

Nos travaux sur *M. genitalium* ne sont que la première des nombreuses étapes à franchir pour modéliser les cellules ou tissus humains à l'échelle des gènes et des molécules. Notre modèle est loin d'être parfait, et les mycoplasmes sont l'expression la plus simple de la vie autosuffisante. Nos simulations, le code source, les données expérimentales, le code de visualisation, la base de connaissances, tout est disponible gratuitement sur Internet. Avec d'autres chercheurs, nous travaillons déjà à améliorer le simulateur et à l'étendre à divers organismes, tels que la bactérie *E. coli* et la levure *Saccharomyces cerevisiae*, toutes deux des organismes utilisés fréquemment tant dans les laboratoires que dans l'industrie.

Chez ces espèces, la régulation des gènes est beaucoup plus complexe, et la localisation des événements dans la cellule joue un rôle bien plus important (la levure, par exemple, présente un noyau). Autant de problèmes à résoudre avant de s'attaquer à des cellules plus complexes encore, tels des macrophages de souris ou

humains. Faciles à cultiver, ces cellules du système immunitaire seront sans doute les premières utilisées pour ajuster et valider le modèle.

Il est trop tôt pour estimer le temps que ces améliorations prendront. Comparées aux bactéries, les cellules humaines comportent beaucoup plus de compartiments. De plus, leur contrôle génétique, plus complexe que celui des mycoplasmes, demeure en grande partie mystérieux. Enfin, de par leurs rôles complémentaires dans les tissus multicellulaires, les cellules humaines interagissent plus avec d'autres types cellulaires que les bactéries.

Début 2008, modéliser la plus simple des cellules ne semblait pas pouvoir se concrétiser avant une décennie, et modéliser une forme de vie plus complexe paraissait inenvisageable. Aujourd'hui, l'idée de simuler une cellule humaine ne paraît plus si saugrenue. Et si le logiciel se révèle inapproprié, au moins cet échec mettra-t-il en lumière les nombreuses inconnues à expliciter encore sur nos propres cellules. ■

■ BIBLIOGRAPHIE

E. Gonçalves *et al.*, **Bridging the layers: toward the integration of signal transduction, regulation and metabolism into mathematical models**, *Molecular Biosystems*, vol. 9, pp. 1576-1583, 2013.

J. R. Karr *et al.*, **A whole-cell computational model predicts phenotype from genotype**, *Cell*, vol. 150, pp. 389-401, 2012.

P. L. Freddolino et S. Tavazoie, **The dawn of virtual cell biology**, *Cell*, vol. 150, pp. 248-250, 2012.

W. Gibbs, **La biologie virtuelle**, *Pour la Science* n°288, pp. 38-44, 2001.

france culture

LA MARCHÉ DES SCIENCES

Découvertes, inventions, aventures savantes au fil de l'Histoire

Aurélie Luneau
14h-15h / chaque jeudi

franceculture.fr

POUR LA SCIENCE



La biologie de vers une ingénierie

François KÉPÈS

Concevoir un génome *ab initio*, inventer une cellule minimale, produire de nouvelles fonctions biologiques... Telles sont quelques-unes des ambitions de la biologie de synthèse. Son objectif : ouvrir de nouveaux horizons en biologie en s'inspirant des méthodes de l'ingénierie.

Au début des années 1960, l'invention des circuits intégrés a métamorphosé l'ingénierie électronique : un schéma imprimé sur un masque suffisait à déposer sélectivement des couches métalliques et chimiques sur un support, et à produire ainsi le circuit imaginé. Bientôt, les ingénieurs ont puisé dans les circuits déjà conçus pour les combiner en des montages toujours plus complexes. La modélisation informatique aidant, les progrès furent considérables, conduisant aux nombreux circuits intégrés que comportent les téléphones portables, ordinateurs et autres dispositifs électroniques.

Et si une telle combinaison de technique et de méthode était transposable en biologie ? S'il était possible, en biologie aussi, de disposer d'une banque d'outils et de

règles rationnelles qui, combinés selon les besoins, aideraient à comprendre le vivant en le fabriquant, ou à concevoir et produire à la demande de nouvelles fonctions biologiques n'existant pas dans la nature ? Depuis quelques années, cette idée fait son chemin parmi les biologistes. Rassemblées sous le nom de biologie de synthèse, leurs approches se construisent aux interfaces de la biologie et de l'ingénierie, avec des apports de la chimie, de la physique, des mathématiques et de l'informatique. Elles ont pour objectif d'injecter en biologie les principes fondant toute ingénierie. De quoi s'agit-il ? Quelles sont les ambitions de la biologie de synthèse ? Que permet-elle aujourd'hui ? Quelles sont ses perspectives et limites ? Telles sont les questions que nous examinerons ici.

La biologie de synthèse est à la biologie des systèmes – l'analyse des interactions des biomolécules et des comportements des systèmes biologiques – ce que la chimie de synthèse est à la chimie analytique (l'analyse des substances chimiques). Après le triomphe de la chimie analytique dans les années 1850, les chimistes se sont tournés vers la synthèse pour obtenir plus de produits naturels à analyser, puis des composés n'existant pas dans la nature. Cette recherche a donné naissance à l'industrie chimique et à la pétrochimie du XX^e siècle. De même, dans les années 1990, les progrès de la biologie analytique ont été considérables grâce au développement des techniques dites « -omiques » : le séquençage du génome (génomique), la cartographie des ARN (transcriptomique) ou des protéines (protéomique) d'un système biologique, ou des modifications non génétiques du génome (épigénomique)... Stimulée par un volume croissant de données, la biologie des systèmes s'est développée, attirant physiciens et informaticiens. Quelques années plus tard est apparue la biologie de synthèse, qui s'appuie sur la compréhension ainsi acquise.

La biologie de synthèse est l'ingénierie rationnelle de la biologie. En d'autres termes,



synthèse du vivant

elle vise à la conception rationnelle et à l'ingénierie de systèmes complexes fondés sur le vivant ou inspirés par le vivant, et dotés de fonctions absentes dans la nature. Un exemple de cette démarche est la synthèse par voie biologique de l'artémisinine, médicament antipaludéen de référence dont l'industrialisation a débuté en 2013.

Un médicament contre le paludisme

L'artémisinine est naturellement produite par une plante, l'armoise annuelle. Son usage dans le traitement du paludisme remonte à plus de deux millénaires en Chine et, depuis 2002, l'Organisation mondiale de la santé recommande son utilisation (combinée à d'autres médicaments pour prévenir le développement de résistances du parasite responsable de la maladie). L'artémisinine fait partie d'une classe de molécules naturelles produites par des plantes, les terpénoïdes, utilisées dans des secteurs aussi variés que l'alimentation (arômes), les cosmétiques (fragrances) ou certaines thérapies anticancéreuses (outre la lutte contre le paludisme). Le menthol et le camphre en sont deux exemples.

L'ESSENTIEL

■ Apparue dans les années 2000, la biologie de synthèse vise à rationaliser les pratiques en biotechnologie.

■ Une fonction biologique peut être vue comme un circuit génétique constitué de plusieurs modules – des gènes produisant des fonctions de base.

■ En modélisant et combinant ces modules, les biologistes espèrent concevoir et fabriquer des fonctions biologiques à la demande.

■ Cette approche aiderait à mieux comprendre le vivant. Ses applications potentielles sont multiples.

© Kenn Brown/Mandolitic Studios

Au début des années 2000, Jay Keasling, de l'Université de Californie à Berkeley aux États-Unis, et ses collègues ont cherché des méthodes alternatives de fabrication de ces molécules, dont la production naturelle était trop faible et fluctuante en qualité et en quantité. Ils se sont notamment concentrés sur l'artémisinine. La synthèse chimique demandant un grand nombre d'étapes et n'offrant elle aussi qu'un faible rendement, ils se sont tournés vers la synthèse par voie biologique, ou bioproduction. En 2013, après une dizaine d'années d'efforts, leurs travaux ont abouti à la bioproduction, par la société pharmaceutique *Sanofi*, d'environ 60 tonnes annuelles d'artémisinine de synthèse aux caractéristiques stables.

Pour parvenir à cette bioproduction, ils ont combiné plusieurs éléments : un micro-organisme connu, à croissance rapide et aisé à modifier génétiquement, des gènes issus de l'armoise annuelle et des techniques de biologie moléculaire pour optimiser l'expression des gènes. Après quelques tâtonnements, le choix s'est porté sur la levure du boulanger. Cette cellule a l'intérêt, comme d'autres cellules eucaryotes (c'est-à-dire à noyau) non végétales, de produire un précurseur de l'artémisinine. En lui ajoutant les gènes de plusieurs enzymes de l'armoise annuelle impliquées dans la production de l'artémisinine, les biologistes ont réussi à construire, dans la levure, une voie de synthèse complète d'un précurseur de l'artémisinine, l'acide artémisinique. Voie qu'ils ont ensuite améliorée en flanquant chaque gène d'une séquence (un promoteur) favorisant son expression et en effectuant quelques autres modifications raisonnées de la levure. Le rendement est ainsi passé de 1,6 à 28 grammes par litre de fermenteur.

Cette concentration élevée entraîne une précipitation du produit, une caractéristique qui permet d'extraire l'acide artémisinique durant la fermentation. Celui-ci est ensuite transformé en artémisinine par voie chimique (chimie de synthèse optimisée par photochimie).

Cet exemple illustre le principe de la biologie de synthèse : concevoir de nouveaux systèmes biologiques en combinant des biomolécules, des circuits biochimiques servant un métabolisme ou la régulation d'une fonction biologique, et des cellules ou des particules qui abriteront l'ensemble. Trois démarches la caractérisent. La première est une approche constructiviste et hiérarchique :

on part de biobriques, des segments d'ADN remplissant des fonctions définies, et on les assemble dans des dispositifs – des circuits de régulation ou métaboliques –, eux-mêmes insérés dans des cellules qui produiront les fonctions. La bioproduction d'artémisinine en est un exemple.

La deuxième démarche est, à l'inverse, déconstructiviste: on part d'un micro-organisme fonctionnel et on remplace une partie de son génome par un bout de génome de synthèse qui produit la molécule ou la fonction souhaitée. Les biologistes espèrent, à terme, réussir l'opération avec le génome entier, nous y reviendrons. La troisième démarche, enfin, consiste à concevoir et construire des systèmes vivants simplifiés, tels que des protocellules – des vésicules aqueuses fermées par une enveloppe lipidique. Des équipes cherchent aussi à synthétiser de nouveaux nucléotides (les constituants de l'ADN) et acides aminés (les constituants des protéines), jamais observés dans la nature.

Bien que ces approches soient très différentes, elles ont en commun une même volonté: fonder ces recherches sur les prin-

cipes de l'ingénierie, afin de les optimiser, avec l'idée que cela facilitera l'émergence de nouvelles idées et la compréhension du vivant. Qu'est-ce que cela signifie? Les ingénieries de systèmes électroniques ou mécaniques nécessitent des cadres bien établis pour gérer la complexité, des outils fiables pour manipuler les états du système, et des plateformes de tests. La biotechnologie, en revanche, est encore dépourvue de tels cadres, outils et plateformes. Pour mesurer la distance à parcourir, reprenons une à une les propriétés qui caractérisent toute ingénierie. Le projet immédiat de la biologie de synthèse est de raccourcir cette distance.

Une propriété fondamentale de l'ingénierie est le découplage de la conception et de la fabrication: on commence par concevoir les plans d'une voiture ou d'un circuit intégré avant de les fabriquer. Actuellement, pour développer des composants ou circuits biochimiques opérationnels, il faut suivre des protocoles compliqués de construction de fragments d'ADN, user de modèles simples pour guider la conception, et procéder par approxima-

■ L'AUTEUR



François KÉPÈS est directeur de recherche à l'institut de Biologie des Systèmes

et de Synthèse (ISSB, Genopole, UEVE, CNRS) et directeur du Programme d'épigénétique à Genopole, à Évry. Il est professeur invité permanent au Collège impérial de Londres.

LA BOÎTE À OUTILS DE LA BIOLOGIE DE SYNTHÈSE

Si, depuis quelques années, il est possible de développer l'ingénierie rationnelle des systèmes biologiques, c'est grâce aux techniques visant à modifier les génomes. Dans le passé, ces méthodes étaient limitées à l'insertion ou à la mutation de gènes soit en des sites aléatoires, soit en des points précis du génome. Les quelques méthodes disponibles pour

diriger les modifications de génomes avaient plusieurs défauts, en particulier une faible efficacité et un coût élevé en main-d'œuvre. L'essor de la biologie de synthèse s'appuie sur plusieurs progrès récents qui ont considérablement accru les capacités d'intervention dirigée et de combinaison des modifications génétiques. Voici les principaux outils utilisés aujourd'hui.



LES RECOMBINASES sont des enzymes qui permettent d'insérer de longs fragments d'ADN dans le génome en des sites spécifiques.



CERTAINS INTRONS DE GROUPE II sont des éléments génétiques « égoïstes » qui s'incorporent eux-mêmes dans l'ADN hôte *via* une molécule intermédiaire d'ARN.



LES NUCLÉASES à doigts de zinc ou à domaines TAL sont des protéines dont la séquence peut être manipulée afin de cibler une séquence particulière d'ADN. Cette activité de fixation spécifique peut être associée à une autre, telle que le clivage de l'ADN ou l'interaction avec une autre protéine.



L'INGÉNIERIE RECOMBINOGÈNE fait appel aux propriétés des virus bactériens [les phages] pour effectuer des recombinaisons ciblées du génome, c'est-à-dire y insérer des gènes en des endroits précis.



LES NUCLÉASES GUIDÉES par ARN du type CRISPR offrent plus de flexibilité que les précédentes. Les CRISPR naturels sont des systèmes immunitaires bactériens dégradant l'ADN étranger de virus. On les utilise dans nombre d'organismes pour effectuer des modifications dirigées par des molécules d'ARN conçues sur mesure.



LE GÉNIE GÉNÉTIQUE EN MULTIPLEX permet d'incorporer simultanément dans la même cellule plusieurs fragments d'ADN en différents points ciblés de son chromosome. Dans la méthode MAGE, le ciblage est effectué grâce à des fragments d'ADN que la cellule assimile à des intermédiaires de la réplication du génome lors de la division cellulaire. Ces fragments se retrouvent ainsi incorporés dans le génome.