
Le Programme d'Épigénomique

Constitution d'une communauté de recherche en modélisation et simulation dans le contexte de la génomique

Gilles Bernot — François Képès

*Programme d'Épigénomique
Genopole® et Université d'Évry
4, Bd François Mitterrand
F-91025 Évry cedex
{bernot, kepes}@epigenomique.genopole.fr*

MOTS-CLÉS : modélisation de processus biologiques, épigénomique, épigénèse, réseaux biologiques.

KEYWORDS: modelling of biological processes, epigenomics, epigenesis, biological networks.

Gilles Bernot est informaticien et François Képès est biologiste cellulaire.

1. « Chef, le simulateur est mort » ...

... dit l'infirmier militaire au médecin chef. Humour noir, mais la même phrase dans une équipe qui modélise et simule les processus cellulaires dans le contexte de la génomique pourrait révéler, dans un futur pas si lointain, des avancées scientifiques non négligeables : compréhension des mécanismes de mort cellulaire, étude détaillée *in silicio* de la cascade de causalités produisant ce phénotype, prédictions de points de contrôles pour le piloter ou l'éviter *in vivo* et leur cortège d'applications comme le traitement du cancer.

La génomique apporte une cartographie de plus en plus complète des interactions moléculaires au sein d'une cellule, qui accrédite l'espoir de simulations informatiques biologiquement crédibles d'une cellule. Là où il y a quelques années la génétique explorait de proche en proche un réseau de tunnels à la manière locale d'une taupe dévoilant certains processus moléculaires, la génomique arase la taupinière et dévoile de larges portions de réseaux d'interactions moléculaires.

Cette nouvelle situation confirme pour longtemps encore la biologie moléculaire comme une exceptionnelle pourvoyeuse d'informations, mais signe aussi la perte d'efficacité de son paradigme ordinaire. Par exemple un biologiste moléculaire reconnaît l'identité d'un compartiment cellulaire par présence une certaine protéine. Si cette protéine est là, ce ne peut être que parce qu'elle y est recrutée par une autre, laquelle est également recrutée, etc. Cette régression infinie est praticable en *terra incognita*. Elle se heurte maintenant au mur dressé par la finitude de la liste des gènes. Or cette liste est le fruit du séquençage des génomes qui représente précisément la victoire la plus indiscutable de la biologie moléculaire.

L'issue consiste à concevoir le vivant (par exemple notre compartiment cellulaire) en terme des processus dynamiques qui le réalisent. L'informatique théorique a cependant vite tempéré l'optimisme suscité par l'afflux de données sur les interactions moléculaires. Que dire d'un gène recevant des influences contradictoires ? Comment prédire les forces respectives d'interactions concurrentes alors que la plupart des données de cette nature ne sont pas connues *in vivo* ? Comment prendre en compte les délais d'actions ? Comment gérer les causalités circulaires ? Bref, connaître le plan de la taupinière ne suffit pas pour modéliser les comportements, il faudrait aussi des informations sur le trafic complexe qui l'habite.

Nombreux sont les hiatus dans notre connaissance d'un processus dynamique. Nous les comblons usuellement par des modèles plus ou moins abstraits, qui désormais peuvent s'alimenter à la corne d'abondance de la biologie moléculaire. C'est pourquoi la modélisation informatique dans le contexte de la génomique s'inscrit dans une tendance de fond à long terme. Mais c'est aussi pourquoi elle doit s'allier à l'induction expérimentale, par simulations informatiques et dans l'éprouvette, et par observation, perturbation et réinvention du vivant.

Pour ce projet de modélisation du vivant nous avons forgé le terme d'« Épigenomique », qui par analogie à « Épigénétique », fait allusion à la

construction des objets biologiques situés à des niveaux croissants d'intégration, mais avec le génome et non le gène pour référentiel. On le voit, relever le défi épigénomique est un projet nécessairement interdisciplinaire qui met à profit à la fois les développements fondamentaux de l'informatique et la nouvelle donne de la biologie pour déchiffrer la complexité de la cellule et de l'organisme vivants.

2. Naissance d'une communauté

Nous avons relevé ce défi épigénomique à Genopole®-Évry dès 2000 et nous avons progressivement créé une communauté de recherche interdisciplinaire sur le territoire français. Il faut faire vite car plusieurs pays se sont dotés d'instituts des sciences de la complexité presque dans chaque université, à proximité immédiate des laboratoires de biologie, informatique, mathématiques et physique ou chimie théoriques, et nous ne resterons pas longtemps en avance sur notre temps sans une institutionnalisation vigoureuse des initiatives épigénomiques. En France la monodisciplinarité est fortement ancrée dans les esprits, les procédures d'évaluation des chercheurs et des unités, à la racine même de l'organisation de la recherche. Avec le soutien de Genopole®, le Programme d'Épigénomique a adapté en conséquence le dispositif de création de la branche française de cette communauté.

Pour susciter un engagement commun dans le domaine, de chercheurs reconnus de leurs disciplines respectives, il fallait lever les appréhensions des difficultés de la pluridisciplinarité. Nous avons mis en place à Évry des formations spécifiques à la biologie à destination des chercheurs des autres disciplines et nous participons à des formations en modélisation à destination des chercheurs en biologie. Cette formation, les *Introductions Avancées à la Biologie*, devenue annuelle d'une durée d'une semaine et anglophone, apporte un flux non négligeable de nouveaux épigénomiciens et accueille de plus en plus de chercheurs étrangers.

Les *Introductions Avancées* permettent aux chercheurs de chaque discipline de comprendre les objectifs de recherche des autres disciplines. Le travail de terrain, lui, est de plus confronté à l'obstacle du langage, des terminologies apparemment partagées mais en fait avec des sémantiques sensiblement différentes. Pour lever ces ambiguïtés, le Programme d'Épigénomique a mis en place des *groupes de travail* : dans un premier temps deux groupes assez généralistes qui ont constitué le langage commun et la culture commune, et maintenant une palette d'une dizaine de groupes focalisés sur des questions biologiques ou des méthodes de modélisation ciblées, tous multidisciplinaires et codirigés par un biologiste et un modélisateur.

Face à l'engouement pour l'épigénomique, nous devons également préserver des évaluations scientifiques individuelles et collectives qui n'enclavent pas le domaine comme une exception au milieu des disciplines classiques. Nous sommes très attentifs à ce que chaque chercheur, et tout particulièrement nos jeunes chercheurs, possèdent *une majeure* c'est-à-dire une discipline reconnue selon les canaux existants, avec des pairs pouvant évaluer la qualité scientifique des recherches selon

les critères établis de la discipline majeure. Plus généralement, il est crucial que chaque discipline liée au Programme d'Épigénomique préserve fermement son identité. Cela implique un respect mutuel des orientations et des intérêts scientifiques. Les *Introductions Avancées* jouent là encore un rôle majeur dans la compréhension et donc le respect de chaque discipline.

Genopole® a joué un rôle fédératif de la communauté nationale dans ce domaine, facilité dès 2000 par un partenariat avec l'*Institut des Hautes Études Scientifiques* (IHÉS). Des chercheurs biologistes, informaticiens, mathématiciens, physiciens ou chimistes théoriciens ont constitué la communauté nationale autour des groupes de travail. Lors de la première évaluation scientifique en septembre 2002, après seulement 20 mois de fonctionnement, un ensemble de 39 articles multidisciplinaires avait été publié par des coauteurs qui ne se seraient pour la plupart même pas connus sans ces groupes de travail. À la suite de cette évaluation, Genopole® en partenariat avec l'Université d'Évry nous a chargés de créer le *Programme d'Épigénomique*. Un an plus tard, en janvier 2004, six bureaux nous ont été attribués ; le premier programme de visiteurs débuta quatre mois plus tard ; une bibliothèque de 300 livres choisis avec soin dans le domaine des sciences de la complexité appliqués à la biologie a été constituée et ouverte à tout le campus génopolitain ; etc. Le maître-mot du Programme d'Épigénomique est « modéliser pour comprendre » les processus biologiques dans le contexte de la génomique.

3. Vie d'une communauté polymorphe

Inspiré en partie du mode de fonctionnement de l'IHÉS, le *Programme d'Épigénomique* constitue un point de rencontre de référence pour la communauté scientifique du domaine. Il offre un ensemble cohérent de manifestations scientifiques qui facilitent les actions pluridisciplinaires et maintiennent une forte ouverture vers de nouveaux arrivants. L'accueil des chercheurs dans le domaine est outillé à tous les niveaux, des *Introductions Avancées* aux groupes de travail en passant par les Écoles Thématiques, les cours de haut niveaux, les manifestations internationales, l'accueil de post-doctorants, les incubations de jeunes équipes et les programmes de visiteurs courts ou longs termes.

Les groupes multidisciplinaires, maintenant au nombre d'une dizaine, sont les lieux concrets de travail de la communauté. Ils se réunissent sur trois jours tous les deux mois : une journée plénière au milieu des deux journées de réunions individuelles des groupes sert de plaque tournante à l'ensemble de la communauté. Elle permet par exemple d'obtenir des informations sur les groupes qui se réunissent les jours précédent et suivant, éventuellement de se faire coopter par un groupe. Elle s'organise autour de 3 ou 4 exposés d'intérêt général.

Le Programme d'Épigénomique veille à la fluidité des groupes. La naissance de nouveaux groupes de travail n'est pas rare ; des ateliers préliminaires servent à la fois de groupes de lecture d'articles et d'incubateurs de futurs groupes de travail

épigénomiciens. Les chercheurs du réseau épigénomicien appartiennent souvent à plusieurs groupes de travail, ainsi la fermeture d'un groupe ayant raisonnablement atteint son objectif se fait sans aucune difficulté ; ses forces rejoignent les autres groupes. On assiste également à des fusions ou des scissions de groupes de travail en fonction du nombre de participants et de l'ampleur et la proximité de leurs thématiques. Tout cela est facilité par la proximité immédiate de toutes les salles de réunion des groupes dans les locaux du Programme d'Épigénomique.

En conclusion, le Programme d'Épigénomique entretient une animation scientifique de qualité qui fédère la communauté de modélisation dans le contexte de la génomique. Il reçoit des visiteurs du monde entier, organise de nombreuses manifestations scientifiques, workshops et symposia (plus de 70 journées pleines par an d'animations diverses). Il met une palette complète d'outils à la disposition des chercheurs européens afin de faciliter l'accès de nouveaux chercheurs à la communauté épigénomicienne, de tenir à jour les connaissances de chacun et de faire évoluer les thématiques du domaine. Parmi ces outils notons : les *Introductions Avancées* déjà mentionnées ; une *École Thématique* annuelle qui permet d'une part de former les nouveaux chercheurs français et européens à la modélisation et la simulation dans le contexte de la génomique et d'autre part de regrouper la communauté francophone ; les *groupes de travail* à contours variables, des *cours de haut niveau* européens, tous les deux ans à Genopole®, qui présentent les dernières développements du domaine et sont ouverts à tous les chercheurs européens en sciences de la complexité ; des *think tank* (cercles de réflexion) d'une semaine sur des sujets ciblés avec 5 à 8 chercheurs de disciplines et pays différents ayant pour mission de produire un article novateur ; l'accueil d'*invités* pour des périodes d'une semaine à trois ans ; la possibilité de *former* des post-doctorants et d'*incuber* de jeunes équipes de recherche via des ATIGES au sein du Programme d'Épigénomique, en bénéficiant de l'environnement de Genopole® à Évry.

L'informatique est en première ligne pour aborder le défi épigénomicien (près de la moitié de nos membres sont chercheurs en informatique) et le Programme d'Épigénomique est un lieu où tout informaticien peut trouver facilement des applications passionnantes et les outils pour se former et développer des projets de qualité.