



ACADÉMIE  
DES SCIENCES  
INSTITUT DE FRANCE

# Les traductions des *Comptes Rendus*

---

## *Biologies*

### La biologie dangereuse des germes pathogènes

Corps de l'article traduit par Patrick Berche<sup>®</sup>

Date de publication : 2024-11-28

<https://doi.org/10.5802/crbior.157-fr>

**Résumé :** La convergence des biotechnologies avec de nombreuses disciplines, notamment l'informatique et l'Intelligence Artificielle (IA), peut permettre des manipulations génétiques dangereuses sur les germes pathogènes. Il peut s'agir d'expériences de gain de fonction exacerbant la virulence, telles que réalisées sur les *Myxovirus* et les coronavirus. On peut aussi synthétiser n'importe quel microorganisme à partir des séquences *in silico*, y compris les virus les plus dangereux (poxvirus, Ebola, etc.) dont les séquences sont accessibles. On pourrait même concevoir par IA des germes nouveaux qui pourraient être des armes biologiques.

**Mots-clés :** Gain de fonction, Biologie synthétique, Arme biologique, Biosécurité

Article original :

Patrick Berche, "The dangerous biology of pathogenic germs". *Comptes Rendus Biologies*, vol. 347 (2024), p. 77-86. <https://doi.org/10.5802/crbior.157>

Les traductions des *Comptes Rendus* sont mises à disposition du public dans un but d'accessibilité mais n'ont pas reçu de validation scientifique avant leur publication, seule la publication initiale engage la responsabilité de l'Académie des sciences. Le projet Traduction des *Comptes Rendus* est soutenu par le ministère de l'Enseignement supérieur et de la Recherche et par la délégation à la langue française et aux langues de France du ministère de la Culture.



Nous assistons à une révolution sans précédent des sciences biologiques, marquée par la convergence de la biologie avec de nombreuses autres disciplines, telles que la chimie, la physique, les biotechnologies, les nanotechnologies, les mathématiques, l'informatique et l'intelligence artificielle, etc. Les impacts technologiques majeurs ont été le séquençage rapide de l'ADN (1977) permettant la constitution de banques de données, la découverte de la transcriptase reverse (1970) permettant de recopier l'ARN en ADN complémentaire (ADNc), et la synthèse chimique de l'ADN à partir de séquences *in silico*.

On peut désormais modifier ou même créer des entités biologiques modifiées (gènes, protéines, micro-organismes, virus, etc.) présentant des propriétés nouvelles, surtout pour des applications médicales. Des expériences inactivant certains gènes de germes pathogènes par des mutations ou des délétions permettent d'identifier ceux impliqués dans la virulence et la contagiosité. Cela permet de lier le phénotype au génotype des pathogènes, ce qui est utile pour préparer des vaccins et des drogues anti-infectieuses. *A contrario*, on peut exacerber la virulence en modifiant ou en introduisant certains gènes impliqués dans la pathogénicité dans d'autres micro-organismes, ce qui est facilité par la technologie CRISP-Cas9 [1]. Ce sont les expériences de gain de fonction (*gain-of-function, GoF*) or DURC (*Dual-use research of concern*). Bien que la majorité de ces expériences soient conçues pour des fins utiles, certaines posent des problèmes de recherche duale et de biosécurité avec d'éventuels échappements accidentels de laboratoires, surtout quand il s'agit de virus pathogènes potentiellement pandémiques (*Potentially Pandemic Pathogens, PPP*), tels que notamment les virus influenza, les coronavirus ou les poxvirus.

Grâce au séquençage de l'ADN, on dispose de dizaines de milliers de séquences complètes d'agents pathogènes accessibles dans les banques de données. Par exemple, on dispose de 11 694 séquences complètes de génomes viraux qui ont été déposées dans le *Viral Genome Resource* [2], incluant tous les virus pathogènes connus. On a désormais la capacité de synthétiser des génomes entiers de micro-organismes à partir des séquences *in silico*. Cela « dé-sanctuarise » les agents infectieux les plus dangereux, comme les poxvirus ou les filovirus, dont les génomes sont accessibles en dehors des stocks hautement protégés. On pourrait même en concevoir des virus nouveaux avec l'aide de l'intelligence artificielle.

## 1. Les expériences de gain de fonction

De nombreux micro-organismes pathogènes inféodés de longue date à des espèces animales sauvages circulent silencieusement chez les hôtes qui constituent des réservoirs viraux. Très adaptés à chaque espèce, ils reconnaissent des récepteurs spécifiques qui leur permettent de coloniser et de pénétrer les cellules-hôtes. Rencontrant de nouveaux hôtes, ils ne sont le plus souvent pas infectieux, car il existe une barrière d'espèce qui prévient une éventuelle colonisation. Le franchissement de cette barrière est rarement possible directement, car elle dépend de la fréquence des interactions entre les espèces animales et de leur proximité génétique, mais aussi de la plasticité des génomes des micro-organismes.

Les micro-organismes pandémiques sont souvent des virus à RNA à fort taux de mutations. Les principaux réservoirs sauvages sont les oiseaux, les chauves-souris et les rongeurs. Chez ces animaux, ces virus existent sous forme de quasi-espèce qui sont des essaims de plusieurs génotypes. Le virus sauvage émergent est celui qui possède le génotype qui prolifère le plus efficacement [3, 4]. Le passage à l'homme est un processus séquentiel nécessitant le plus souvent des adaptations progressives à travers des hôtes animaux intermédiaires chez qui des mutations adaptatives surviennent et s'accumulent [5]. Ces hôtes intermédiaires sont identifiés pour les myxovirus aviaires (porc), le SARS-CoV1 de 2002 (civettes, chiens viverrins, blaireaux), le MERS-CoV (dromadaires) et le virus Nipah (chevaux) [6]. Le plus souvent les premiers contacts humains avec des virus sauvages sont des échecs, aboutissant au mieux à des cas sporadiques d'infections non contagieuses ou à des séroconversions asymptomatiques [7]. Par exemple, c'est le cas de la grippe aviaire H5N1 chez l'homme [8] ou du SARS de 2002, où près de 80 % des différents animaux des marchés de Guangzhou présentaient des anticorps anti-SARS-CoV1 [9].

Les expériences de gain de fonction peuvent mimer ce processus d'adaptation des virus sauvages. On peut augmenter la virulence et la contagiosité d'un agent infectieux par plusieurs approches. Depuis longtemps, on

sait créer des mutations aléatoires qui augmentent progressivement la virulence par passages sur animaux ou cultures cellulaires. On peut aussi réaliser des manipulations génétiques induisant des mutations, par la mutagenèse dirigée ou l'insertion de matériel génétique. Plus récemment, on a réussi à créer des mutants chimériques par recombinaison de certains virus pathogènes, par exemple pour modifier leur tropisme cellulaire. Il peut arriver qu'une expérience aboutisse à un gain de fonction de façon inattendue. Par exemple en 2001, des chercheurs australiens voulant développer des vaccins viraux immunocontraceptifs ont généré un nouveau virus de la variole des souris (ectromélie) d'une extrême virulence. En ajoutant le gène murin de l'interleukine -4 (IL-4), le virus recombinant a entraîné une mortalité de toutes les lignées de souris, y compris après vaccination [10]. Le nouveau virus a entraîné une suppression totale des réponses NK et T cytotoxiques. Sachant que le virus de la variole humaine est apparenté à ce virus et présente la même organisation génétique, on peut craindre d'avoir publié la recette d'une arme biologique redoutable. Fallait-il publier ses résultats?

Devant les menaces réitérées de pandémies, liées à la grippe aviaire (H5N1, H7N9, etc.), au SARS (2002) et au MERS (2012), les chercheurs ont tenté de comprendre les mécanismes moléculaires de la pathogénicité de ces virus en étudiant les principaux facteurs de virulence comme notamment l'hémagglutinine HA des myxovirus et la protéine Spike des coronavirus. Le but était de prévoir l'émergence d'un nouveau virus potentiellement mieux adapté à l'homme à partir des mutations créées dans certains gènes de virulence putatifs. C'est ainsi que des expériences dangereuses de gain de fonction ont été réalisées sur des virus influenza et des coronavirus.

### 1.1. *Virus influenza*

Le virus de la grippe, *Myxovirus influenza*, est constitué d'ARN monocaténaire (12-15 kb), codant 8 gènes portés par des segments séparés, dont les deux gènes de virulence codant l'hémagglutinine HA et la neuraminidase NA, exprimées à la surface de l'enveloppe du virus. Lors de co-infections avec plusieurs virus grippaux chez le porc, des réassortiments peuvent survenir, permettant l'émergence de nouveaux virus potentiellement pandémiques. On craint aujourd'hui l'émergence d'une nouvelle pandémie de grippe à partir du virus aviaire H5N1, qui a déclenché des épizooties avec des cas humains à partir de 1996 [11]. Ce virus reconnaît les acides sialiques (AS) aviaires (AS- $\alpha$  2,3 galactose), mais ne reconnaît pas les AS humains (AS  $\alpha$ -2,6 galactose) présents sur les voies respiratoires supérieures, expliquant sa non-contagion pour l'homme. Cependant, après l'inhalation profonde au contact des volailles, le virus peut atteindre les bronchioles et les alvéoles humaines où le récepteur aviaire est présent. Cela explique la forte mortalité et la faible contagiosité du virus pour l'homme. Une alerte pandémique a été déclenchée en 2005 à Hong Kong, à la suite de plusieurs cas humains groupés, mais sans véritable épidémie [12]. Entre 2005 et 2024, on a dénombré 888 cas de grippe aviaire chez l'homme, dont 463 décès (mortalité, 52 %) [13]. Cependant, on peut craindre que quelques mutations suffisent pour transformer le virus H5N1 en un virus pandémique redoutable. Pour tenter de prédire le risque de l'émergence d'une nouvelle pandémie, deux équipes ont cherché en 2012 à identifier les mutations cruciales du gène de l'HA qui permettraient au virus H5N1 de devenir contagieux pour l'homme en reconnaissant les récepteurs AS humains. Des chercheurs de Madison ont remplacé le gène H1 de l'hémagglutinine d'une souche H1N1 par le gène H5 d'une souche aviaire H5N1 [14]. Le nouveau virus n'est pas contagieux chez le furet, contrairement à la souche H1N1. Des mutations aléatoires du gène H5 ont été créées in vitro, puis sélectionnées par adsorption sur des globules rouges de dinde exprimant l'AS humain. Ainsi, 370 mutants ont été obtenus et criblés individuellement pour sélectionner ceux capables de se lier très efficacement à l'AS humain. Sur les neuf mutants ainsi isolés, tous présentaient des mutations dans le *Receptor Binding Domain* (RBD), dans la région 120-259 de l'hémagglutinine H5. Quatre d'entre eux étaient contagieux par voie nasale pour le furet et peu virulents (mutations N186K, S227N, Q226L et G228S). Ces mutants provoquent une mortalité en 3 à 7 jours après inoculation intra-trachéale et intra-nasale ( $10^6$  virus). Le virus réassorti H5 transmissible reconnaît préférentiellement les récepteurs de type humain, se réplique efficacement chez les furets, provoque des lésions pulmonaires et une perte de poids, mais n'est pas hautement pathogène et n'entraîne pas de mortalité. Une autre équipe, à Rotterdam (Pays-Bas), a réalisé une mutagenèse dirigée pour introduire deux mutations dans le gène de l'hémagglutinine d'une souche H5N1 (Q222L et G224S), pour lui permettre de se lier à l'AS humain, et une mutation dans

le gène PB2 codant l'ARN-polymérase pour la rendre active à 33 °C au lieu de 40 °C. Ce mutant H5N1, qui n'est pas transmissible par voie aérienne, a ensuite été inoculé par voie nasale à des furets pour forcer l'adaptation par des passages répétés (10 au total). Des virus transmissibles par voie aérienne ont été identifiés, tous porteurs de deux mutations supplémentaires dans le gène de l'hémagglutinine (H103Y et T156A). Cela suggère que seules cinq mutations (quatre dans le gène HA et une dans le gène PB2) pourraient suffire à conférer au virus H5N1 une transmission par voie aérienne entre mammifères [15].

Constatant la fréquence des co-infections de porcs par des souches aviaires H5N1 et des souches humaines H1N1, une équipe chinoise a reproduit en culture cellulaire une co-infection, pour générer des souches pandémiques émergentes [16]. Ils ont généré tous les réassortiments possibles entre deux souches : une souche H5N1 hautement pathogène pour la souris (portant deux mutations facilitant la réplication et la liaison AS chez les mammifères) mais incapable de se transmettre efficacement par les gouttelettes respiratoires chez le cobaye, et une souche H1N1 transmissible par voie aérienne chez le cobaye mais peu virulente chez la souris. Sur les 127 virus hybrides, 35 étaient plus pathogènes que la souche H5N1. Parmi ceux-ci, certains étaient également capables d'une transmission très efficace par gouttelettes respiratoires entre cobayes. Ces expériences dangereuses, publiées en 2012 et 2013, montrent que quelques mutations permettant une meilleure réplication et une meilleure liaison au récepteur SA étaient suffisantes pour franchir la barrière d'espèces.. Ces expériences de gain de fonction générant des virus potentiellement pandémiques ont entraîné un moratoire décrété aux États-Unis d'octobre 2014 à décembre 2017. En octobre 2022, une importante épizootie à virus H5N1 est survenue dans des élevages de visons en Espagne, touchant près de 52 000 animaux, avec une mortalité passant de 0,7 % au début à 4,3 % au bout de trois semaines. L'analyse des génomes des virus a permis de détecter des mutations portant sur la polymérase PB2 et des mutations du gène codant HA différentes de celles prédites antérieurement, ainsi que des mutations supplémentaires dans d'autres gènes [17].

## 1.2. *Coronavirus*

Les coronavirus sont constitués d'un ARN monocaténaire de 26-32 kb et d'une enveloppe hérissée de spicules constituées de la protéine S (Spike). Le génome du virus du SARS (*Severe Acute Respiratory Syndrome*) de 2002 comporte 11 gènes. Il existe chez l'homme 7 coronavirus pathogènes. Quatre (HCoV-229E, HCoV-NL63, HCoV-OC43, HCoV-HKU1) entraînent des infections respiratoires épidémiques bénignes (rhinites, etc.) chez les enfants. Trois autres sont responsables d'infections respiratoires sévères : SARS-CoV1 du SARS en 2002-2003 (8 346 cas, 646 décès, mortalité 7,8 %); le MERS-CoV du MERS (*Middle East Respiratory Syndrome*) apparu au Moyen-Orient en 2012 (> 2000 cas de 2012 à 2024, mortalité 35 %); le SARS-CoV2 responsable de la pandémie de Covid-19 apparue en décembre 2019 et évoluant par vagues itératives jusqu'à aujourd'hui [18] (mortalité 0,6 %, > 14 millions de décès). Le SARS-CoV1 et le SARS-CoV2 reconnaissent les récepteurs ACE2 (*Angiotensine-Converting Enzyme 2*), le MERS-CoV le récepteur DPP4 (*human Dipeptidyl-Peptidase 4*). Les chauves-souris constituent le réservoir sauvage des 3 coronavirus pathogènes pour l'homme. On a retrouvé des hôtes intermédiaires pour le SARS-CoV1 (civette palmée, chien viverrin, blaireau) et pour le MERS-CoV (dromadaire). Pour le SARS-CoV2, aucun hôte intermédiaire n'a été identifié après trois ans de pandémie. Son origine reste controversée [19].

À Wuhan, l'émergence du SARS en 2002 a stimulé une recherche importante sur les coronavirus avec, à partir de 2004, la constitution d'une collection de prélèvements provenant des chauves-souris capturées dans les grottes du sud de la Chine et du Sud-Est asiatique. Ainsi a été constitué jusqu'à aujourd'hui un stock de près de 15 000 échantillons (sang, salive, urine, etc.), permettant d'identifier par RT-PCR et séquençage plus de 200 coronavirus apparentés au SARS-CoV-1, dont une centaine de séquences ont été publiées. Le réservoir du virus SARS-CoV1 a été identifié en 2005 chez *Rhinolophus spp*) [20]. En 2013, des chercheurs chinois découvrent dans des prélèvements de chauves-souris plusieurs coronavirus dont les séquences du gène spike sont très similaires à celles du SARS-CoV1. Ces virus peuvent se lier et se multiplier dans des cellules HeLa exprimant des récepteurs ACE-2 de l'homme, de la civette ou de *Rhinolophus affinis* [21]. Dans une grotte de la province du Yunnan en 2017, ces chercheurs chinois ont découvert 8 virus inconnus de chauves-souris étroitement

apparentés au SARS-CoV1, dont 2 avec une protéine Spike fonctionnelle et 6 avec des délétions du RBD. La Spike d'un des virus (Rs4874) est identique (99,9 %) à celle du SARS-CoV1 [22].

En 2008, une équipe américaine a réalisé une expérience de gain de fonction sur une souche sauvage de SARS-like coronavirus (Bat-ScoV). Ils ont produit un ADNc qu'ils ont introduit par transfection dans des cellules en culture, obtenant un virus répliquatif qui pénètre difficilement dans les cellules humaines exprimant ACE2. Le domaine RBD de Spike a été alors remplacé par celui du SARS-CoV1. Le nouveau virus pénètre et se réplique alors facilement dans les cellules de primates et de souris [23]. En 2015 des expériences gain de fonction ont été réalisées par des équipes chinoises et américaines en utilisant le coronavirus « squelette » MA15 [24], qui est un virus SARS-CoV1 adapté par passages sur des souris BALB/c [25]. Ce virus MA15 est avirulent et incapable de pénétrer les cellules humaines. L'ajout au virus MA15 du gène codant la Spike d'un SARS-like coronavirus sauvage (SHC014-CoV) permet d'obtenir un nouveau virus capable de reconnaître le récepteur ACE2 humain. Ce virus a été transmis en série *in vitro* sur des cellules humaines de l'épithélium respiratoire, où il a atteint des titres proches de la souche épidémique de SARS-CoV1. Les passages *in vivo* chez la souris transgéniques ACE2, précédemment décrite [26], ont montré une importante répllication dans les poumons. Le virus chimérique a été ensuite entièrement synthétisé, confirmant sa forte répllication et sa virulence *in vivo*. Il n'était plus neutralisé par les anticorps anti-SARS-CoV1, ni protégé par les vaccins anti-SARS [24]. Les auteurs concluent qu'il existe un risque potentiel de réémergence du SARS-CoV1.

En 2014, il a été montré que la protéine Spike du MERS-CoV possède deux sites furine qui la sensibilisent aux protéases cellulaires [27]. En 2017, une équipe américaine a pratiqué des expériences de gain de fonction sur le MERS-CoV. Le virus sauvage a été propagé chez des souris transgéniques exprimant le récepteur humain DPP4 de ce virus, ce qui n'entraîne pas de maladie au départ, mais après 30 passages, le virus acquiert la propriété de croître 100 fois plus dans les poumons que le virus parental, entraînant une infection léthale des souris. L'analyse génétique des virus après passages montre l'acquisition de 13 à 22 mutations, dont plusieurs dans la protéine Spike du MERS [28]. En 2021, une équipe internationale a montré que la délétion du gène 5 de ce virus entraînait une hypervirulence d'un mutant du MERS-CoV. Ce gène stimule normalement la production d'interféron au cours de l'infection des souris transgéniques DPP4 [29]. Cela montre qu'un gain de fonction peut être obtenu par délétion d'un gène.

Récemment en 2023, une équipe américaine a réalisé une expérience de gain de fonction sur le variant Omicron du SARS-CoV2. Ce virus porteur de nombreuses mutations dans le RBD de Spike échappe à l'immunité humorale du vaccin et il est plus contagieux et moins virulent que le virus initial Wuhan-Hu-2. Le variant Omicron entraîne une infection modérée non mortelle pour les souris transgéniques ACE2. Le remplacement du gène spike du virus Omicron par celui du virus Wuhan-Hu-2 crée un virus chimérique qui entraîne chez la souris une infection sévère mortelle à 80 %. Une étude ultérieure a montré que la mutation de la protéine non structurale 6 (nsp6) en plus de la protéine S était suffisante pour récupérer le phénotype atténué d'Omicron [30].

## 2. La biologie synthétique

En 1965, la première séquence déchiffrée d'un polynucléotide naturel est un ARNt-Ala de levure (76 nt) [31], lequel a été ensuite synthétisé en 1970 sous forme d'ADN double brin [32]. Puis les chimistes ont mis au point la synthèse d'oligonucléotides en phase solide, ce qui a permis les premières synthèses nucléotidiques par segments de 40 à 80 nt [33]. Le premier gène fonctionnel ainsi synthétisé est un ADN de 207 pb codant l'ARNt suppresseur de tyrosine d'*Escherichia coli* [34]. De plus grands gènes ont ensuite été synthétisés, en assemblant de grands polynucléotides obtenus par ligature d'oligonucléotides se chevauchant [35, 36], un procédé fastidieux et coûteux. On a ainsi synthétisé des plasmides contenant le gène de 2 700 pb codant la  $\beta$ -lactamase [37] et le gène de 4 917 pb codant la protéine de surface du mérozoïte (MSP-1) de *Plasmodium falciparum* [38]. On assiste ensuite à des avancées considérables dans les techniques de séquençage et de synthèse nucléotidiques avec la mise en place de plates-formes à haut débit, réduisant considérablement les délais et les coûts [39]. On peut désormais synthétiser chimiquement des fragments de 8 à 30 kb, ce qui suffit

pour reconstituer la plupart des virus à ARN. Plusieurs génomes bactériens ont également pu être synthétisés. En 2008, on a réalisé la synthèse du génome de 582 970 pb de *Mycoplasma genitalium* [40], puis en 2010 celle de *Mycoplasma mycoides* de 1,08 Mb [41] et en 2017 la création de 9 souches de *Saccharomyces cerevisiae* dans lesquelles 1 ou 2 des 16 chromosomes ont été remplacés par de l'ADN synthétique [42]. Enfin la synthèse chimique complète d'*E. coli* a été réalisée en 2019 [43]. Désormais, on peut synthétiser et assembler le génome de n'importe quel micro-organisme à partir de la séquence *in silico*.

## 2.1. Synthèse des virus

Il a été possible très tôt de produire des virus par synthèse enzymatique à partir de génomes viraux, en utilisant une ADN polymérase pour les virus à ADN ou une transcriptase inverse pour les virus à ARN. En 1978 on a ainsi converti le génome du phage Q $\beta$  à ARN (4127 nt) en ADNc bicaténaire [44], incorporé dans un plasmide et transfecté dans *E. coli* pour produire des phages fonctionnels. Cette approche, qui permet des manipulations génétiques de l'ARN, a révolutionné la connaissance des virus à ARN [45]. Par la suite, l'ADNc d'un poliovirus inséré dans un plasmide a été introduit dans des cellules HeLa, permettant une production en très faible quantité de poliovirus infectieux [46]. Ainsi, des ADNc des virus à ARN peuvent être facilement préparés et utilisés pour régénérer soit des virus à brin positif (polarité de l'ARNm), soit des virus à brin négatif, selon les virus. Cette approche sera utilisée largement pour synthétiser de nombreux virus à RNA, notamment le virus de la rage [47], le virus respiratoire syncytial [48], le virus de la grippe A [49, 50], le virus de la rougeole [51], le virus Ebola [52], le bunyavirus (arbovirus) [53] et le rotavirus [54].

En 2002, on a réussi à séquencer le virus de la grippe espagnole à partir des séquences virales obtenues par transcriptase inverse à partir de multiples prélèvements humains datés de 1918. Le virus a été ensuite reconstitué en fragments réassemblés en ADNc correspondant aux 8 fragments du virus. Ces fragments ont été transférés dans des cellules MDCK de singe ou des cellules chorioallantoïdes d'embryons de poulet, coinfectées avec des virus H1N1 peu virulents [55, 56]. Le virus viable H1N1 de la grippe espagnole a ainsi été ressuscité et s'est avéré beaucoup plus virulent que les virus H1N1 contemporains [57]. D'autres virus à ARN ont été aussi reconstitués à partir de fragments viraux. En 2006, on a réussi la synthèse du génome entier (9 472 nt) d'un rétrovirus HERV à partir de vestiges rétroviraux endogènes insérés dans les chromosomes humains depuis plus d'un million d'années [58], puis d'un autre provirus infectieux HERV-K, désigné *Phoenix*, capable de se répliquer dans des cellules humaines en culture [59]. De même à partir des fragments d'ADN viral obtenus par RT-PCT sur des fèces de chimpanzés, on a réussi en 2007 à reconstituer *de novo* un virus infectieux de l'immunodéficience simienne (SIVcpz), qui est le virus le plus proche du VIH-1 [60]. Les génomes des grands virus à ARN, tels que les coronavirus (29 kb), sont difficiles à cloner et à manipuler dans *E. coli* en raison de leur taille et de l'instabilité de leur génome. Des fragments viraux sont générés à partir d'isolats viraux, d'ADN viral cloné, d'échantillons cliniques ou d'ADN synthétiques. Ces fragments sont ensuite réassemblés en une seule étape et clonés dans un chromosome artificiel de la levure *S. cerevisiae*. L'ARN polymérase T7 est ensuite utilisée pour générer le virus viable dans la levure [61].

On peut aussi synthétiser n'importe quel génome de virus par synthèse chimique *de novo* à partir de la séquence *in silico*. En 2000, la première structure répliquative ainsi synthétisée est un réplicon du virus de l'hépatite C dépourvu des gènes des protéines structurales [62]. En 2002, la première synthèse complète d'un virus à ARN (7 500 pb) d'un poliovirus a été réalisée en l'absence d'une matrice moléculaire [63]. L'ADNc obtenu a été transcrit *in vitro* en ARN viral infectieux après incubation dans un extrait de cellules HeLa. Conjointement, le génome complet d'un virus à ADN, le phage  $\Phi$ X174 (5 386 pb), a été synthétisé en deux semaines par synthèse chimique. L'ADN a été ensuite transfecté dans des bactéries d'*E. coli* qui ont produit des bactériophages viables [64].

## 2.2. Remodelage des génomes entiers

La conception de nouveaux micro-organismes passe par l'ingénierie et la modification de châssis microbiens synthétiques qui constituent l'un des meilleurs moyens de découvrir les principes fondamentaux de la

vie, permettant d'améliorer les applications dans de nombreux domaines, notamment la médecine et l'industrie. Le principe est inspiré de la pratique du « refactoring » utilisée en informatique pour améliorer les logiciels existants. Le but général est d'améliorer la structure interne d'un système existant pour une utilisation future, tout en maintenant les fonctions principales.

### 2.2.1. Recodage des génomes

À l'échelle du génome entier, on peut modifier le code génétique d'un virus ou d'une bactérie. Du fait de la dégénérescence de ce code, il existe plusieurs codons synonymes pour un même acide aminé, dont certains sont préférentiellement utilisés dans les génomes selon les micro-organismes. L'introduction d'un biais de codons peut freiner la production de protéines, du fait de l'utilisation de codons rares. Ce procédé a été utilisé pour atténuer la prolifération et la virulence des poliovirus [65]. Ceci peut aussi être réalisé de façon ciblée pour restreindre l'expression de certains gènes (dé-optimisation), sans modifier un seul acide aminé des protéines codées par ces gènes [66]. On a aussi eu recours au biais des paires de codons qui sont inégalement répartis selon les génomes considérés, certaines paires étant très sous-représentées, ce qui a des effets sur le taux de réplication virale [67]. On a pu ainsi synthétiser et recoder en 2019 le génome entier d'*E. coli* en utilisant 59 codons au lieu des 64 pour la souche sauvage [43].

Pour améliorer la performance d'un micro-organisme dans un but industriel, on a cherché à identifier les gènes non essentiels en modifiant les châssis microbiens. Cela permet de comprendre quels gènes sont essentiels pour les propriétés cellulaires et requis pour maintenir la vie cellulaire. En général, un certain nombre d'analyses informatiques ont été utilisées pour définir les gènes essentiels au maintien de la vie. Ces gènes sont généralement impliqués dans le métabolisme de base, le métabolisme de la paroi cellulaire, la division cellulaire et le métabolisme de l'ADN. Grâce à la biologie synthétique, on peut ainsi reconstruire à partir des génomes naturels bactériens ou viraux, de nouvelles entités biologiques qui n'existent pas dans la nature, principalement pour des applications médicales ou industrielles. Cela peut aider à concevoir des virus nouveaux en manipulant les séquences *in silico*, éventuellement avec l'aide de l'intelligence artificielle (IA).

Du fait de la dissémination de synthétiseurs de paillasse, cette approche par synthèse biologique doit être contrôlée, car elle peut constituer un réel danger. L'ingénierie et la modification de châssis microbiens synthétiques constituent l'un des meilleurs moyens non seulement de découvrir les principes fondamentaux de la vie, mais aussi d'améliorer les applications dans les domaines de la santé, de la médecine, de l'agriculture, de la médecine vétérinaire et de la santé publique. Mais cela donne aussi des perspectives d'utilisation malveillante dans le cas d'agents pathogènes.

### 2.2.2. Construction du châssis microbien

Il existe deux stratégies de construction de châssis microbiens, l'une descendante et l'autre ascendante [68]. La stratégie descendante consiste à réduire progressivement le génome du micro-organisme. Ce processus permet de déterminer les gènes indispensables à la survie de l'organisme en combinant l'analyse informatique des systèmes, les données expérimentales et les modèles tels que ceux des réseaux métaboliques, de régulation et de signalisation. Puis, un châssis est construit et modifié en supprimant les gènes non essentiels par diverses stratégies, notamment l'utilisation de procédures médiées par l'ADN et de recombinases spécifiques à un site, la mutagenèse par transposon ou le système CRISPR/Cas. Les exemples de cette approche concernent la réduction du génome de *E. coli* et de *Bacillus subtilis*. Cela a permis d'obtenir une souche de *E. coli* à croissance plus rapide que la souche parentale [69] et d'élucider les fondements génétiques des protéines sécrétées chez *B. subtilis* [70].

La stratégie ascendante a recours à la synthèse chimique à faible coût de segments d'ADN provenant de divers génomes qui seront fusionnés, ce qui permet d'élaborer des châssis polyvalents provenant de différentes entités biologiques. Les nouveaux micro-organismes ainsi construits et assemblés après synthèse sont transplantés dans des organismes (cellules, bactéries, etc.) pour créer et produire un génome microbien entièrement synthétisé. Des entités génétiques peuvent facilement être conçues par biologie synthétique avec l'aide de l'informatique. Par exemple en 2016, on a réussi à synthétiser une bactérie viable dérivée de

*M. mycoides* avec un génome bactérien minimal synthétisé chimiquement à partir d'une séquence *in silico* de 531 kb au lieu des 1079 kb de la souche sauvage [71]. Cette stratégie a aussi été utilisée en 2019 pour synthétiser la bactérie *Caulobacter crescentus* [72]. On peut ainsi construire des micro-organismes personnalisés pour la production de produits à visée industrielle. Il est possible aussi de reformater des virus pour améliorer leurs performances à partir de génomes connus. Cela a été fait en 2005 avec le bactériophage T7, qui a été séparé physiquement en segments génomiques qui ont été ensuite réassemblés en un virus ergonomisé aux performances améliorées. Le virus chimérique est viable avec toutes les propriétés du virus sauvage, mais beaucoup plus facile à manipuler [73]. Ainsi les génomes viraux peuvent être facilement reconstruits pour obtenir une meilleure performance. En 2018, les chercheurs ont réussi à synthétiser *in vitro* le virus *horsepox* (212 kb) de la variole équine en fusionnant 10 grands fragments synthétisés d'ADN de 10-30 kb provenant de la séquence virale *in silico* [74, 75]. L'ADN ainsi obtenu est introduit dans des cellules infectées par un poxvirus apparenté, le virus du fibrome de Shope. Cela permet la production du virus vivant en culture cellulaire. Ce virus est moins virulent pour les souris que celui de la vaccine et induit une réponse protectrice contre lui. C'est la première synthèse complète d'un poxvirus par biologie synthétique. Il s'agit de la première synthèse complète d'un poxvirus utilisant la biologie synthétique. Cette publication est un exploit technologique discutable et fortement critiqué, car elle révèle la stratégie pour construire au laboratoire un virus très dangereux, le virus *smallpox* de la variole [76, 77].

### 3. Intelligence artificielle.

La capacité de l'IA à traiter de grandes quantités de données brutes et non structurées (séquences d'ADN, protéines, etc.) permet de réduire le temps et le coût de certaines expériences, d'en réaliser d'autres non envisageables auparavant et de contribuer au domaine plus large de l'ingénierie génétique [78]. Une des applications de l'IA est la *machine learning* qui permet d'extraire des connaissances à partir de données et d'en tirer des enseignements de manière autonome. En utilisant des algorithmes pour analyser de grandes quantités de données, cette approche permet à une machine d'apprendre et de s'améliorer automatiquement. On peut accroître la capacité humaine de modifier du matériel génétique pour obtenir des fonctions spécifiques. Cela induit des préoccupations concernant les utilisations potentielles sur la biosécurité [79].

Les applications aux systèmes biologiques, notamment le *machine Learning*, permettent de modifier et de programmer des systèmes génétiques avec des fonctions nouvelles. Les principes d'ingénierie et l'utilisation d'outils de conception systématique sont utilisés pour reprogrammer de systèmes cellulaires. L'utilisation du *machine learning* a conduit à des progrès étonnants dans la conception computationnelle de protéines, permettant des applications industrielles et biomédicales [80]. Des protéines synthétiques sont ainsi conçues pour réaliser des fonctions cellulaires. On peut dès lors synthétiser les gènes correspondants et les insérer dans le génome microbien. Cela peut être appliqué aux facteurs de virulence, tels que les toxines ou les adhésines des micro-organismes pathogènes. On peut par exemple concevoir de nouvelles toxines à partir de modèles moléculaires, et même créer des armes redoutables. On peut citer l'exemple récent de l'entreprise *Collaborations Pharmaceuticals* qui utilise un système IA baptisée *MegaSyn*, entraîné avec des pesticides, des toxines environnementales et des médicaments, pour tenter de trouver de nouveaux médicaments. Au lieu de chercher les molécules ayant la plus faible toxicité possible, les chercheurs ont fait la requête de chercher la toxicité et la bioactivité les plus élevées. En moins de 6 heures, l'IA a généré plus de 40 000 molécules neurotoxiques, tous plus toxiques les unes que les autres. Elle a notamment découvert le VX, un des agents innervants les plus dangereux, inventé en 1952 et dix fois plus mortel que le sarin. Elle a également découvert d'autres armes chimiques connues, mais aussi de nombreuses molécules encore inconnues et potentiellement plus toxiques que le VX [81]. Les chercheurs ont dit qu'ils n'avaient jamais pensé auparavant à la façon dont leur outil pourrait être détourné, et ont été très surpris par les résultats. Avec le développement de l'IA, la conception d'armes encore inimaginables aujourd'hui pourrait bientôt être à la portée de tous.



#### 4. Conclusion

Les progrès des biotechnologies ouvrent une nouvelle ère, celle de progrès scientifiques sans précédent sur les connaissances du vivant, mais aussi celle d'une biologie dangereuse pour laquelle il faut être très vigilant. On peut désormais facilement modifier le génome des micro-organismes (perte ou gain-de-fonction) et synthétiser des séquences de gènes et même des micro-organismes entiers (bactéries et virus), ce qui permet de grandes avancées dans de nombreux domaines de la science. Toutes ces manipulations peuvent être aidées par l'utilisation de l'IA pour concevoir les manipulations ou des synthèses de génomes *de novo*. On peut citer notamment une meilleure compréhension de l'évolution et des propriétés d'agents pathogènes dangereux. Cependant cela a des implications non seulement en termes de double usage et aussi de disponibilité d'agents infectieux très dangereux qui peuvent être synthétisés à partir des séquences *in silico* provenant de banques de données. En outre, la synthèse du génome des virus à ADN et à ARN offre des possibilités sans précédent de modification des génomes naturels, permettant ainsi de créer de nouveaux micro-organismes infectieux potentiellement dangereux. Ainsi, l'Humanité pourrait être confrontée à la menace d'agents pathogènes bien pires que tout ce que la nature pourrait créer. Le prochain grand défi sera de concilier les progrès scientifiques avec la biosécurité et la biosûreté.

#### 5. Déclaration d'intérêt

L'auteur ne travaille pas, ne conseille pas, ne possède pas d'actions ou ne reçoit pas de fonds de toute organisation qui pourrait bénéficier de cet article, et n'a déclaré aucune affiliation autre que son institution de recherche.

#### Références

- [1] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, « A programmable Dual-RNA-guided DNA endonuclease in adaptive bacterial immunity », *Science* **337** (2012), p. 816-821.
- [2] NCBI Virus, « Viral genomes », Online at <https://www.ncbi.nlm.nih.gov/genome/viruses/> (accessed August 9, 2024).
- [3] J. Holland, K. Spindler, F. Horodyski, E. Grabau, S. Nichol, S. VandePol, « Rapid evolution of RNA genomes », *Science* **215** (1982), p. 1577-1585.
- [4] M. Eigen, « Viral quasispecies », *Sci. Am.* **269** (1993), p. 42-49.
- [5] C. Tu, G. Cramer, X. Kong *et al.*, « Antibodies to SARS coronavirus in civets », *Emerg. Infect. Dis.* **10** (2004), p. 2244-2248.
- [6] J. H. Ellwanger, J. A. Bogo Chies, « Zoonotic spillover : understanding basic aspects for better prevention zoonotic spillover », *Genet. Mol. Biol.* **44** (2021), n° S1, article no. e20200355.
- [7] B. L. Pike, K. E. Saylor, J. N. Fair *et al.*, « The origin and prevention of pandemics », *Clin. Infect. Dis.* **50** (2010), p. 1636-1640.
- [8] R. K. Plowright, C. R. Parrish, H. McCallum *et al.*, « Pathways to zoonotic spillover », *Nature Rev. Microbiol.* **15** (2017), p. 502-510.
- [9] Z. W. Ye, S. Yuan, K. S. Yuen *et al.*, « Zoonotic origins of human coronaviruses », *Int. J. Bio. Sci.* **16** (2021), p. 1686-1697.
- [10] R. J. Jackson, A. J. Ramsay, C. D. Christensen, S. Beaton, D. F. Hall, I. A. Ramshaw, « Expression of mouse interleukin-4 by a recombinant ectromelia virus suppresses cytolytic lymphocyte responses and overcomes genetic resistance to mousepox », *J. Virol.* **75** (2001), p. 1205-1210.
- [11] L. D. Sims, T. M. Ellis, K. K. Liu *et al.*, « Avian influenza in Hong Kong 1997–2002 », *Avian Dis.* **47** (2003), p. 832-838.
- [12] K. Ungchusak, P. Auewarakul, S. F. Dowell *et al.*, « Probable person-to-person transmission of avian influenza A (H5N1) », *N. Engl. J. Med.* **352** (2005), p. 333-340.

- [13] OMS, « Bulletins d'information sur les flambées épidémiques Grippe aviaire A(H5N1)—Viet Nam », 2 avril 2024. Online at <https://www.who.int/fr/emergencies/disease-outbreak-news/item/2024-DON511> (accessed August 9, 2024).
- [14] M. Imai, T. Watanabe, M. Hatta *et al.*, « Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets », *Nature* **486** (2012), p. 420-428.
- [15] S. Herfst, E. J. A. Schrauwen, M. Linster *et al.*, « Airborne transmission of influenza A/H5N1 virus between ferrets », *Science* **336** (2012), p. 1534-1541.
- [16] Y. Zhang, Q. Zhang, H. Kong *et al.*, « H5N1 hybrid viruses bearing 2009/H1N1 virus genes transmit in guinea pigs by respiratory droplet », *Science* **340** (2013), p. 1459-1463.
- [17] M. Agüero, I. Monne, A. Sánche *et al.*, « Highly pathogenic avian influenza A(H5N1) virus infection in farmed minks, Spain, October 2022 », *Euro. Surveill.* **28** (2023), n° 3, article no. 2300001.
- [18] Y. Buisson, « Covid-19, an unfinished story », *Presse. Med. (QMR)* **51** (2022), n° 3, article no. 104131.
- [19] P. Berche, « Gain-of-function and origin of Covid19 », *Presse. Med. (QMR)* **51** (2022), n° 3, article no. 104167.
- [20] W. Li, Z. Shi, M. Yu *et al.*, « Bats are natural reservoirs of SARS-like coronaviruses », *Science* **310** (2005), p. 676-679.
- [21] X. Y. Ge, J. L. Li, X. L. Yang *et al.*, « Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor », *Nature* **503** (2013), p. 535-538.
- [22] B. Hu, L. P. Zeng, X. L. Yang *et al.*, « Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus », *PLoS Pathog.* **13** (2017), n° 11, article no. e1006698.
- [23] M. M. Becker, R. L. Graham, E. F. Donaldson *et al.*, « Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice », *Proc. Natl. Acad. Sci. USA* **105** (2008), p. 19944-19449.
- [24] V. D. Menachery, B. L. Yount, K. Debbink *et al.*, « A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence », *Nat. Med.* **21** (2015), p. 1508-1513.
- [25] A. Roberts, D. Deming, C. D. Paddock *et al.*, « A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice », *PLoS Pathog.* **3** (2007), n° 1, article no. e5.
- [26] P. B. McCray, L. Pewe, C. Wohlford-Lenane *et al.*, « Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus », *J. Virol.* **81** (2007), p. 813-821.
- [27] J. K. Millet, G. R. Whittaker, « Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein », *Proc. Natl. Acad. Sci. USA* **111** (2014), p. 15214-15219.
- [28] K. Li, C. L. Wohlford-Lenane, R. Channappanavar *et al.*, « Mouse-adapted MERS coronavirus causes lethal lung disease in human DPP4 knockin mice », *Proc. Natl. Acad. Sci. USA* **114** (2017), p. 3119-3128.
- [29] J. Gutierrez-Alvarez, L. Wang, K. L. Fernandez-Delgado *et al.*, « Middle east respiratory syndrome coronavirus gene 5 modulates pathogenesis in mice », *J. Virol.* **95** (2021), n° 3, article no. e01172-20.
- [30] D.-Y. Chen, C. V. Chin, D. Kenney *et al.*, « Spike and nsp6 are key determinants of SARS-CoV-2 Omicron BA.1 attenuation », *Nature* **615** (2023), p. 143-150.
- [31] R. W. Holley, J. Apgar, G. A. Everett *et al.*, « Structure of a ribonucleic acid », *Science* **147** (1965), p. 1462-1465.
- [32] K. L. Agarwal, H. Büchi, M. H. Caruthers *et al.*, « Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast », *Nature* **227** (1971), p. 27-34.
- [33] M. H. Caruthers, « Gene synthesis machines : DNA chemistry and its uses », *Science* **230** (1985), p. 281-285.
- [34] H. G. Khorana, K. L. Agarwal, P. Besmer *et al.*, « Total synthesis of the structural gene for the precursor of a tyrosine suppressor transfer RNA from *Escherichia coli*. 1. General introduction », *J. Biol. Chem.* **251** (1978), p. 565-570.
- [35] L. Stewart, A. B. Burgin, « Whole gene synthesis : a gene-o-matic future », *Front. Drug Des. Disc.* **1** (2005), p. 297-341.
- [36] Y. Sanghvi, « A roadmap to the assembly of synthetic DNA from raw materials », in *Working Papers for Synthetic Genomics : Risks and Benefits for Science and Society* (M. S. Garfinkel, D. Endy, G. L. Epstein, R. M. Friedman, éds.), Craig Venter Institute, Rockville, Md, 2007, p. 17-33.
- [37] W. P. Stemmer, A. Cramer, K. D. Ha *et al.*, « Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides », *Gene* **164** (1995), p. 49-53.

- [38] W. Pan, E. Ravot, R. Tolle *et al.*, « Vaccine candidate MSP-1 from *Plasmodium falciparum* : a redesigned 4917 bp polynucleotide enables synthesis and isolation of full-length protein from *E. coli* and mammalian cells », *Nucleic Acids Res.* **27** (1999), p. 1094-1103.
- [39] J. A. Shendure, G. J. Porreca, G. M. Church, « Overview of DNA sequencing strategies », in *Current Protocols in Molecular Biology* (F. M. Ausubel *et al.*, éd.), John Wiley & Sons, Hoboken, NJ, 2008.
- [40] D. G. Gibson, E. Ravot, R. Tolle *et al.*, « Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome », *Science* **319** (2008), p. 1215-1220.
- [41] D. G. Gibson, J. I. Glass, C. Lartigue *et al.*, « Creation of a bacterial cell controlled by a chemically synthesized genome », *Science* **329** (2010), p. 52-56.
- [42] S. M. Richardson, L. A. Mitchell, G. Stracquadanio *et al.*, « Design of a synthetic yeast genome », *Science* **355** (2017), p. 1040-1044.
- [43] J. Fredens, K. Wang, D. de la Torre *et al.*, « Total synthesis of *Escherichia coli* with a recoded genome », *Nature* **569** (2019), p. 514-518.
- [44] T. Taniguchi, M. Palmieri, C. Weissmann, « Q $\beta$  DNA-containing hybrid plasmids giving rise to Q $\beta$  phage formation in the bacterial host », *Nature* **274** (1978), p. 223-228.
- [45] C. Weissmann, H. Weber, T. Taniguchi *et al.*, « Reversed genetics : a new approach to the elucidation of structure–function relationship », *Ciba. Found. Symp.* **66** (1979), p. 47-61.
- [46] V. R. Racaniello, D. Baltimore, « Cloned poliovirus complementary DNA is infectious in mammalian cells », *Science* **214** (1981), p. 916-919.
- [47] M. J. Schnell, T. Mebatsion, K. K. Conzelmann, « Infectious rabies viruses from cloned cDNA », *EMBO J.* **13** (1994), p. 4195-4203.
- [48] P. L. Collins, M. G. Hill, E. Camargo *et al.*, « Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5' proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development », *Proc. Natl. Acad. Sci. USA* **92** (1995), p. 11563-11567.
- [49] G. Neumann, T. Watanabe, H. Ito *et al.*, « Generation of influenza A viruses entirely from cloned cDNAs », *Proc. Natl. Acad. Sci. USA* **96** (1999), p. 9345-9350.
- [50] E. Fodor, L. Devenish, O. G. Engelhardt, P. Palese, G. G. Brownlee, A. García-Sastre, « Rescue of influenza A virus from recombinant DNA », *J. Virol.* **73** (1999), p. 9679-9682.
- [51] K. Takeuchi, M. Takeda, N. Miyajima, « Toward understanding the pathogenicity of wild-type measles virus by reverse genetics », *Jpn. J. Infect. Dis.* **55** (2002), p. 143-149.
- [52] G. Neumann, H. Feldmann, S. Watanabe *et al.*, « Reverse genetics demonstrates that proteolytic processing of the Ebola virus glycoprotein is not essential for replication in cell culture », *J. Virol.* **76** (2002), p. 406-410.
- [53] A. K. Överby, V. Popov, E. P. Neve, R. F. Pettersson, « Generation and analysis of infectious virus-like particles of uukuniemi virus (bunyaviridae) : a useful system for studying bunyaviral packaging and budding », *J. Virol.* **80** (2006), p. 10428-10435.
- [54] S. Komoto, J. Sasaki, K. Taniguchi, « Reverse genetics system for introduction of site-specific mutations into the double-stranded RNA genome of infectious rotavirus », *Proc. Natl. Acad. Sci. USA* **103** (2006), p. 4646-4651.
- [55] A. H. Reid, T. G. Fanning, T. A. Janczewski *et al.*, « Characterization of the 1918 Spanish influenza virus matrix gene segment », *J. Virol.* **76** (2002), p. 10717-10723.
- [56] J. K. Taubenberger, A. H. Reid, R. M. Lourens *et al.*, « Characterization of the 1918 influenza virus polymerase genes », *Nature* **437** (2005), p. 889-893.
- [57] T. M. Tumpey, C. F. Basler, P. V. Aguilar *et al.*, « Characterization of the reconstructed 1918 Spanish influenza pandemic virus », *Science* **310** (2005), p. 77-80.
- [58] Y. N. Lee, P. D. Bieniasz, « Reconstitution of an infectious human endogenous retrovirus », *PLoS Pathog.* **3** (2007), article no. e10.
- [59] M. Dewannieux, F. Harper, A. Richaud *et al.*, « Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements », *Genome Res.* **16** (2006), p. 1548-1556.

- [60] J. Takehisa, M. H. Kraus, J. M. Decker *et al.*, « Generation of infectious molecular clones of simian immunodeficiency virus from fecal consensus sequences of wild chimpanzees », *J. Virol.* **81** (2007), p. 7463-7475.
- [61] T. T. N. Thao, F. Labroussaa, N. Ebert *et al.*, « Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform », *Nature* **582** (2020), p. 561-565.
- [62] K. J. Blight, A. A. Kolykhalov, C. M. Rice, « Efficient initiation of HCVRNA replication in cell culture », *Science* **290** (2000), p. 1972-1974.
- [63] J. Cello, A. V. Paul, E. Wimmer, « Chemical synthesis of poliovirus cDNA : generation of infectious virus in the absence of natural template », *Science* **297** (2002), p. 1016-1018.
- [64] H. O. Smith, C. A. Hutchison, C. Pfannkoch, J. C. Venter, « Generating a synthetic genome by whole genome assembly :  $\Phi$ X174 bacteriophage from synthetic oligonucleotides », *Proc. Natl. Acad. Sci. USA* **100** (2003), p. 15440-15445.
- [65] C. C. Burns, J. Shaw, R. Campagnoli *et al.*, « Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region », *J. Virol.* **80** (2006), p. 3259-3272.
- [66] S. Mueller, D. Papamichail, J. R. Coleman *et al.*, « Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity », *J. Virol.* **80** (2006), p. 9687-9696.
- [67] J. R. Coleman, D. Papamichail, S. Skiena *et al.*, « Virus attenuation by genome-scale changes in codon pair bias », *Science* **320** (2008), p. 1784-1787.
- [68] H. Chia, X. Wanga, Y. Shaoa *et al.*, « Engineering and modification of microbial chassis for systems and synthetic biology », *Synth. Syst. Biotechnol.* **4** (2019), p. 25-33.
- [69] H. Mizoguchi, H. Mori, T. Fujio, « *Escherichia coli* minimum genome factory », *Biotechnol. Appl. Biochem.* **46** (2007), p. 157-167.
- [70] H. Tjalsma, H. Antelmann, J. D. Jongbloed *et al.*, « Proteomics of protein secretion by *Bacillus subtilis* : separating the 'secrets' of the secretome », *Microbiol. Mol. Biol. Rev.* **68** (2004), p. 207-233.
- [71] C. A. Hutchison, R. Y. Chuang, V. N. Noskov *et al.*, « Design and synthesis of a minimal bacterial genome », *Science* **351** (2016), n° 6280, article no. aad6253.
- [72] J. E. Venetz, L. Del Medico, A. Wölfle *et al.*, « Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality », *Proc. Natl. Acad. Sci. USA* **116** (2019), p. 8070-8079.
- [73] L. Y. Chan, S. Kosuri, D. Endy, « Refactoring bacteriophage T7 », *Mol. Syst. Biol.* **1** (2005), article no. 2005.0018.
- [74] E. R. Tulman, G. Delhon, C. L. Afonso *et al.*, « Genome of horsepox virus », *J. Virol.* **80** (2006), p. 9244-9258.
- [75] R. S. Noyce, S. Lederman, D. H. Evans, « Construction of an infectious horsepox virus vaccine from chemically synthesized DNA fragments », *PLOS One* **13** (2018), article no. e0188453.
- [76] G. D. Koblentz, « The *de novo* synthesis of horsepox virus : implications for biosecurity and recommendations for preventing the reemergence of smallpox », *Health Secur.* **15** (2017), n° 6, p. 620-628.
- [77] T. Inglesby, « Horsepox and the need for a new norm, more transparency, and stronger oversight for experiments that pose pandemic risks », *PLOS Pathog.* **14** (2018), article no. e1007129.
- [78] A. Bhardwaj, S. Kishore, D. K. Pandey, « Artificial intelligence in biological sciences », *Life* **12** (2022), article no. 1430.
- [79] T. Kuiken, « Artificial intelligence in the biological sciences : uses, safety, security, and oversight, congressional research service », R47949, Washington, Nov. 22, 2023. Online at <https://crsreports.congress.gov/product/pdf/R/R47849> (accessed August 9, 2024).
- [80] S. Kohyama, B. P. Frohn, L. Babl, P. Schwillie, « Machine learning-aided design and screening of an emergent protein function in synthetic cells », *Nat. Commun.* **15** (2024), article no. 2010.
- [81] F. Urbina, F. Lentzos, C. Invernizzi, S. Ekins, « Dual use of artificial intelligence-powered drug discovery », *Nat. Mach. Intell.* **4** (2022), n° 3, p. 189-191.