

# La RECHERCHE

## L'ACTUALITÉ DES SCIENCES



Le magazine La Recherche a publié en février un dossier sur « La cellule en équation ». Voici un extrait

## Et la cellule devint virtuelle

« **S**i quelqu'un devait analyser les idées du moment et les slogans à la mode, il trouverait "systèmes" au sommet de la liste », écrivait, en 1967, l'influent théoricien de la biologie Ludwig von Bertalanffy [1]. Il est décédé en 1972, mais son propos rencontre aujourd'hui une résonance nouvelle : en ce début de XXI<sup>e</sup> siècle, impossible d'échapper à la « *systems biology* ». Les programmes de recherche qui en appellent à cette « biologie intégrative » ou « biologie systémique » se succèdent, les revues scientifiques telles que *Nature* ou *Science* lui consacrent des dossiers entiers [2].

De quoi s'agit-il ? En termes lapidaires, de revenir à l'approche dégagée par Bertalanffy : une cellule, un organe ou un organisme sont autant de systèmes dont le fonctionnement dépend certes des entités en présence, mais aussi et surtout des interactions qui les lient. Autrement dit, le tout est plus que la somme des parties. C'est réinventer la roue ! s'exclament certains, tant il est vrai que les biologistes – en particulier les physiologistes – ont toujours accordé de l'importance aux interactions. Pourtant, il y a bien nouveauté. Les techniques d'analyse à haut débit de la génomique et de la protéomique ont engendré une si grande quantité de données qu'il est aujourd'hui envisageable, en faisant appel aux mathématiques et à l'informatique, de réaliser la modélisation et la simulation de systèmes qui intègrent des niveaux jusque-là étudiés de façon distincte [3].

Parmi ces systèmes, figure, bien sûr, la cellule. Et le terme de « cellule virtuelle » commence, peu ou prou, à émerger sur la scène scientifique. Que veut-on dire par là ? Quels sont les objectifs visés, et les approches employées ? Est-ce une recherche purement académique, ou l'industrie s'y intéresse-t-elle aussi ? Le point, dans ce dossier de *La Recherche*.

[1] L. von Bertalanffy, *General System Theory: Foundations, Development, Applications*, George Braziller Inc., New York, 1976 (édition révisée).

[2] Dossiers spéciaux « *Systems biology* » : *Science*, 295, 1661, 2002 ; *Nature*, 420, 205, 2002 ; *Nature Biotechnology*, 22, 1249, 2004.

[3] H. Kitano (éd.), *Foundations of Systems Biology*, MIT Press, 2001.

## L'industrie cherche son modèle

Si les groupes pharmaceutiques semblent, pour l'instant, plutôt indifférents à la modélisation cellulaire, plusieurs sociétés de biotechnologie ont choisi d'occuper ce créneau. Les premiers résultats sont timides, mais encourageants.

**Olivier Blond**  
est journaliste scientifique.

**Q**u'il s'agisse de séquençage du génome ou d'analyse des protéines, les technologies à haut débit ont rempli les disques durs des chercheurs de quantités énormes d'informations. La question est désormais de les interpréter. « Certains pensent que la biologie systémique [les modèles informatiques] détient la clé qui permettra de réassembler toutes ces informations et d'accélérer la découverte de nouveaux médicaments », lit-on dans la revue *Nature Biotechnology* d'octobre 2004. Mais Frank Molina, du laboratoire de bioinformatique pour la biologie systémique du CNRS, à Montpellier, nous met en garde : « Les ambitions sont énormes, et nous sommes à l'aube d'une révolution en biologie. Mais il faut savoir faire la part de la réalité et des discours marketing. Non pas que la science-fiction ait tort, mais les résultats qu'elle promet n'arriveront que dans quinze ans. »

## Relier les informations

À l'instar de nombreux acteurs de la recherche privée dans ce secteur, François Iris, président fondateur de la société Bio-Modeling Systems (BMSystems), adhère à cette prudence: « *Un fossé colossal sépare nos capacités actuelles de la réalisation d'une cellule hépatique ou neuronale virtuelle. Mais il est impossible d'attendre que ces modèles soient achevés. C'est pourquoi nous en cherchons d'autres, plus simples, qui amènent des prédictions que l'on puisse confronter à des expériences.* »

Pour l'instant, les projets privés sont encore peu nombreux, et la plupart sont l'œuvre de petites entreprises de biotechnologie. En effet, les industriels de la pharmacie hésitent encore à entrer dans la course. « *Ils sont méfiants, car trop d'argent a été dilapidé dans des start-up qui promettaient monts et merveilles* », explique Laurent Buffat, directeur scientifique de IT-Omics. Et si Eli Lilly peut se targuer d'avoir créé, à Singapour, le plus grand centre en ce domaine, il n'emploie pas plus de 60 personnes.

Où en est-on en France? Parmi les acteurs, la société Atragene Bioinformatics développe des outils informatiques pour identifier de nouveaux facteurs de transcription. Ces molécules se fixent sur l'ADN en des zones particulières et contrôlent ainsi l'expression d'autres gènes. Chaque facteur de transcription en active d'autres qui, fonctionnant en cascade, jouent un rôle fondamental. Le dérèglement de la chaîne est à l'origine de nombreux cancers. Comme l'explique Alain Malpertuy, directeur scientifique d'Atragene: « *Ces facteurs de transcription sont mal connus. Chez l'humain, on estime leur nombre à plus de 3000, mais on ne connaît qu'environ 300 sites de fixation.* »

Atragene Bioinformatics commercialise deux logiciels permettant d'analyser les profils d'expression des gènes obtenus grâce à des puces à ADN\*, et d'identifier d'éventuels éléments génétiques de régulation. Elle met également en place des plates-formes d'analyse bioinformatique. L'enjeu est ici d'acquiescer suffisamment d'informations pour construire des modèles (relativement simples) qui expliquent la régulation des processus étudiés.

Certaines « biotechs » veulent aller plus loin: mettre en relation non pas quelques dizaines de molécules, mais des milliers. C'est l'ambition de IT-Omics. Cette société construit des outils de gestion de données appelés graphes. L'idée est de fouiller automatiquement l'ensemble de la littérature scientifique et d'en extraire des « relations » entre des objets biologiques – gènes, protéines ou autres molécules. Concrètement, cela signifie croiser la base de publications Medline (qui contient plus de 12 millions d'articles de biologie et médecine, la quasi-totalité de ceux publiés depuis 1960) avec les bases de données de gènes ou de protéines (GenBank et Swiss-Prot, par exemple), en recherchant, dans les résumés des articles, toute mention d'un lien entre deux entités A et B.

« *Le graphe de IT-Omics contient 500 000 relations entre 18 000 objets*, explique Laurent Buffat. *C'est une sorte de représentation pragmatique de la connaissance* ». Comment utiliser un tel outil? « *Il faut poser une question précise. Sans cela, on se noie dans les données.* » Il peut s'agir d'interpréter les résultats d'une puce à ADN sur laquelle on a observé la surexpression simultanée de 1 500 gènes: on plonge ces données dans le graphe et on essaie d'en sortir, par exemple, un lien fonctionnel entre ces gènes. La société a déjà vendu plusieurs licences de son logiciel à des groupes pharmaceutiques comme Aventis ou UCB\*. IT-Omics assure aussi des fonctions de service, en proposant des interprétations à la carte.

Toutefois, selon François Iris, de BMSystems, le système des graphes souffre de deux inconvénients majeurs: d'une part, il est difficile de savoir si l'information est fiable, car la littérature scientifique regorge d'erreurs. D'autre part, la complexité d'un graphe croît exponentiellement avec le nombre de ses éléments, jusqu'à devenir ingérable. BMSystems a donc élaboré un système de « sélection négative »: un outil semi-automatique de réfutation qui génère des hypothèses de relation entre entités cellulaires et les teste les unes après les autres en les recroisant avec les bases de données. Le but est d'en éliminer le plus possible jusqu'à n'en conserver qu'un nombre suffisamment restreint pour permettre de proposer des expériences. « *À l'inverse des graphes, plus l'information est riche et diversifiée, meilleures sont nos hypothèses. Nous avons même un problème quand il n'y a pas assez d'informations. Quant aux incohérences internes, elles nous servent car elles permettent d'éliminer les hypothèses fausses* », explique François Iris.

## Débouché médical

Le « business modèle » de BMSystems diffère lui aussi de celui de ses concurrents. Manuel Géa, PDG de l'entreprise, a choisi non de vendre une technologie particulière (le logiciel reste confidentiel) ou de développer un modèle exhaustif, mais de proposer des « solutions » à des problèmes industriels. Par exemple, BMSystems a testé son approche avec une forme de cancer du sein. Ses chercheurs ont identifié plusieurs voies métaboliques, qu'ils ont rassemblées dans un modèle. Et il est apparu que le blocage simultané de trois d'entre elles devait stopper la prolifération cancéreuse. Or, en pratique, chacune de ces voies peut être bloquée par une molécule peu toxique, à l'inverse des anticancéreux classiques. Sur des cellules en culture, chaque molécule utilisée seule n'apporte qu'une légère amélioration. Mais, utilisées ensemble, elles sont efficaces dans 75 % des cas! Ces résultats ont été obtenus indépendamment par trois hôpitaux publics français et publiés [1] – sauf pour la troisième molécule, qui reste secrète. Manuel Géa affirme être en négociation avec un industriel pour développer cette approche thérapeutique. Il affirme également que BMSystems a déjà réalisé neuf modèles prédictifs de pathologies humaines, dont deux ont déjà été validés et publiés (pour le cancer du sein), et deux autres (dont un pour la maladie de Creutzfeldt-Jakob) sont en « évaluation expérimentale » dans des laboratoires de recherche publics.

### \*Une puce à ADN

est un dispositif qui permet de visualiser l'expression d'un très grand nombre de gènes (plus de 10 000) de façon rapide et simultanée.

\*UCB est une société pharmaceutique basée à Bruxelles et productrice, notamment, des anti-allergiques Zyrtec et Atarax.

[1] F. Gadal, *Nucl. Acids Res.*, 31, 5789, 2003; F. Gadal et al., *J. Mol. Endoc.*, sous presse.

Certains des projets de la recherche privée portent donc leurs fruits. Mais comparés à ceux développés dans le monde académique, les systèmes utilisés ne sont pas des modèles itératifs évoluant avec le temps (voir l'article de F. Rechenmann et H. de Jong). Par ailleurs, reconnaît Laurent Buffat : « *Nos modèles sont des outils de connaissance et d'interprétation des données qui permettent d'identifier, dans le flux des données, ce qui n'est pas connu et ce qui est intéressant. Ce ne sont pas encore des modèles qui permettent une approche rationnelle de la recherche pharmaceutique.* » On est encore loin des méthodes qui permettraient enfin à l'industrie pharmaceutique de découvrir de nouvelles catégories de molécules. Loin aussi de la toxicogénomique : prévoir si une molécule aura des effets secondaires néfastes et dans quelle mesure. Mais cela pourrait changer. ■ O. B.

#### POUR EN SAVOIR PLUS

■ Les sites des sociétés  
[www.atragene.com/atragene/](http://www.atragene.com/atragene/)  
[www.it-omics.com/v3gb/index.html](http://www.it-omics.com/v3gb/index.html)  
[www.bmsystems.net/](http://www.bmsystems.net/)

© LA RECHERCHE, N°383, FÉVRIER 2005, pp. 30-39.

## Poursuivez votre lecture !

### Abonnez-vous à La Recherche

A retourner sous enveloppe affranchie à La Recherche - Service Abonnements  
 B604 - 60732 Ste-Genevieve Cedex

**OUI**, je souhaite m'abonner et recevoir **en cadeau** la clé USB

La Recherche, je choisis :

**L'offre couplée** : 1 an (11 numéros dont 4 spéciaux)  
 + 4 numéros des Dossiers de La Recherche  
 pour **72€ seulement** au lieu de **112€**, soit une économie de 35 %.  
 Je suis étudiant(e) ou enseignant(e)\*, je ne paie que **65€**.

**L'offre simple** : 1 an (11 numéros dont 4 spéciaux) pour  
**52€ seulement** au lieu de **86€** soit une économie de 40 %.  
 Je suis étudiant(e) ou enseignant(e)\*, je ne paie que **46€**.

\*Pour bénéficier de ce tarif préférentiel, merci de nous envoyer un justificatif de votre statut.

Je préfère profiter de la formule « 4 fois sans frais »

Je choisis l'offre couplée et je règle dès maintenant **18€** par chèque ou carte bancaire, ce qui correspond au quart du montant de l'abonnement (11 numéros de La Recherche + 4 Dossiers).

Je choisis l'offre simple et je règle dès maintenant **13€15** par chèque ou carte bancaire, ce qui correspond au quart du montant de l'abonnement.

Puis je recevrai par courrier l'autorisation de prélèvements qui me permettra de régler les trois autres versements sans aucun frais supplémentaire par prélèvement automatique trimestriel.

J'indique ci-dessous mes coordonnées :

Nom : \_\_\_\_\_

Prénom : \_\_\_\_\_

Adresse : \_\_\_\_\_

Code Postal : [ ][ ][ ][ ][ ][ ]

Ville : \_\_\_\_\_

E-mail : \_\_\_\_\_

Règlement à l'ordre de La Recherche par :

Chèque  Carte Bancaire

Numéro CB : [ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ][ ]

Notez aussi les 3 derniers chiffres du numéro inscrit  
 au dos de votre carte bancaire, au niveau de la signature [ ][ ][ ]

Expire fin : [ ][ ][ ][ ][ ]

Signature obligatoire :

\_\_\_\_\_

Le prix acquiesçamment les numéros normaux au prix de 9€90, les numéros spéciaux au prix de 9€20, les Dossiers de La Recherche au prix de 6€50 et la clé USB La Recherche au prix de 20€. Offre réservée à la France métropolitaine, valable uniquement jusqu'au 31/03/2005 et dans la limite des stocks disponibles. La formalité et identifiants - vous disposez d'un droit d'accès et de rectification vous concernant et vous pouvez vous opposer à leur cession.