

# BIOINFORMATIQUE

## Quand le labo se fait virtuel

Le programme BioIntelligence, auquel a participé l'Inserm, a permis de mettre au point une plateforme informatique collaborative capable de simuler l'action de médicaments sur une cellule, de modéliser sa toxicité ou encore d'anticiper l'issue de différentes prises sur des malades. Retour sur l'aventure concrète d'un laboratoire virtuel.



Plus de 100 millions de données dans l'application BIOVIA

puisque leur métabolisme diffère. La bioinformatique nous y donne accès facilement », explique Sophie Visvikis-Siest (☛), directrice de l'unité de recherche Interactions gène-environnement en physiopathologie cardiovasculaire (IGE-PCV). Pour parvenir à de telles prouesses, les partenaires de ce projet se sont inspirés du principe « Product Lifecycle Management » (PLM), ou « gestion du cycle de vie des produits ». Développé initialement dans le domaine de la métallurgie et plus largement pour les

Et si, plutôt que de tester de nouveaux médicaments in vivo, les chercheurs le faisaient virtuellement en modélisant, notamment, les réactions du corps humain vis-à-vis de différentes molécules à potentiel pharmaceutique ? C'est l'objectif dont s'est emparé le projet BioIntelligence. De 2009 à 2014, un consortium de partenaires\* a souhaité bouleverser la R&D pharmacologique en développant un ensemble d'applications logicielles. Objectif : proposer aux industriels et aux chercheurs une plateforme d'outils de modélisation, de simulation et de prédiction permettant de mener à bien leurs recherches et la mise au point de nouveaux produits. À l'issue de cinq années de développement, une quinzaine de logiciels a été peaufinée et est aujourd'hui en cours de commercialisation.

« Un laboratoire ne peut pas fonctionner aujourd'hui sans la bioinformatique. Grâce à cette approche multidisciplinaire, nous sommes en mesure de modéliser jusqu'à l'environnement pathologique dans lequel se trouve une molécule. C'est important, car elle n'a pas le même rôle ni le même fonctionnement chez un individu ou un autre,

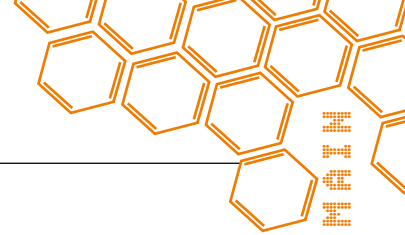
sciences de l'ingénieur, il permet d'optimiser la conception de nouveaux produits grâce à une suite de logiciels fonctionnant ensemble autour d'une plateforme commune : logiciels pour traiter de grandes quantités de données, pour rechercher et agréger toutes les informations disponibles sur un sujet précis, visualiser en 3D les résultats des études ou encore pour faciliter le contact entre toutes les équipes travaillant à un même projet. Afin d'utiliser ces outils dans le domaine pharmacologique, il a fallu les adapter.

« Dans le cas des sciences de l'ingénieur, on a une idée de ce que l'on invente, de ce que l'on veut concevoir : un nouveau pare-choc de voiture, une coque de bateau, etc. Les outils PLM sont alors utilisés dès le début de la conception. Tout est différent avec la pharmacologie : dans les phases amont, il faut explorer, découvrir des mécanismes biologiques profonds et l'on ne dispose pas forcément de l'ensemble des informations associées ! Nous ne connaissons pas, par exemple, certaines interactions biologiques à cibler. Dans notre domaine, nous découvrons autant que nous concevons des produits. Sans compter que nous

\* Le consortium réunissait le chef de file Dassault Systèmes, cinq partenaires industriels (Ipsen, Pierre Fabre, Sanofi-Aventis, Servier et Bayer CropScience), deux PME spécialisées dans la bioinformatique (Sobios et Aureus Pharma) et trois partenaires publics (l'INRIA, le Genopole et l'Inserm).

\*\*Voir S&S n° 18, Découvertes « Chémo-informatique – De l'algorithme au médicament », p. 12-13

☛ Sophie Visvikis-Siest : Unité 1122 Inserm – Université de Lorraine



LE MOYEN DE MODÈLES INFORMATIQUES

© DASSAULT SYSTÈMES



**Application BIOVIA : carte des cancers humains**

*avons affaire à des processus de nature bien différente, et nous ne maîtrisons pas encore entièrement la complexité incroyable du corps humain !* », explique Patrick

Johnson, vice-président *Corporate Research & Sciences* pour Dassault Systèmes.

**Une plateforme nourrie aux Big Data**

Afin de prévoir des mécanismes complexes, tels que le bénéfice ou la toxicité de molécules ou les interactions entre cellules, la collection d'informations est une des clés de voûte : en accumulant les données cliniques issues du suivi de patients ou celles issues d'expériences en laboratoire de recherche, des résultats peuvent être corrélés et la simulation du vivant *in silico*\*\* (🔗), approchée. Dans le cadre des maladies cardiovasculaires, par exemple, l'étude du profil génétique des malades, croisée avec l'identification en laboratoire des mécanismes de fonctionnement des gènes impliqués dans le développement de ces pathologies, permet de simuler comment la présence de tels ou tels variants (🔗) de ces gènes influe sur la survenue de ces cancers.

**Esperanto scientifique**

Pour faire fonctionner un tel projet de « laboratoire virtuel », il a fallu passer par une étape essentielle : standardiser les termes utilisés par tous les participants afin qu'ils parlent tous le même langage. Problème : pour désigner telle molécule ou tel processus, chaque spécialité scientifique possède des termes spécifiques, lesquels, une fois intégrés à un modèle informatique ne donnent qu'un brouhaha incompréhensible ! C'est, par exemple, le cas pour le terme « interactions », qui peut soit désigner des interactions entre voies métaboliques, entre gènes, entre gènes et environnement, avec une conséquence biologique, mais qui peut aussi désigner une interaction statistique. Sophie Visvikis-Siest et ses collègues ont ainsi travaillé à la mise au point d'une sorte d'« esperanto » scientifique afin que tout le monde se comprenne et que les résultats des recherches menées par tous les partenaires puissent être compatibles avec la plateforme BioIntelligence : pour « interactions », il est désormais demandé de préciser pour chaque utilisation le domaine respectif.

L'intégration des informations contenues dans des bases de données déjà existantes a été primordiale. C'est le cas de

la plateforme Biostatistiques/Bioinformatique, administrée par l'IGE-PCV. « Pour comprendre les mécanismes impliqués dans la physiopathologie des maladies multifactorielles liées au vieillissement, il faut exploiter une quantité importante de données expérimentales hétérogènes : biologiques, protéomiques (🔗), génétiques, épigénomiques (🔗) ou encore transcriptomiques (🔗). Pour extraire des informations, notre plateforme assure une analyse à haut débit et une veille méthodologique constantes. De telles informations, qui décrivent notamment les liens entre des éléments moléculaires et des observations physiopathologiques sur une population donnée, sont très intéressantes lorsqu'il s'agit de modéliser le vivant », explique Sophie Visvikis-Siest. En retour, la chercheuse et son équipe ont bénéficié des apports de BioIntelligence pour alimenter le consortium international fédératif de génomique « *Vascular Endothelial Growth Factor* » (VEGF ou facteur de croissance de l'endothélium (🔗) vasculaire) dont le but est de développer un programme personnalisé de prévention des maladies cardiovasculaires fondé sur l'analyse du VEGF. Produite par les cellules, cette protéine a la capacité de stimuler la croissance des vaisseaux sanguins : ce facteur jouerait de fait un rôle important dans le développement de maladies cardiovasculaires. Au sein des projets du consortium VEGF, des outils « -omiques » (🔗) seront développés, tels que des biopuces et des algorithmes : grâce à ceux-ci, toutes les données accumulées (profils « -omiques » des individus, facteurs environnementaux, styles de vie, etc.) seront intégrées et serviront d'indicateurs prédictifs et diagnostiques de maladies cardiovasculaires. Un logiciel développé dans le cadre de

BioIntelligence sera appliqué à l'analyse des voies de signalisation (🔗) de deux des variants du VEGF.

Aujourd'hui, les logiciels développés pour alimenter une plateforme de recherche collaborative en pharmacologie, baptisée 3DEXPERIENCE, sont désormais opérationnels et en cours de commercialisation auprès de l'industrie pharmaceutique et des chercheurs. Pari réussi donc pour le programme BioIntelligence, collaboration entre le monde de la biologie et celui de l'informatique ! ■

Alice Bomboy

**In silico**

Recherche ou essai réalisé au moyen de modèles informatiques

**Variants**

Différentes versions existantes d'un gène

**Protéomique**

Science qui étudie les protéomes, soit l'ensemble des protéines d'une cellule, d'un tissu, etc.

**Épigénomique**

Étude de l'ensemble des modifications épigénétiques, soit des changements de l'activité des gènes sans modifications de l'ADN

**Transcriptomique**

Étude de l'ensemble des ARN messagers, copies transitoires de portions de l'ADN, au cours du processus de transcription

**Endothélium**

Couche la plus interne des vaisseaux sanguins, en contact avec le sang

**-omiques**

Ce suffixe correspond aux technologies qui permettent d'étudier un grand nombre de données, tel que le séquençage du génome à grande échelle ou « génomique ».

**Voie de signalisation**

Ensemble de mécanismes de communication qui régissent le fonctionnement et l'activité des cellules.