

Les verrous scientifiques et technologiques dans la phase conceptuelle de la biologie de synthèse

Le génie biologique est actuellement en rapide évolution. Ses formes avancées, actuellement rassemblées sous l'expression *biologie de synthèse*, couplent itérativement la conception et la fabrication d'objets complexes basés sur (ou inspirés par) la biologie. Comme la nanotechnologie, la biologie de synthèse est susceptible de changer totalement notre approche de certaines technologies clés. Ainsi, elle ouvre la voie à une nouvelle génération de produits, d'industries et de marchés construits sur notre capacité de manipuler la matière à l'échelle moléculaire.

Les applications potentielles de la biologie de synthèse se situent principalement dans les domaines de la santé, de l'agroalimentaire, de l'environnement, de l'énergie et des matériaux.

La phase de fabrication relève de méthodes biomoléculaires dites post-génomiques qui font l'objet d'intenses efforts de recherche depuis plusieurs décennies.

Mais c'est la phase de conception qui est affectée des verrous scientifiques et méthodologiques les plus saillants, une situation qui reflète probablement un investissement bien moindre.

Dans cet article, nous allons justement identifier certains de ces verrous.

Par **François KÉPÈS***

* Programme d'Épigénomique (Genopole®), institut de Biologie des Systèmes et de Synthèse (Genopole®, UEVE, CNRS), *Center for Synthetic Biology and Innovation (Imperial College London)*.

L'ÉTAT DE L'ART

Les ingénieries des systèmes mécaniques, électriques ou chimiques sont rendues possibles par l'existence de cadres bien établis pour en gérer la complexité, d'outils fiables pour manipuler les états d'un système donné, et de plateformes de test.

A contrario, l'ingénierie de la biologie est, elle, largement dépourvue de ces cadres, outils et plateformes. Actuellement, pour développer des composants ou des circuits biochimiques fonctionnant correctement, il faut suivre des protocoles compliqués de construction de séquences d'ADN, utiliser des modèles rudimentaires pour en guider le *design*, et faire appel à des cycles répétés d'essais-erreurs se terminant par d'interminables réglages tous plus fins les uns que les autres. À l'intérieur d'une cellule vivante, les interactions entre composants synthétiques et composants cellulaires sont difficiles à prévoir ou à caractériser. En conséquence, les cycles de conception, de fabrication et de dépannage de composants et de circuits biochimiques restent lents et sujets à erreurs.

Cependant, quelques progrès ont été enregistrés dans la fabrication de systèmes biologiques artificiels. Ces progrès concernent en particulier les ressources et les méthodes permettant de construire des dispositifs ou des systèmes à partir de petits composants, en général des segments d'ADN bien définis. En outre, les technologies de lecture et d'écriture de l'ADN à grande échelle ont graduellement rendu possible la programmation directe de gènes faisant éventuellement partie de circuits biochimiques. En coopération avec la bioinformatique, il est également devenu possible d'optimiser ces séquences d'ADN en fonction d'objectifs ingénieraux.

Voici quelques jalons fondamentaux balisant ce terrain.

* En 1995, le premier génome bactérien est séquencé et le premier gène long est synthétisé.

* En 1999, le premier génome viral est synthétisé.

* En 2000, est démontrée la faisabilité de concevoir rationnellement des circuits de régulation et de les implanter dans des bactéries, et le génome humain est séquencé.

* En 2003, un repli (*fold*) protéique non naturel est conçu, puis réalisé.

* En 2008, un génome bactérien est, pour la première fois, entièrement re-synthétisé chimiquement, mais c'est la copie quasi à l'identique d'un génome naturel ; il n'a fait l'objet d'aucune conception globale.

* En 2010, ce génome bactérien est introduit avec succès dans une bactérie hôte capable de se reproduire.

* En 2011, une bactérie vit et se reproduit avec un génome dont l'un des quatre nucléotides constitutifs a été quasi entièrement remplacé par un autre chimiquement différent.

Malgré ces quelques avancées, la conception et la fabrication de circuits biochimiques artificiels qui remplissent les fonctions souhaitées, restent un art très difficile et peu répandu. En outre, la conception *ab initio* de génomes, qui étendrait fortement la taille et la complexité des circuits envisageables, reste irréalisable.

VERROUS SCIENTIFIQUES ET VERROUS TECHNOLOGIQUES

Le génie biologique est actuellement en rapide évolution. Ses formes avancées, rassemblées sous l'expression *biologie de synthèse*, couplent itérativement la conception et la fabrication d'objets complexes basés sur (ou inspirés par) la biologie [1]. La phase de fabrication repose principalement sur les méthodes de la biologie moléculaire. Ces méthodes progressent sur une large ligne de front qui concerne un nombre considérable d'acteurs. En revanche, la phase de conception est encore balbutiante et ne concerne que peu d'acteurs hautement spécialisés. Nous présenterons les verrous scientifiques et technologiques de cette phase de conception, ceux de son articulation avec la fabrication, mais nous ne traiterons pas de la phase de construction proprement dite.

Les objets concernés par les formes avancées du génie biologique sont typiquement des biomolécules et des circuits biochimiques assurant un métabolisme ou la régulation d'une fonction biologique, ainsi que des cellules ou des particules abritant ces biomolécules et ces circuits.

Ce sont donc typiquement des objets micrométriques ou nanométriques. Les propriétés de ces objets minuscules exigent une intermédiation technique, qui fournira la dimension méthodologique de la présente discussion.

Développer cette intermédiation revient non seulement à bénéficier de l'usage direct de nos capacités cognitives et sensorielles, dont on sait qu'elles sont considérables, mais aussi à outiller notre intelligence collective. Cela participe également à démystifier des biotechnologies qui nous sont sensibles parce qu'elles touchent au biotope.

Cependant, pour pouvoir enregistrer durablement des progrès dans ce domaine, les défis à relever sont également d'ordre scientifique. Ces défis scientifiques touchent aussi bien à une meilleure compréhension du vivant qu'à des progrès dans des domaines théoriques aussi variés que l'intelligence artificielle ou la physique des milieux désordonnés. Notons que, séparément, les domaines techniques et scientifiques concernés existent bien, mais qu'ils devront progresser et, surtout, coopérer entre eux. C'est là que réside le véritable défi, qui exigera l'ouverture d'un espace de

créativité à la mesure de ce que fut l'élan vers la biologie moléculaire, à la fin des années 1960.

Globalement, notre thèse est que la phase de conception souffre d'un outillage insuffisant à la fois pour épauler et exploiter les facultés humaines.

Épauler

Le processus de conception d'objets biologiques est actuellement obéré par la fragmentation de ses segments logiciels, qui ont été développés indépendamment et dont les interfaces n'ont pas été accordées. La raison est d'ordre historique, car chaque segment concrétise un état de l'art pointu dans un domaine distinct, avec ses notions idiosyncratiques de ce que doivent être ses entrées et ses sorties. Il serait très utile de fluidifier ces processus. La jeune pousse américaine *Gingko Bioworks* [2], établie dans la région de Boston, a notablement avancé dans cette direction. Elle a ainsi démontré la validité et la faisabilité d'une telle approche. En revanche, ses solutions ne présentent pas de caractère générique et servent uniquement à ses travaux internes. Aussi, le défi consistant à établir un cadre procédural à valeur quasi-universelle reste entier.

En amont de la conception, se trouve le cahier des charges du circuit biochimique souhaité. Ce cahier des charges comporte deux versants distincts, l'un technique et l'autre réglementaire (droit, normes et propriété intellectuelle). Prenons un exemple qui vaut pour l'ensemble du versant réglementaire. La propriété intellectuelle des composants biochimiques à assembler dans un circuit un tant soit peu complexe est une véritable jungle. Pendant un certain temps, la négociation des « paquets » de brevets permettra la survie dans cette jungle, mais elle atteindra vite ses limites.

Il serait donc utile, et bientôt nécessaire, d'assister l'expert en formalisant les aspects réglementaires, normatifs et de propriété intellectuelle. Pour cela, il conviendrait de s'appuyer sur des bases de données qui souvent existent (comme, par exemple, les bases d'ordre réglementaire sur la toxicité ou, en matière de propriété intellectuelle, sur les biobriques).

Au plan technique, le facteur très limitant est le niveau d'expertise nécessaire pour rédiger un cahier des charges et en éliminer les erreurs en l'absence d'une médiation adéquate. Pour que cette maîtrise se répande, il est nécessaire d'outiller l'ingénieur. Cela représente un défi majeur. En effet, depuis les débuts de la révolution industrielle, il est clair qu'il n'existe pas d'autre approche pour développer massivement un domaine technologique que celle de le « démocratiser », et donc d'outiller, en l'ouvrant à un grand nombre d'ingénieurs, une expertise jusque-là réservée

à un groupe très restreint de personnes hyperspécialisées.

En outre, la phase de conception n'est pas actuellement alimentée en retour par l'analyse des résultats et par celle de leur reproductibilité expérimentale. Pour mettre en place ces boucles de rétroaction, il faudrait intégrer l'expérimentation robotisée et la documentation digitale complète des protocoles expérimentaux. Durant la phase de conception, une approche clé du génie biologique avancé consiste à identifier le (ou les) génotype(s) qui conduira(ont) au phénotype désiré. Par exemple, dans le cadre du *design* de protéines, il s'agirait de découvrir les séquences d'ADN dont le décodage aboutirait à une protéine présentant telle ou telle structure ou ayant telle ou telle activité. C'est ce que l'on appelle le problème inverse, par opposition au problème direct, qui consisterait dans notre exemple, à partir d'une séquence – le génotype – pour calculer la structure repliée d'une protéine – le phénotype.

Or, dans tous les cas, le problème inverse est plus complexe à traiter que le problème direct.

Pour poursuivre avec le même exemple dans lequel le problème direct est déjà en lui-même malaisé à résoudre sauf à recourir à des analogies avec des protéines connues, le problème inverse consisterait à replier un nombre énorme de séquences possibles pour identifier celles qui donneront la structure désirée. L'espace des génotypes est démesuré et chaque test par repliement est coûteux. La parallélisation et l'augmentation constante des capacités de calcul scientifique vont certes dans le bon sens. Cependant, nos capacités calculatoires resteront incommensurables, face à la démesure de l'espace des génotypes. Il sera donc indispensable de développer des heuristiques sophistiquées pour n'avoir à explorer qu'un petit sous-espace utile des génotypes.

Un dernier verrou, enfin, tient à la difficulté de fabriquer un produit en très petite série dans des conditions économiquement viables.

Prenons le cas du traitement d'une seule personne ou d'une petite population, par exemple pour des maladies du système immunitaire ou pour des maladies rares. Le nombre de tels cas augmente actuellement à vive allure, et ce pour deux raisons. La première raison est la constante augmentation du nombre des personnes atteintes de maladies immunitaires. La seconde raison est la fragmentation des maladies non rares. En effet, les méthodes moléculaires de diagnostic affinent les distinctions entre des sous-cas, dont le pronostic et le traitement doivent être différenciés. Ce phénomène fragmente de fait les cohortes atteintes de maladies non rares en plusieurs petites populations, chacune étant affectée par une maladie rare. Ces cas de plus en plus nombreux sont incompatibles avec le schéma traditionnel de production thérapeutique et diagnostique de masse. Certes, ce schéma traditionnel ne se dévalorisera pas, mais une

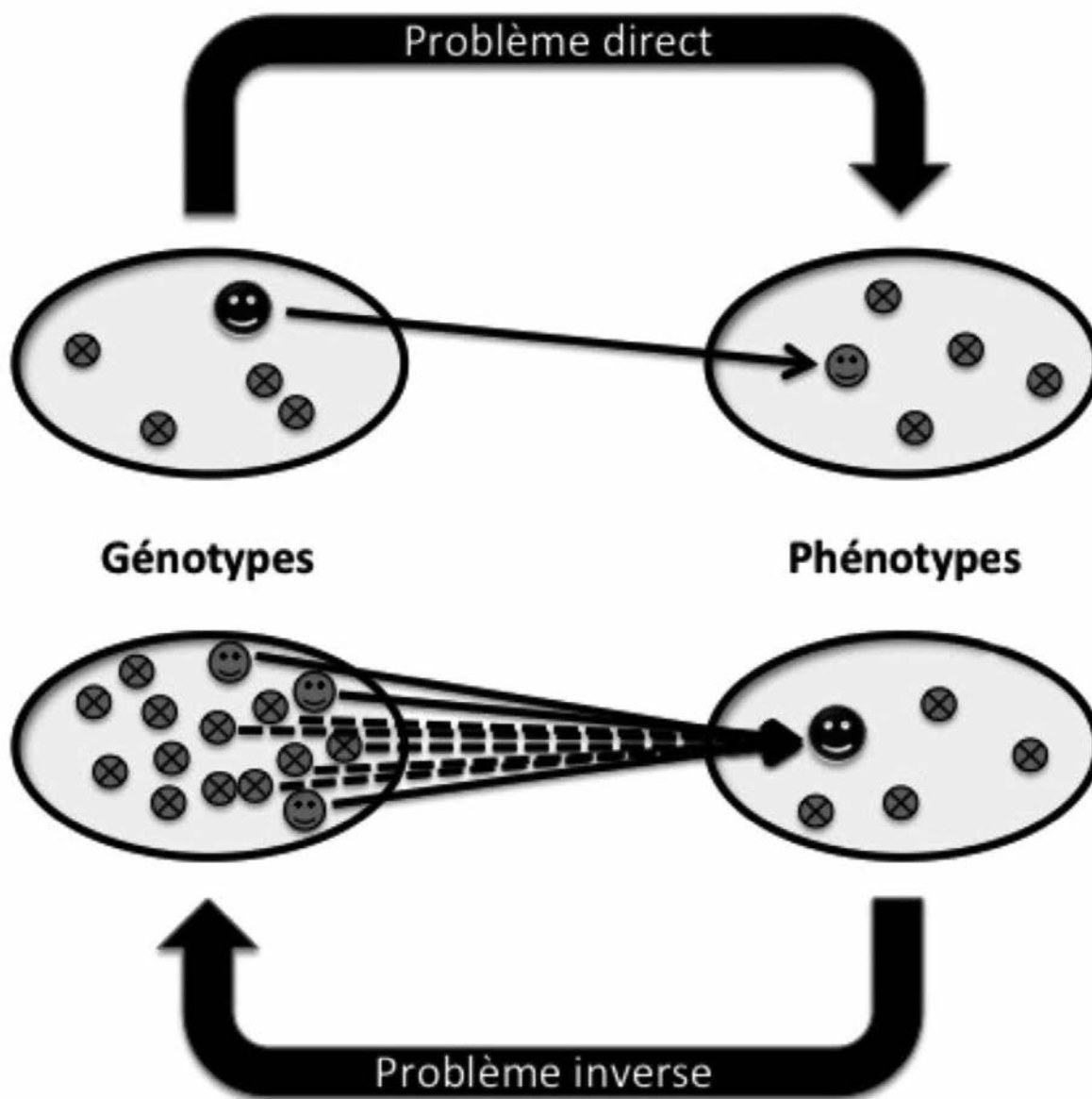


Image 1 : Schématisation des notions de *problèmes directs* et *problèmes inverses*.

approche supplémentaire, à petite échelle, devrait émerger pour permettre une personnalisation de la production thérapeutique, diagnostique, ou dans d'autres domaines.

Exploiter

Il est paradoxal et dommageable que le génie biologique avancé, alors même qu'il résulte de concepts découverts et orchestrés par des hommes, sous-exploite les facultés humaines durant sa phase de conception. Les facultés utiles dans ce contexte résident dans une exceptionnelle capacité de chaque individu à résoudre des énigmes et à reconnaître des configura-

tions ou des patrons *via* l'un ou l'autre de ses cinq sens. En outre, l'intelligence humaine distribuée permet de résoudre de manière collaborative des énigmes plus complexes qu'un individu ne saurait aborder seul. Cependant, que cela soit au niveau individuel ou collectif, un outillage approprié est absolument nécessaire pour que certains de nos sens puissent appréhender et manipuler des objets virtuels, ne fût-ce qu'en raison de leur taille micrométrique ou nanométrique. À ce titre, la vision, l'audition et le toucher devraient être immergés dans une boucle de rétroaction sensorimotrice.

Cependant, l'intelligence collective des experts ne concerne par définition qu'un très petit nombre d'individus. Le succès récent du jeu *Foldit* dans le domaine du repliement des protéines, avec 50 000 joueurs dans



le monde, suggère de tirer profit d'une source plus large, même si elle est moins experte, d'intelligence collective. *Foldit* est un jeu coopératif en ligne qui permet à des volontaires, issus du grand public, de contribuer à l'intelligence collective en résolvant des problèmes scientifiques [3]. De plus, ces volontaires peuvent éventuellement apporter de la puissance de calcul à travers leurs ordinateurs, tout en s'éduquant et en contribuant à diffuser largement la culture scientifique.

BIBLIOGRAPHIE

- [1] KÉPÈS (F.), « La biologie de synthèse : développements, potentialités et défis », *in* « Des nanotechnologies à la biologie de synthèse », *Réalités Industrielles (Annales des Mines)*, février 2010.
- [2] <http://ginkgobioworks.com>
- [3] <http://fold.it/portal/info/science>