuel de l'intelligence artificielle. »

### **Vous dirigez une équipe spécialisée dans l'imagerie\*. En quoi le *Big Data* a-t-il modifié la donne ?**

Associé à l'augmentation des capacités de calcul des ordinateurs, le *Big Data* a remis au goût du jour des techniques assez anciennes d'intelligence artificielle, qui donnent aujourd'hui des résultats précisément parce qu'on dispose de données massives. Il s'agit en particulier du *deep learning* ou apprentissage profond, dont on ne parlait pas il y a encore 10 ans. Grâce à ces techniques, on peut désormais par exemple entraîner un algorithme à distinguer des images de chiens et de chats – en lui donnant des milliers d'images d'apprentissage (annotées avec des 0 pour les chiens et des 1 pour les chats). Après un très grand nombre de calculs sur ces données d'entraînement, l'algorithme devient capable de généraliser et de reconnaître un chat ou un chien à partir d'une image inconnue. Et plus on entre de données, moins il se trompe. La reconnaissance vocale de votre portable, la traduction automatique des textes sur internet, la voiture automatique, les robots capables de manipuler des objets font tous appel au *deep learning*, qui arrive également dans le domaine médical. S'il commence à être utilisé en génomique, l'imagerie médicale est sans doute aujourd'hui le domaine dans lequel ses percées sont les plus évidentes.

### **Quelles perspectives d'applications médicales permet l'intelligence artificielle ?**

Nous travaillons par exemple avec l'unité de Neuropathologie expérimentale, dirigée par Fabrice Chrétien, sur le diagnostic de tumeurs cancers du cerveau de l'enfant à

partir de biopsies après opération chirurgicale, provenant des hôpitaux Sainte-Anne et Necker à Paris. Actuellement, pour une même biopsie, deux pathologistes peuvent diverger sur le diagnostic dans un pourcentage notable de cas. Or diagnostiquer avec la plus grande précision le type de tumeur est déterminant pour la suite du traitement : chimiothérapie, radiothérapie, chirurgie supplémentaire, etc. Pour analyser la biopsie, l'expert dispose d'une image contenant typiquement 5 milliards de pixels – pour comparaison, une photo classique est composée d'environ 1 million de pixels seulement – et n'a pas le temps de l'inspecter entièrement au niveau maximal de résolution. Comme dans l'exemple des chats et des chiens, nous apprenons à un algorithme à déterminer le degré de sévérité d'une tumeur en l'entraînant à partir de centaines d'images de biopsies bien caractérisées par un groupe d'experts. Notre objectif à terme est de faciliter et compléter l'analyse histopathologique

par l'expert humain, pour rendre le diagnostic plus fiable mais aussi plus rapide, car il peut s'écouler plusieurs jours avant qu'un spécialiste ne dispose de l'intégralité des tests pour expertiser une biopsie. En attendant d'atteindre cet objectif, les outils que nous développons pourraient déjà servir à attirer l'attention de l'expert sur des zones de l'image particulièrement informatives.

### **Ces méthodes peuvent-elles aussi accélérer la recherche ?**

Absolument. À l'Institut Pasteur, l'équipe d'Arnaud Echard (unité de Trafic membranaire et division cellulaire) étudie la toute dernière étape de la division cellulaire, l'abscission, qui aboutit à la séparation des cellules, et dont les dysfonctionnements sont associés à la formation de nombreux cancers. L'équipe d'Arnaud a identifié 400 gènes potentiellement impliqués dans ce processus. Pour déterminer le rôle exact de ces candidats, il faut inactiver chaque gène un par un, filmer les divisions cellulaires et déterminer le moment de la scission, qui varie d'une cellule à l'autre. Les chercheurs font cela en inspectant les films à l'œil, mais pour déterminer l'effet d'un seul gène, il faut observer environ 300 divisions cellulaires, soit 15 heures de travail, à faire trois fois pour s'assurer de la reproductibilité, donc au total 45 heures par gène. Multiplié par 400, on arrive à 18 000 heures, soit plusieurs années de travail à temps plein ! Nous mettons donc en place un algorithme de *deep learning* en vue d'analyser chaque film presque instantanément. Si nous réussissons, cette équipe pourra rapidement connaître quels gènes influencent l'abscission, ce qui pourrait ouvrir la voie à des traitements contre les cancers.

“

Pour l'imagerie médicale, notre objectif est de compléter l'analyse par l'être humain pour rendre le diagnostic plus fiable et plus rapide. ”

\* Christophe Zimmer a une formation d'astrophysicien. Son équipe comprend des mathématiciens, des physiciens, des informaticiens, des biologistes et une chimiste.



# 1 000 volontaires « en bonne santé » pour une médecine de précision

« *Quels sont les facteurs génétiques, environnementaux ou liés au mode de vie qui interviennent sur notre réponse immunitaire et font que nous sommes plus ou moins enclins à développer une maladie, à réagir efficacement ou non à tel traitement ou à un vaccin ? C'est à cette question ambitieuse, clé d'une médecine plus personnalisée, que nous cherchons à répondre* », explique **Lluis Quintana-Murci** (voir aussi p.7), responsable à l'Institut Pasteur de l'unité de Génétique évolutive humaine et coordinateur du Consortium Milieu Intérieur\*, qui réunit 30 équipes de diverses disciplines scientifiques (immunologie, génomique, biologie moléculaire, bioinformatique) de plusieurs institutions de recherche et hôpitaux. « *Nous étudions une cohorte\*\* de 1 000 volontaires sains de 20 à 70 ans, 500 femmes et 500 hommes. Nous avons déjà recueilli pour chacun une masse considérable de données et en ce moment même, leurs génomes complets sont en train d'être séquencés* », précise le chercheur. Des échantillons de sang, de salive, de selles (pour l'étude du microbiote intestinal) ont été collectés pour chaque volontaire et un questionnaire d'environ 200 variables renseigné : âge, poids, taille, antécédents médicaux, habitudes de sommeil, nutrition, tabagisme ou non, données socio-démographiques (revenus, niveau éducatif...), etc. Pour chaque individu, une partie du sang prélevé a été divisé



en une trentaine de tubes, chacun immédiatement stimulé avec un agent pathogène (virus de la grippe, BCG, staphylocoques...) pour mimer une infection et mesurer la réponse immunitaire. « *Ces multiples données nous permettent une analyse très poussée combinant – à l'aide d'outils informatiques – les facteurs environnementaux et socio-démographiques avec des facteurs biologiques, notamment le génome, pour savoir avec un haut niveau de précision ce qui participe à la variabilité du système immunitaire.* » ajoute Lluis Quintana-Murci. « *L'espoir avec ces Big Data est d'appliquer les connaissances issues de cette étude à une médecine*

*de précision, consistant à adapter la bonne stratégie thérapeutique à un patient selon ses caractéristiques personnelles, son profil immunitaire propre, que ce soit vis-à-vis de maladies infectieuses, inflammatoires ou dans le cadre de l'immunothérapie des cancers.* » Les premiers résultats ont déjà montré l'impact majeur du tabagisme sur la composition des globules blancs dans le sang, expliquant une plus grande sensibilité aux infections, et permis d'identifier des centaines de variations génétiques modifiant la production de molécules clés de la réponse immunitaire – dont certaines associées à un plus grand risque de développer des maladies comme l'allergie aux pollens, le lupus érythémateux ou le diabète de type 1.

\* Le Consortium Milieu Intérieur a été labellisé « Laboratoire d'excellence » dans le cadre du programme gouvernemental des Investissements d'Avenir.

\*\* Cohorte gérée par un centre de recherche biomédicale à Rennes.

Les serveurs de stockage des données occupent des volumes de plus en plus considérables.

UK Biobank des découvertes utilisant ces données « afin que les autres scientifiques puissent en bénéficier ». Ceci illustre un phénomène concomitant à l'avènement du *Big Data* : le partage des données, le *data sharing*. Les chercheurs du monde entier mettent désormais leurs ressources en commun, collaborent pour regrouper leurs données et les analyser et donnent accès aux bases de données qu'ils organisent (lire p.3). Une condition *sine qua non* pour exploiter aux mieux les kyrielles d'informations existantes.

## Un défi de taille : le stockage des données

Un autre défi considérable du *Big Data* – technique et économique – est de stocker les données, les organismes de recherche devant acquérir toujours plus de serveurs de stockage et de supercalculateurs. Celles

d'un seul génome humain peuvent occuper un espace équivalent à 100 000 photos de vacances. Dans le cadre d'une vaste banque de données européenne sur l'autisme (lire p.3), 800 000 euros ont par exemple été consacrés au seul stockage des données génomiques des patients...

Mais des avancées inestimables sont attendues du *Big Data* en recherche biomédicale : des déductions beaucoup plus robustes que par le passé sont permises par des études sur de grands nombres de patients, l'accès à de nombreuses informations sur chaque individu permet de mieux connaître les spécificités de chacun, la part d'influence de son génome, de son microbiote, de son environnement (mode de vie) sur sa réponse aux maladies, aux traitements ou aux vaccins (lire ci-dessus). À la clé : une médecine de précision – préventive, personnalisée, prédictive et participative.

DOSSIER RÉALISÉ PAR LA RÉDACTION



# 3 découvertes, 3 regards sur le *Big Data*

## Le 10<sup>e</sup> gène du **VIH**



**Olivier Gascuel** s'est notamment illustré en 2015 en débusquant, grâce à l'analyse génomique de 25 000 souches du virus du sida (VIH)\* un gène « caché », dont l'existence était suspectée depuis la fin des années 80, mais très controversée. Il a ainsi confirmé que le génome du virus

était bien composé de 10 gènes et non de 9, ouvrant ainsi de nouvelles perspectives thérapeutiques.

Ce chercheur pionnier de la bioinformatique dirige aujourd'hui l'unité de Bioinformatique évolutive et le Centre de bioinformatique, biostatistique et biologie intégrative (C3BI) de l'Institut Pasteur. Le C3BI regroupe une cinquantaine de bioinformaticiens recrutés depuis 2014 pour épauler les équipes du campus en développant des méthodes informatiques et statistiques pour analyser leurs quantités massives de données. « **Jusqu'à la fin des années 90, la biologie se faisait devant des tubes à essai ou des boîtes de Petri. Maintenant, elle se pratique aussi bien derrière un ordinateur qu'à la paillasse.** » souligne Olivier Gascuel.

**Le virus du sida a été isolé pour la première fois à l'Institut Pasteur en 1983, découverte qui a valu un prix Nobel de médecine à deux de ses chercheurs. En janvier 1985, d'autres scientifiques pasteurien – dont l'actuel directeur général de l'Institut Pasteur Stewart Cole – obtenaient la première séquence génétique du virus, une prouesse à cette époque « pré-génomique », avant l'avènement des techniques de séquençage à haut débit.**

## L'influence de nos gènes sur **l'empathie**



« Le *Big Data* a remis en question certaines connaissances », explique **Thomas Bourgeron**, responsable de l'unité de Génétique humaine et fonctions cognitives à l'Institut Pasteur. « Il était par exemple acquis que les autistes présentaient des anomalies dans une

région du cerveau, le corps calleux. Ceci avait été avancé par des études sur de petits nombres de patients. Mais en les reproduisant non pas sur 20 personnes mais sur 500 ou plus, il s'est avéré que cette conclusion était fautive. **Aujourd'hui, grâce au *Big Data*, nos analyses sont beaucoup plus robustes.** »

Le chercheur a notamment participé à une étude menée sur 46 000 personnes sur l'empathie, cette faculté d'apporter une réponse émotionnelle adaptée face aux pensées et sentiments d'autrui. **Elle suggère que notre empathie n'est pas seulement le résultat de notre éducation et de notre expérience, mais est aussi en partie influencée par nos gènes :** un dixième de la variation du degré d'empathie entre les individus serait dû à des facteurs génétiques...

\* Menée par des chercheurs de l'Université de Cambridge, de l'Institut Pasteur, de l'université Paris Diderot, du CNRS et de la société de génétique 23andMe.

## Héritage de l'Homme de **Néandertal** et immunité



« L'arrivée du *Big Data* nous a ouvert la fenêtre d'une façon spectaculaire », remarque **Lluís Quintana-Murci**, responsable de l'unité de Génétique évolutive humaine. « Nous pouvons par exemple aujourd'hui étudier le système immunitaire inné, notre première ligne de défense contre les pathogènes, dans son ensemble. Alors qu'il y a peu nous analysons deux ou trois gènes impliqués dans cette immunité, nous en étudions désormais plus de 1 500. » Son équipe a récemment montré que les gènes impliqués dans notre système immunitaire inné sont enrichis en héritage de l'Homme de Néandertal. Cette connaissance de l'histoire évolutive humaine – tout comme la comparaison des génomes actuels de

différentes populations (Européens, Africains, Asiatiques...) – permet de repérer les gènes les plus importants, ceux qui nous ont permis de nous adapter au mieux et de survivre aux pathogènes. « **Le *Big Data* a révolutionné nos recherches sur la biodiversité des génomes. Or étudier la variabilité génétique des populations est primordial pour comprendre nos inégalités face à la maladie, mais aussi face à la réponse aux traitements ou aux vaccins.** »



### Unis pour comprendre l'émergence des maladies

Ces trois chercheurs et plusieurs autres équipes de l'Institut Pasteur participent au vaste projet **Inception**, lancé en 2016 dans le cadre du programme gouvernemental « Investissements d'avenir ». À la croisée de la biologie « intégrative », des sciences sociales et de l'analyse des données, il a pour but d'étudier l'émergence des maladies infectieuses (Ebola, Zika, sida...) et autres (cancers, autisme...) chez les individus et dans les populations.

## HÉPATITE C

# Nouveau test de diagnostic

**Un test de dépistage du virus de l'hépatite C, transportable au chevet du patient, rapide et fiable, vient d'être développé\*.**

Ce nouveau test de détection, accessible aux populations des pays aux ressources limitées, permettra une prise en charge immédiate du patient suite au diagnostic – réalisable en une heure – et une mise en place sans tarder d'un traitement contre la maladie. L'hépatite C chronique touche environ 1 % de la population mondiale (71 millions de personnes) et est responsable de 400 000 décès par an, dus à l'évolution sévère de la maladie vers la cirrhose ou le cancer du foie.

Son dépistage précoce est d'autant plus important que la plupart des personnes infectées ne connaissent par leur statut alors que les antiviraux les plus récents permettent de traiter 95% des cas... s'ils sont pris à temps.

\* Par l'équipe de Darragh Duffy dans l'unité d'immunobiologie des cellules dendritiques (Institut Pasteur / Inserm), en collaboration avec la société Genedrive.



## MÉNINGITE À MÉNINGOCOQUES

# Mal de ventre : un symptôme qui doit alerter

**Les patients infectés par des méningocoques présentent de la fièvre, des vomissements, des raideurs dans la nuque... mais ils peuvent aussi avoir simplement mal au ventre.**

Tellement mal, qu'on les opère parfois, à tort, d'appendicite. Une étude\* vient de montrer que 10% des patients infectés par une souche de méningocoques en pleine expansion en Europe présentent des douleurs abdominales. Une forme atypique et de plus en plus fréquente de la maladie, dont les médecins doivent rapidement prendre conscience. Rappelons que l'infection invasive à méningocoques est une maladie mortelle dans quasiment 100% des cas s'il l'on n'administre pas rapidement des antibiotiques.

\* Menée par le Centre national de référence des méningocoques à l'Institut Pasteur, dirigé par Muhamed-Kheir Taha, en collaboration avec une équipe du service de pédiatrie de l'hôpital Bicêtre AP-HP.

## E-SANTÉ

# Toi Même® : une application pour mesurer l'évolution des troubles de l'humeur

**Une application destinée à mesurer l'activité des troubles de l'humeur, en particulier les troubles bipolaires, est en développement à l'Institut Pasteur\*.**

Issu de la recherche clinique et mis au point avec des médecins psychiatres, le système Toi Même® est entré en phase test de faisabilité clinique en avril dernier, et sera utilisé par une centaine de patients au cours des prochains mois. Caractérisés par la récurrence d'épisodes d'exaltation de l'humeur (manie) et d'états dépressifs, et par des perturbations marquées de la régulation émotionnelle, de l'activité psychomotrice et cognitive, les troubles bipolaires sont associés à un risque élevé de rechutes et d'hospitalisations, souvent dues à des interventions thérapeutiques tardives par rapport à l'apparition des symptômes et à des réponses partielles au traitement.



Toi Même®, en proposant à l'utilisateur de remplir des tests évaluant son humeur et ses comportements, permet de mesurer les variations de l'humeur et les variations comportementales et émotionnelles en temps réel chez un individu. Le système collecte aussi, grâce aux capteurs de son smartphone, des informations concernant l'activité physique et sociale de l'utilisateur. L'évaluation répétée en temps réel de multiples paramètres chez les personnes ayant un trouble bipolaire devrait aider à préciser les diagnostics, à mieux connaître les facteurs prédictifs de réponses aux

traitements et à terme de mieux comprendre les mécanismes qui sous-tendent ces troubles. Toi Même® devrait aussi permettre aux patients de mieux connaître leur maladie et, en lien avec leur psychiatre, de devenir plus actif dans leur prise en charge.

Cette étude a reçu un financement d'AG2R LA MONDIALE.

\* Par le Dr Aroldo A. Dargél et le Pr Chantal Henry, psychiatres-chercheurs de l'unité Perception et mémoire, dirigée par Pierre-Marie Lledo, avec la Direction des systèmes d'information et le Centre de recherche translationnelle de l'Institut Pasteur.



GÉNOMIQUE

# Qu'est-ce que le séquençage ?

**L**e séquençage génétique ou séquençage de l'ADN consiste à décrypter tout ou partie de l'information génétique d'un organisme.

Il s'agit précisément de révéler le « texte génétique » constitué d'un alphabet à quatre lettres – **A, T, G et C** – qui correspondent aux quatre molécules ou bases se succédant sur l'ADN (**Adénine, Thymine, Guanine et Cytosine**). **L'ADN étant le support des gènes (23 000 chez l'Homme), chaque gène correspond à une phrase particulière de ce texte. Connaître la séquence d'un gène peut par exemple servir à poser un diagnostic chez une personne en identifiant des erreurs dans le texte (mutations)**. Les premières techniques de séquençage ont été développées dans les années 1970, permettant en 1977 d'obtenir la première séquence d'un génome\* entier, celui d'un bactériophage (virus de bactéries). Des centaines de génomes d'organismes vivants (bactéries, mammifères, plantes...) ont depuis été intégralement séquencés. Initié en 1989, le séquençage du génome humain (**3 milliards de lettres !**) fut lui finalisé en 2003 grâce au développement des séquenceurs automatiques, pour un coût de 2,7 milliards de dollars. Aujourd'hui, avec l'apparition du séquençage à haut débit au milieu des années 2000, l'intégralité du génome d'un individu peut être séquencée en quelques jours pour moins de 1 000 euros ! Ce séquençage individuel est de plus en plus utilisé en médecine pour poser le diagnostic de certaines maladies, évaluer certaines prédispositions à des pathologies (médecine prédictive) ou encore ajuster certains traitements (médecine personnalisée). En recherche, le séquençage de masse des génomes de très nombreux patients est par exemple effectué pour identifier des gènes associés à des maladies.

\* Le génome est l'ensemble de l'information génétique d'un organisme et correspond à l'ADN (sauf chez certains virus où il s'agit d'ARN).

FOCUS

## Cancers du sein, du côlon, du rectum : à quel stade sont-ils diagnostiqués en France ?



**Un rapport\* estime pour la première fois la répartition des stades au diagnostic de trois types majeurs de cancers : 60% des cancers du sein, 44% des cancers du côlon et 47% des cancers du rectum sont diagnostiqués à un stade précoce en France, et cette précocité est davantage constatée chez les personnes de moins de 75 ans.**

Ces cancers sont en effet davantage diagnostiqués à un stade précoce chez les 40-74 ans tandis qu'une part plus importante de cancers diagnostiqués au stade avancé est observée chez les personnes âgées de 75 ans et plus.

### Cancer du sein: seul 1 cancer sur 10 est diagnostiqué à un stade avancé

Avec plus de 59 000 nouveaux cas chaque année et 12 000 décès par an, le cancer du sein est plus fréquent chez la femme, et l'un des plus mortels. 6 cancers du sein sur 10 sont diagnostiqués à un stade précoce (extension locale limitée) et plus fréquemment chez les 50-74 ans, classe d'âge cible du dépistage organisé ; 3 sur 10 à un stade intermédiaire (extension régionale) ; 1 sur 10 à un stade avancé.

### Les cancers du côlon et du rectum sont rares avant 50 ans

Dans l'étude, 1% avaient entre 15 et 39 ans, 3% entre 40 et 49 ans. Avec 45 000 nouveaux cas détectés et 18 000 décès chaque année, le cancer du côlon-rectum reste le 2<sup>e</sup> cancer le plus meurtrier. Il est diagnostiqué à un stade avancé dans environ un tiers des cas, chez l'homme comme chez la femme.

Ces informations sur le degré de sévérité de ces cancers lors du diagnostic vont permettre « d'améliorer le suivi épidémiologique des cancers les plus fréquents pour lesquels un diagnostic précoce est possible et d'adapter les politiques publiques de prévention et de lutte contre le cancer », conclut le rapport.

\* Réalisé à partir des registres de cancer sur la période 2009-2012 en France métropolitaine par Santé publique France, l'Institut national du cancer (INCa), le Réseau français des registres des cancers (réseau Francim) et le Service de biostatistique des Hospices civils de Lyon (HCL).

**DR HALIMA  
MAÏNASSARA**

Directrice générale du Centre de Recherche Médicale et Sanitaire (CERMES) à Niamey au Niger



“

Au sein du Réseau international des instituts Pasteur, on ne se sent pas seul, on est dans un groupe. C'est une force immense!”

**P**remière femme nigérienne à occuper le poste de directrice générale du Centre de Recherche Médicale et Sanitaire (CERMES) à Niamey, depuis février 2017, le Dr Halima Maïnassara est médecin, épidémiologiste et spécialiste des méningites. Elle était entrée en 2005 comme enquêtrice de terrain au sein de cet établissement public de référence, sous tutelle du ministère nigérien de la Santé Publique et membre du Réseau international des instituts Pasteur.

Dès l'école primaire, Halima Boubacar Maïnassara, née en 1975 d'une enseignante et d'un conseiller pédagogique, rêvait de devenir médecin : « *je voulais être la personne en blouse qui soigne les petits.* » Son bac avec mention lui permet d'obtenir une bourse pour aller étudier à l'étranger, au Mali, où elle obtient en 2004 son diplôme à la Faculté de Médecine et de Pharmacie de Bamako : « *j'aimais le contact avec les patients, la confiance qu'ils nous accordent. Le médecin a une place importante dans la société.* ». De retour à Niamey, sa thèse de médecine portant sur l'échographie, la jeune docteure travaille dans ce domaine mais aussi aux urgences de l'Hôpital National. Le CERMES cherche alors des médecins pour réaliser des enquêtes épidémiologiques, et Halima est recrutée comme enquêtrice. « *C'est comme cela que j'ai mis le pied dans la recherche. J'aimais beaucoup ce travail de terrain. En parallèle, je consultais pour le dispensaire du Centre [fermé depuis] et restais donc en contact avec les patients.* »



En 2007, Halima Maïnassara participe à une enquête de santé dans le village de Mozagué (sud du Niger).



Le CERMES à Niamey (Niger).

Elle complète ensuite son cursus et enchaîne les diplômes : en *statistique appliquée à la médecine* à l'Université Pierre-et-Marie-Curie (UPMC) à Paris, par des cours à distance, en *Pratique de la santé publique dans les pays en situation de développement* à l'université de Montpellier ou encore en *Essais cliniques et maladies infectieuses et tropicales* à l'Institut Pasteur. Dès 2012, Halima, mère de trois enfants, multiplie les allers-retours entre Niamey et Paris pour préparer sa thèse de science, et obtient en 2017 un doctorat en *Santé publique, option Épidémiologie et sciences de l'information biomédicale* à l'UPMC.

Elle est nommée cette même année directrice du CERMES : « *Lors de l'appel à candidatures, j'étais occupée par ma thèse. Mais on m'a encouragé à postuler et j'ai tenté ma chance!* ».

La méningite – « *un sujet de travail historique au CERMES, toujours d'actualité car il y a des épidémies chaque année au Niger* » – reste une de ses priorités : « *les décideurs attendent beaucoup des chercheurs, pour repérer le germe en cause afin d'utiliser le vaccin approprié.* ». Mais, ajoute-t-elle : « *il ne faut pas se limiter aux maladies infectieuses, même si cela reste majeur. Nous avons aussi des problèmes de santé publique comme l'hypertension ou le diabète qui méritent des investigations dans nos pays. La santé mère-enfants est aussi l'une de mes ambitions.* ». Pour mener à bien ses missions, Halima Maïnassara sait qu'elle peut s'appuyer sur le Réseau des instituts Pasteur : « *au sein du Réseau, on ne se sent pas seul, on est dans un groupe. Nous menons beaucoup de projets communs, nous partageons des idées et des valeurs. C'est une force immense!* ».





# Le prélèvement à la source et le don

Pour la santé de tous, vous donnez, vous déclarez, vous déduisez !

## Exemple

**Vous effectuez un don de 100 € le 23 septembre 2018 en faveur de l'Institut Pasteur, qui vous adressera alors un reçu fiscal pour l'année 2018.**

**Au printemps 2019, vous indiquerez votre don sur votre déclaration de revenus 2018. Votre réduction d'impôt de 66 € vous est alors restituée en septembre 2019 par l'administration fiscale.**



## Le prélèvement à la source en 2019 ne changera rien à vos déductions fiscales

Nous savons que vous vous interrogez sur les déductions fiscales liées à vos dons effectués en 2018, lors de l'entrée en application du prélèvement à la source à partir du 1<sup>er</sup> janvier 2019. En réalité, le prélèvement à la source implique uniquement un changement dans le mode de collecte de votre impôt, retenu chaque mois sur vos sources de revenus (fiche de paie, pension retraite, indemnités chômage etc.) afin de s'adapter immédiatement à votre situation. **Ainsi tous les crédits et réductions d'impôt sont maintenus dans les mêmes conditions que les années passées, y compris pour vos dons effectués sur l'année 2018.**

## Vous devrez toujours déclarer vos dons réalisés en 2018

**Au printemps 2019**, vous devrez effectuer votre déclaration d'impôts sur le revenu comme par le passé. À ce moment-là, vous pourrez déclarer vos dons effectués en 2018 au profit d'organismes comme l'Institut Pasteur.

## Vos déductions fiscales sont maintenues

Concrètement, comme l'année précédente, vos dons réalisés en 2018 ouvriront le droit à une déduction fiscale en 2019. Le pourcentage sur vos dons sera identique à aujourd'hui, c'est-à-dire **66 % pour l'impôt sur le revenu ou 75 % pour les personnes assujetties à l'impôt sur la Fortune Immobilière (IFI).**

## Vous bénéficiez de votre avantage fiscal en 2019

**L'État vous restituera le montant correspondant à votre réduction fiscale, en septembre 2019**, au moment du nouveau calcul de l'impôt sur le revenu sur la base des déductions fiscales prévues en 2018.

## Donner 3 fois plus ne coûte pas plus !

Votre déduction fiscale		
Vous voulez donner	Donnez plutôt	L'État vous rembourse
30 €	90 €	59,40 €
50 €	150 €	99,00 €
100 €	300 €	198,00 €

Ainsi, un don de 100 € ne vous revient en réalité qu'à 34 €, alors que nos chercheurs disposent bien de 100 € pour agir pour notre santé à tous.

## VOTRE DON DE 100 €

Votre don vous coûte réellement

34 €

Votre déduction fiscale

66 €

Grâce à votre don, les chercheurs pourront continuer d'œuvrer au quotidien pour notre santé à tous ! **MERCI.**

PASTEURDON

12<sup>e</sup> édition  
« Défendons la recherche » !



**E**n cette année célébrant ses 130 ans, l'Institut Pasteur, créé grâce à la générosité de nombreux donateurs, a toujours, et plus que jamais, besoin de dons pour faire avancer ses recherches d'excellence au bénéfice de la santé humaine.

Nos chercheurs espèrent donc une mobilisation sans précédent pour la 12<sup>e</sup> édition du Pasteurdon qui se tiendra du 11 au 14 octobre prochains. Alexandra Lamy, fidèle marraine de l'opération, 24 chaînes de télévisions, 15 radios, soutiendront le même message :

« Depuis 130 ans l'Institut Pasteur se bat avec détermination et pourtant... face aux nouvelles menaces qui pèsent sur la santé – résistance aux antibiotiques, cancers, maladies liées au vieillissement tel qu'Alzheimer ou Parkinson, maladies infectieuses comme le sida ou encore Zika –, l'Institut Pasteur manque de moyens. C'est seulement avec votre soutien qu'il pourra continuer ses grandes découvertes médicales.

**Pour le progrès et pour l'avenir de tous, défendons la recherche. »**

Merci d'avance !

Pour en savoir plus :  
[pasteurdon.fr](http://pasteurdon.fr)

SPORTIFS SOLIDAIRES

Vivons Vélo pour le Pasteurdon !

**F**aire du vélo peut être un acte solidaire ! Le groupe de protection sociale AG2R LA MONDIALE organise plusieurs fois par an des rassemblements à vélo au profit des chercheurs de l'Institut Pasteur.

Le prochain aura lieu le **samedi 13 octobre** dans de nombreuses villes de France à l'occasion du Pasteurdon. Dans chaque ville, un parcours original est proposé, et pour chaque cycliste, **les kilomètres parcourus sont transformés en dons** par AG2R LA MONDIALE. Une autre façon d'alimenter le compteur au profit de l'Institut Pasteur est de télécharger l'application « Vivons Vélo » (sur smartphone et tablette) qui permet de mesurer ses performances lors d'une sortie vélo et là encore de transformer les kilomètres pédalés en dons, à tout moment de l'année. Parlez-en autour de vous !

Pour s'inscrire à l'opération Vivons Vélo du 13 octobre : [www.inscription-vivonsvelo.fr/](http://www.inscription-vivonsvelo.fr/)



Protection des données à caractère personnel

**L**e 25 mai dernier entré en vigueur la nouvelle réglementation de l'Union européenne relative au traitement des données à caractère personnel et à leur protection : le règlement général sur la protection des données ou RGPD.

Il vise à créer un cadre harmonisé pour la protection des données à caractère personnel dans l'ensemble des pays de l'Union européenne et renforce les droits des citoyens.

En tant qu'organisme à but non lucratif faisant appel à la générosité du public, l'Institut Pasteur s'est toujours attaché à faire de la protection des données personnelles un principe fondamental guidant ses pratiques. Les données personnelles vous concernant sont ainsi utilisées dans le seul but de vous informer et de collecter des dons pour faire avancer la recherche. Votre soutien et votre générosité sont nos plus précieux atouts.

Nous sommes vigilants à respecter vos droits et souhaits relatifs à vos données personnelles et notre service donateurs est à votre disposition les concernant :

- par email à [dons@pasteur.fr](mailto:dons@pasteur.fr),
- par téléphone au 01 40 61 33 33,
- par **courrier**.

Vous pouvez également accéder à vos informations dans votre espace donateurs sur [www.pasteur.fr](http://www.pasteur.fr).



BULLETIN D'ABONNEMENT et/ou DE SOUTIEN

Merci de bien vouloir nous le retourner à : Institut Pasteur – 25 rue du Docteur Roux – 75015 Paris



Je fais un don de :

30€  45€  60€  75€  100€  Autre montant.....€

Sur [www.pasteur.fr](http://www.pasteur.fr)

Par chèque bancaire à l'ordre de l'Institut Pasteur

Je veux continuer à recevoir la Lettre de l'Institut Pasteur et je vous joins le montant de mon abonnement pour un an : soit 4 numéros au prix de 6 euros (non déductible).

Les données recueillies vous concernant sont nécessaires au traitement de votre don et à l'émission de votre reçu fiscal. Conformément à la loi Informatique et Libertés, vous disposez d'un droit d'accès, de rectification, de radiation sur simple demande écrite à l'Institut Pasteur – 25-28, rue du Docteur Roux-75724 Paris Cedex 15. Vos coordonnées peuvent être communiquées à d'autres organismes ou associations faisant appel à la générosité du public ou envoyées hors Union Européenne, sauf avis contraire de votre part en cochant la case ci-contre ( ).

MES COORDONNÉES

Nom .....

Prénom .....

Adresse .....

La lettre de l'Institut Pasteur Institut Pasteur

Lettre trimestrielle éditée par l'Institut Pasteur

Directeur de la publication : Stewart Cole • Directeurs de la rédaction : Jean-François Chambon, Frédérique Chegaray • Rédactrice en chef : Corinne Jamma • Ont participé à la rédaction de ce numéro : Adeline Fougère, Aurélie Perthuisson, Myriam Rebeyrotte, Olivier Rescaillère • Direction artistique, réalisation : BRIEF • Crédit photos : ©Institut Pasteur/F. Gardy, ©Institut Pasteur/T. Lang, ©Giovanni Cittadini Cesi, Réseau International des Instituts Pasteur, Shutterstock, D.R. • Impression : Imprimerie Guillaume • N° de commission paritaire : 0122 H 88711 • ISSN : 1243-8863 • Abonnement : 6 euros pour 4 numéros par an • Contact : Institut Pasteur – 25, rue du Docteur Roux 75015 Paris – Tél. 01 40 61 33 33

Cette lettre a été imprimée sur du papier et selon des procédés de fabrication respectueux de l'environnement.

[www.pasteur.fr](http://www.pasteur.fr) [dons@pasteur.fr](mailto:dons@pasteur.fr)