Biologie systémique

C. Gautier 29 avril 2014

La biologie systémique

La plus grande part de ce chapitre repose sur J.M. Legay [1,2,3] et sur le colloque CNRS « la Notion de Système dans les sciences » qui s'est tenu en 1980 à Lyon en appui sur le GRECO (une sorte d'ancêtre aux GDR pour les plus jeunes) « Analyse de systèmes » [4]. On peut trouver de nombreux textes plus récents qui abordent des aspects plus théoriques de la systémique. Il existe en particulier des sociétés savantes qui sont consacrées à la systémique[1] et qui ouvrent vers d'autres concepts que ceux qui sont présentés ici.

Il est apparu très tôt que le cerveau humain était très peu efficace pour appréhender des objets dont les constituants entretenaient entre eux des relations multiples[2]. La première démarche méthodologique explicite pour surmonter notre infirmité est celle de Descartes (en 1637 dans le discours de la méthode) « découper la réalité en autant de parties qu'il était nécessaire pour que l'objet retenu pour l'étude soit simple » [3]. L'un des exemples les plus fameux (exemple que des générations d'étudiants ont subi!) de l'échec de la démarche réductionniste de Descartes est la dynamique des populations de lynx et de lièvres et sa représentation par le modèle de Lotka-Volterra (1925)[4]. L'observation d'oscillations dans les tailles de populations des deux espèces ne peut pas se déduire de l'examen séparé des deux espèces. Une explication plausible est fournie par un modèle dans lequel le nombre de rencontres entre un lynx et un lièvre augmente la mortalité des lièvres et la reproduction des lynx. Cette simple relation entre les deux constituants du système suffit à générer l'évolution des tailles de populations observées. Si la relation entre les deux populations (lièvres et lynx) constituant ce système n'est pas prise en compte le comportement du système ne peut pas être compris (même si le nombre total d'animaux est la somme exacte des tailles de population des deux constituants, ce qui montre bien que la complexité d'un système dépend avant tout de la question posée et pas de l'objet que l'on étudie).

La notion de système (dans le sens de la systémique) est due à un biologiste L. von Bertalanffy (General system theory 1968)[5]. La « théorie » du système a ensuite évolué dans plusieurs directions, je ne retiendrais que celle désignée par Jean Louis Le Moigne comme « l'analyse de système considérée comme une science de l'ingénieur pour la science des méthodes d'aide à la modélisation, d'aide à la conception, d'aide à la décision, voire d'aide à la l'intelligence de la complexité ». On retrouve tout à fait ici la philosophie de JM Legay dans ce qu'il appelle la méthode des modèles, le modèle apparaissant comme un instrument pour aider à comprendre le système. La notion de science de l'ingénieur introduite par JL Le Moigne me semble très importante, y compris pour insérer correctement la démarche dans une politique scientifique plus large.

L'idée centrale est en effet de résoudre un problème, répondre à une question, les éventuels progrès disciplinaires ne sont que des sous-produits. Il est important de souligner que **la notion de discipline est absente de la systémiqu**e. Plus exactement la question posée voire le point de vue de celui qui la pose peut être liée à une discipline : on peut s'intéresser à la dynamique des lynx et des lièvres avec un point de vue écologique ou économique dans le cadre du commerce des peaux. Par

1 L'AFSCET (association française de science des systèmes : www.afscet.assoc.fr) est une porte d'entrée vers plusieurs associations au niveau européen ou international. On peut trouver sur son site des présentations intéressantes sur la systémique (SystemicApproach.pdf , ressystemica/sysmimat1.html). Ces présentations ne sont pas spécifiques de la biologie et restent dans leur ensemble plus pédagogique ou théorique qu'opérationnelles.

2 Ici les mots « objet » et « constituants » n'ont aucune définition précise et encore moins formelle.

3 Voir une discussion dans JM Legay [1] p14

4 Voir une description dans : https://interstices.info/jcms/n_49876/des-especes-en-nombre

5 On trouve la notion de système dans sa thèse datée de 1937

contre les disciplines sollicitées dans les phases suivantes de l'analyse du système ne sont pas déterminées a priori . On reviendra sur cet aspect dans les relations entre biologie systémique, recherche fondamentale et recherche appliquée.

Même si dans les domaines scientifiques il n'y a jamais de démarche totalement répétitive, la biologie systémique s'appuie sur un schéma assez simple :

***** 1)** comme dit plus haut on commence par poser une question.

***** 2)** On délimite ensuite le système nécessaire pour répondre à cette question. La difficulté est que en biologie tout est relié à tout (ne serait-ce qu'en passant par LUCA, « the last universal common ancestor »). La délimitation du système est donc une démarche réductionniste, explicite et donc pouvant être remise en cause en particulier dans l'étape 6 ci-dessous. Cette réduction porte à la fois sur l'objet lui même et sur les connaissances existantes sur cet objet. C'est au cours des étapes 1 et 2 que le chercheur décide de la complexité ou pas du système qu'il va étudier.

***** 3)** Le système est décrit dans un langage formel, en pratique les mathématiques au sens large (incluant l'informatique, la physique théorique, ...). Cette réécriture formelle du système constitue le modèle. Le modèle dépend du couple question/système et pas uniquement du système. Par exemple en fonction de la question le modèle peut ne rendre compte que du comportement global d'un constituant du système ou au contraire en décrire aussi précisément que possible le fonctionnement.

***** 4)** On utilise les mathématiques pour déduire du modèle les propriétés que notre cerveau n'est pas capable de trouver sans aide. Il est important de bien reconnaître que dans cette étape on ne fait que retrouver ce que l'on a mis dans le modèle, même si cela demande un effort théorique parfois insurmontable. Ceci étant l'analyse peut mettre en évidence des conséquences des hypothèses sur le comportement du système contre-intuitives. Avant que cela devienne un comportement enseigné dans dans les premières années de l'université les oscillations lynx/lièvres n'apparaissait pas à l'écologiste comme la conséquence naturelle d'une dépendance prédateur/proie. De même j'appartiens à l'immense majorité des chercheurs pour lesquels les systèmes dont le comportement présentent des sensibilités extrêmes à certains paramètres (chaos, …) ne repose pas sur une intuition réelle (de là l'idée que leur comportement résulte d'un mécanisme d'émergence, idée sur laquelle je reviendrais plus loin pour tenter d'en démontrer les dangers).

***** 5)** Il convient ensuite de revenir du modèle au système, c'est à dire à la biologie. Dans le cas d'une recherche non strictement finalisée le résultat le plus important est souvent l'émergence d'une idée nouvelle voire d'une nouvelle voie de recherche. L'intégration des connaissances du moment dans un modèle formel, « jouer » avec ce modèle par exemple en en manipulant les paramètres sont des moteurs efficaces à l'innovation scientifique. La réflexion sur la délimitation du système peut mettre en évidence des facteurs négligés à tort par « habitude » ou car leur accessibilité expérimentale est difficile. Par ailleurs le comportement du modèle, sa sensibilité à certains paramètres peuvent suggérer de nouveaux mécanismes fonctionnels.

En dehors de cet aspect exploratoire, la première confrontation est celle entre le comportement du modèle tel que déduit de son analyse mathématique et les connaissances biologiques. Une incompatibilité permettra de détecter une erreur dans la modélisation depuis le fait d'avoir négligé des facteurs importants dans l'étape de délimitation du système jusqu'à une mauvaise représentation formelle de ce système. C'est souvent au travers de la réanalyse de ces modèles incohérents avec les connaissances du domaine que les progrès dans la compréhension du fonctionnement du système sont réalisés.

Si aucune incongruence majeure entre le modèle et les connaissances du domaine n'apparaît, le modèle peut être utilisé comme un outil d'**aide à la planification d'expérience**. Le modèle en tant qu'aide à la planification expérimentale est sans doute l'aspect le plus important de la biologie systémique. En dehors de quelques rares expériences d'exploration massive, tout plan expérimental

devrait être le prolongement d'une modélisation. Pour donner un exemple des liens entre modèle et expérimentation on peut citer la mise en évidence des paramètres auquel le modèle est le plus sensible et dont l'estimation et le contrôle seront nécessaire pour comprendre le fonctionnement mais également pour le perturber. Il faut réellement insister sur cette importance du dialogue entre modélisation et expérimentation, la stratégie qui consiste à réaliser une expérience puis aller voir un modélisateur et espérer une publication après cet aller simple est à bannir vigoureusement[6] !

On peut enfin citer une approche plus récente dans ce contexte parfois décrite comme « **biologie synthétique** ». Si une bonne représentation d'un système est atteinte une validation intéressante est de tenter de le reconstituer à partir d'objets aussi élémentaires que possibles. Les premières approches à grandes échelles se trouve probablement dans le domaine de l'écologie avec le développement des écotrons. Ces plates-formes expérimentales fournissent les outils de construction, de suivi et de contrôle d'écosystèmes complets. La pression pour une reconstruction d'écosystèmes écologiques a été renforcée par la grande difficulté d'expérience in natura pour des raisons évidentes d'éthique environnementale. Plus récemment la biologie synthétique a permis la construction de systèmes cellulaires, parmi les plus spectaculaires on peut citer la réassignation de codons pour faire synthétiser à des bactéries des protéines possédant des acides aminés non « conventionnels ». Cette démarche nécessite la reconstruction partielle de la machinerie de traduction. Ceci étant les « objets » ainsi reconstruits sont des modèles « biophysiques » certainement efficaces pour la compréhension des systèmes biologiques, voir utile par leur production. Il ne sont pas des substituts aux systèmes biologiques et leur bon fonctionnement ne peut être considéré comme une preuve de la parfaite compréhension du fonctionnement et de l'évolution du système biologique étudié.

***** 6)** Soit on a obtenu une réponse satisfaisante à la question initiale et l'analyse s'arrête soit on réinitialise le processus à la première étape en prenant en compte ce qu'à appris le précédant modèle. Ceci constitue ce qui est souvent désigné comme le cercle vertueux de la modélisation.

L'objectif n'est pas ici de faire un cours de modélisation ! il me semble cependant important de discuter de quelques aspects particuliers de l'analyse de systèmes.

La biologie systémique est souvent présenté comme l'étude des **systèmes complexes**. Le mot « complexe » est très ambiguë, dans certaines disciplines une définition formelle lui a été donnée (par exemple la complexité algorithmique) mais ces définitions ne peuvent pas s'appliquer à un « système » lui même n'ayant pas de définition formelle au sens des mathématiques. Des définitions du type « le comportement du système n'est pas la somme des comportements de ses parties » sont parfois proposées mais mais avec une notion très ambiguë de « somme » . JM Legay caractérisait un système complexe par le fait que le retrait d'un quelconque de ses constituants en changeait la nature voire lui faisait perdre sa qualité de complexe. Cette définition à le grand intérêt d'être opérationnelle et de conduire à la réduction d'un système à son « noyau » complexe. Le deuxième point sur lequel JM Legay insistait, et je le rejoint totalement sur ce point, est que la complexité est une décision du modélisateur. Cette décision dépend à la fois de la question posée et des connaissances disponibles sur l'objet biologique. Il serait évidemment contre productif de ne pas reconnaître la « simplicité » d'un système. Une voie classique de « simplification » d'un système est l'encapsulation d'une partie complexe du système quand seule son interface avec le reste du système est importante pour la question posée (c'est une démarche très classique dans les modèles à compartiments) : évitons la complexité si la simplicité suffit !

6 D'ailleurs certain modélisateurs refusent, à juste raison, de prendre en compte des expériences à la conception desquelles il n'ont pas participé.

L'existence de **propriétés émergentes** est souvent signalé comme centrale dans la démarche de biologie systémique. Cette notion d'émergence me semble dangereuse dans le dialogue biologiste-mathématicien. L'émergence suggère une dynamique, la propriété n'existait pas avant et elle émerge. Cela renvoie assez naturellement le biologiste à des notions d'évolution, comment a émergé le découpage des gènes en exons, comment a émergé la reproduction sexuée, … Dans le cas de l'analyse de système rien de tel : la propriété est totalement écrite dans le modèle, simplement notre cerveau, sans une analyse mathématique parfois difficile, ne sait pas la reconnaître.

Le dernier point que je voudrais discuter est l'**aspect globaliste** (ou holiste) souvent associé à la biologie systémique face au réductionnisme cartésien. Il est clair de l'intérêt majeur de la modélisation est la possibilité de prendre en compte les relations entre les constituants d'un système. Ainsi bien que de nature très différente on peut modéliser simultanément un réseau de gènes et un réseau métabolique, modéliser à la fois les caractéristiques des mutation liées aux propriétés des processus de réparation et l'évolution de ces mutations au sein d'une population. Ceci étant l'étape centrale de l'approche systémique est la délimitation du système qui est une étape fondamentalement réductionniste. L'important est que la réduction nécessaire du système est explicite et fait partie du travail de modélisation. Il est donc possible de la remettre en cause, de la modifier, d'identifier les expériences qui permettrait qu'elle soit moins drastique , ...

Panorama rapide de la biologie systémique en 2013

Si on interroge le réseau sur la définition de la biologie systémique les deux premières définitions que l'on trouve sont les suivantes :

« Aujourd'hui, cette approche réductionniste est remise en cause par la renaissance de la biologie systémique, qui a pour objets d'étude les systèmes formés par les produits de gènes en interaction »(*Biologie systémique Des concepts d'hier aux découvertes de demain. A-R Carvunis et al ; MEDECINE/SCIENCES 2009 ; 25 : 578-84*)

"Systems biology is the science of discovering, modeling, understanding and ultimately engineering at the molecular level the dynamic relationships between the biological molecules that define living organisms."*Lee Hood, President and Co-founder, Institute for Systems Biology*

Si la biologie systémique s'est principalement développée dans le cadre général de l'écologie avec en particulier de nombreux travaux en génétique et en dynamique des populations ces deux définitions du réseau ci-dessus correspondent à une forte part des activités qui se réclament actuellement de la biologie systémique. L'écologie utilise plus la notion de modèle que de systémique même si, comme discuté plus haut, ces deux notions se superposent dans le cas de la biologie au moins. Les travaux utilisant la biologie systémique ont été profondément influencés ces dernières années par les avancées technologiques ouvrant le niveau moléculaire à l'expérimentation. En particulier les données de séquençage massif mais aussi celles relatives à l'interactome entre protéines et au métabolome ont conduit à une nouvelle vision des interactions entre des constituants cellulaires. Cela s'est traduit par l'émergence de nombreux réseaux d'interactions (génique et transcriptomique par exemple). L'estimation de ces réseaux a impliqué la mise en place de nouvelles modélisations mathématiques ou informatique. Le couplage de ces différents réseaux est un des principaux défis que pose, à la recherche actuelle, cette explosion technologique. L'apparente appropriation du terme de « biologie systémique » par le niveau moléculaire et cellulaire résulte, de manière assez naturelle, de l'ouverture récente à cette approche (on avait vécu il y a quelques années le même phénomène avec le terme « bioinformatique »)

Cependant les autres domaines de la biologie ont également subi un choc technologique majeur. En écologie la disponibilité des cartes satellitaires, des mesures en continu de paramètres du milieu ou de l'état physiologique d'un animal, le repérage par des balises Argos de plus en plus miniaturisée des suivis individuel d'animaux jusqu'à des taille de l'ordre d'un oiseau marin ont fourni des données d'un volume comparable à celui du séquençage massif. Les capteurs de contacts permettent de construire le graphe dynamique des proximités entre individus d'une population ouvrant par exemple la voie à des modèles épidémiologiques intégrant de tels graphes dynamiques. Au delà des masses des données disponibles, leur diversité d'échelle et de nature rend de plus en plus nécessaire les outils d'intégration que fournit la systémique.

En effet il faut insister sur le fait que depuis quelques années les questions posées sont de plus en plus **intégratives** au sein de la biologie. C'est une phénomène nouveau, il y a moins d'une dizaine d'année les champs disciplinaires au sein de la biologie était globalement bien définis. Le découpage du CNRS en deux blocs l'un autour de l'écologie caractérisé par des relations privilégiées avec SHS, la physique du globe, l'agronomie et un deuxième blocs ou la « biologie moléculaire » se rapprochait de la santé, des physique et de la chimie reflétait une partition bien acceptée par une majorité de la communauté des « biologistes ». Ce découpage posait bien le problème de l'évolution biologique revendiquée par lesdeux instituts, mais cela semblait le problème majeur de ce découpage. Je ne critique pas la création des deux instituts de biologie du CNRS je trouve que cette création est un bon indice de l'évolution récente de la nécessité

d'intégration en biologie. Je vais donner ici 3 exemples de questions transversales à la biologie et impliquant les deux instituts du CNRS et qui ont par ailleurs des relations avec SHS.

La **génomique environnementale** est un premier champ reconnu par le CNRS[7] qui s'appuie sur un large spectre de champs disciplinaires au sein de la biologie. Je cite ci dessous le cahier prospective (oct 2013) de l'INEE qui le défini :

« La génomique environnementale regroupe l'ensemble des connaissances acquises sur les organismes et écosystèmes présents et passés par l'analyse de la séquence des gènes, génomes, métagénomes, transcrits, transcriptomes et métatranscriptomes. Ainsi, en combinaison avec d'autres technologies et observations, la génomique environnementale informe sur la taxinomie et la diversité des organismes actuels et fossiles (individus, populations, communautés), leur phylogénie et évolution, leurs potentialités et capacités d'adaptation et d'acclimatation, leur biologie, leurs traits fonctionnels, et leurs interactions avec l'environnement dans ses dimensions biotique et abiotique. »

Je compléterais juste ce thème en remarquant que les connaissances de plus en plus précises sur le paysage génomique tel qu'il résulte des mécanismes de gestion du génome (réparation, recombinaisons, insertion/délétion), des différentes formes d'information génomique (génes protéiques, RNA de diverses tailles et fonctions, épigénétique, …) et de l'action des mécanismes évolutifs dans les populations offre à la biologie systémique un terrain de jeu exceptionnel pour un enrichissement mutuel de raisonnements concernant la structure et le fonctionnement au niveau moléculaire, les phénomènes adaptatifs au sein des populations et l'évolution à grande échelle de temps.

Les **relations entre un « hôte » et des « partenaires » microbiens et/ou viraux** ont me semble-t-il une identité scientifique et sociétal qui, bien qu'il appartiennent à la génomique environnementale, justifie une description spécifique. On parle donc ici de microbiome, de symbiose, de pathogène et ceci à toutes les échelles de temps et d'espace. Les systèmes en cause peuvent impliquer de très nombreuses espèces microbiennes, des symbioses imbriquées comme des poupées russes. Les dialogues moléculaires entre partenaires nécessitent des modélisations conjointes des fonctionnements cellulaires des partenaires, par exemple comment l'interaction à conduit à la modification des réseaux métaboliques de l'hôte et de son ou de ses symbiontes. Les données d'interactome protéique entre hôte et symbiote (pathogène ou non) constituent également une voie d'accès à ce dialogue. Les questions relatives à ce thème interpellent bien entendu l'infectiologie, l'épidémiologie mais aussi la biologie cellulaire, l'évolution et finalement l'ensemble de la biologie. Enfin l'interaction avec la santé directement dans le cas de pathogènes ou indirecte dans le cas du microbiome crée des liens avec la médecine mais aussi SHS. Ce qui fait le lien avec le dernier thème que je souhaite citer.

Le dernier thème que je veux mettre en évidence est la **santé globale** à laquelle j'associerais la **santé environnementale**. L'université de Lyon ayant sous l'impulsion de J. Comby une initiative dans ce domaine je n'ai pas réellement osé parlé ici de ce domaine que je ne connais qu'au travers de la conversation que j'ai eu avec J. Comby. Je souligne simplement ici l'importance de ce champ et l'enthousiasme communicatif de J. Comby. Je suis aussi pleinement d'accord avec son point de vue qu'il faut laisser aux acteurs de la recherche le soin d'en définir les contours et plus travailler sur l'environnement de la recherche, je soutiens cette idée également pour la biologie systémique.

7 http://www.cnrs.fr/fr/pdf/inee/inee-genomique-environnementale/

Esquisse pour un projet

Deux parties s'inscrivent naturellement dans ce dernier thème : les points forts régionaux sur lesquels peut se bâtir un projet et le projet lui-même. Pour le projet sa définition comporte la difficulté de ne pas porter sur une thématique biologique mais sur une méthodologie. Je n'ai aucune légitimité pour restreindre le champ biologique sur lequel devrait prioritairement porter la biologie systémique régionale (en plus mon passé ne pourrait que rendre partiale une une telle restriction!). Je me contenterais donc dans cette deuxième partie de discuter des stratégies qui me semblent favorables au développement de la biologie systémique. Mon penchant écologique me pousse à croire que si on crée une niche favorable, cette niche ne restera pas vide : pour attirer des activités nouvelles de haut niveau en biologie systémique il est essentiel de créer et promouvoir un environnement en terme de formation, de recherche, de plates-formes technologiques, d'interaction recherche/applications lui-même de haut niveau. Cela est d'autant moins difficile que, comme va le montrer la première partie de ce chapitre, la région est déjà bien armée dans ces domaines. Dans une deuxième partie je discuterais quelques pistes pour renforcer cet environnement.

1) les points forts régionaux :

- Un aspect historique avec deux personnalités fortes Gustave Malécot qui a profondément marqué les mathématiques de la génétique des populations et Jean-Marie Legay pour sa formalisation de la démarche d'analyse des systèmes biologiques au travers de la modélisation mais aussi pour la mise en œuvre de cette politique dans la création d'un laboratoire qui sur 40 années d'histoire a conservé comme axe central la modélisation des systèmes biologiques.
- Un enseignement des mathématiques aux biologistes qui est tourné vers la modélisation depuis près de 40 ans avec entre autre un DEA « analyse et modélisation des systèmes biologiques » et qui au fils du temps s'est modifié et enrichi par exemple d'une filière de licence et de master MIV (Modélisation et Informatique du Vivant), d'une filière ingénieur de l'INSA BIM (Bioinformatique et Modélisation) (à compléter avec les enseignements que je connais mal en particulier à l'ENS, le master « modélisation des systèmes complexes » a évidemment sa place ici). Je reviendrais cependant sur le futur de cet enseignement dans la partie projet.
- Un effort de fédération de la biologie systémique sur Lyon à l'initiative de Chantal Rabourdin-Combe qui a conduit à BioSyl (O. Gandrillon) et a montré la très grande diversité de l'intervention de la biologie systémique sur Lyon.
- Un investissement régional très fort de l'INRIA dans l'interface (informatique, mathématique) / Biologie. La région compte actuellement 5 équipes INRIA dans cette interface (Dracula, Bamboo, Beagle, Ibis, Numed) une équipe internationale (Erable) est en cours de formation (la première de ce type à l'INRIA) et l'INRIA est partenaire avec le CNRS et UCBL d'un LIA (Laboratoire international associé LIRIO).
- Des plates-formes qui apparaissent naturellement comme un support technologique important à la biologie systémique :
    - le Centre européen de résonance magnétique à très hauts champs (CRMN) a une

activité en biologie en particulier dans le cadre de la métabolomique.

- En acceptant une nouvelle extension de la notion de région le pôle Ecotox de Rovaltain est une plate-forme à échelle européenne permettant la reconstitution de milieu dans le cadre de la santé environnementale. Cette plate-forme est lié à la biologie systémique régionale à la fois par son aspect de biologie synthétique mais aussi par sa liaison forte à la notion de santé globale. Elle pourrait plaider pour la promotion d'une activité forte en écotoxicologie dans la région qui impliquerais bien entendu les universités Claude Bernard et Joseph Fourier mais aussi l'IRSTEA (ex CEMAGREF). Dans le cadre des écotrons il faut aussi citer l'éco-aquatron plate-forme régionale piloté par IFR41 (Lehna).
- Une diversité de ressources de calculs permettant à la bioinformatique, complément souvent indispensable à une démarche de biologie systémique, de répondre efficacement à des problèmes variés. Le CC-IN2P3 s'est largement investi dans les aspects biologiques, il offre à la fois une très forte puissance de calcul adaptée au parallélisme par les données et une très grande capacité d'archivage. Le PRABI avec ses 5 sites offre une puissance de calcul intermédiaire avec des machines adaptés à certaines tâches classique du traitement des NGS (assemblage par exemple demandant de grandes tailles mémoire). Le PRABI est à la fois intégré dans une structure régionale (FLMSN fédération lyonnaise de modélisation et sciences numériques) et une structure nationale : France Génomique.
- Les bases de données (et de connaissances) constituent une composante très importante de l'environnement nécessaire à la systémique. La sollicitation de multiples disciplines impose un accès simple et rapide à de multiples connaissances, y compris hors du champ habituel de compétence du modélisateur. Les laboratoires de la région ont développé de nombreuses bases de données à visibilité internationales. La section « Databases » du PRABI (www.prabi.fr) en cite la plupart, il faut cependant ajouter Virhostnet dans ses deux versions privée et publique qui fournit les données d'interactome protéique entre virus et hôtes. Enfin il faut citer le projet européen eTRICKS (utilisant le système transMART) qui vise à développer un ensemble logiciel qui pourrait être un outil intéressant pour gérer, associer et requêter des données provenant de domaines scientifiques différents sur des problématiques médicales. On retrouve associé ou intégré des outils que l'on connaît bien en modélisation comme R (logiciel de programmation adapté aux statistiques et à la modélisation) ou Galaxy (un logiciel permettant des gestions simples de connaissances méthodologiques). Je n'ai pas trouvé d'articles théoriques d'informatique sur transMART je ne sais donc pas s'il repose sur des concepts innovant (il me semble qu'il repose sur postgreSQL et donc que le modèle de données est sans doute ancien). L'évaluation menée actuellement par l'EISBM de l'utilisation de eTRICKS (implanté sur le CC-IN2P3) dans le cadre d'inflammation pulmonaire (projet U-BIOPRED) est donc très intéressante et pourrait avoir des retombées méthodologiques au delà de la médecine. Le développement de bases de données hétérogènes est nécessaire dans de nombreux domaines, parmi les efforts actuels dans ce champ on peut citer comme exemple des approches en métagénomique couplant données environnementales et génomiques.

- Enfin le dernier point important est que la région a su créer, grâce à des reconnaissances nationales ou internationales, un ensemble cohérent de structures. Au niveau de la recherche les nouveaux labex, en particulier Ecofect et IMMU mettent clairement l'accent sur des pluridisciplinarités et sur la modélisation. Par ailleurs Lyon1 et Lyon3 sont acteurs d'un

DIPEE[8] qui dans le cadre de l'écologie globale est évidemment lié à la biologie systémique. Les applications de la recherche en biologie sont largement soutenues dans le contexte régional. Il faut en premier lieu citer les volets médicaux des universités qui font partie des plus grandes universités médicales de province et qui ont tissés des liens étroits avec la modélisation (sans doute plus particulièrement à Lyon). De même les plates-formes qui ont été citées précédemment ont, bien entendu, des activités vers les entreprises. Enfin le rôle majeur de Lyon Biopole dans cette relation vers l'application de la recherche est un atout régional fort auquel se joint sur un plan plus opérationnel des structures en émergence Bioaster et l'institut européen de médecine systémique (EISBM).

2) quel environnement pour la biologie systémique dans la région

Le bilan précédant concerne l'environnement scientifique de la biologie systémique. L'examen des descriptions des équipes rattachée à BioSyl donne une vision, principalement lyonnaise, de la diversité des questions scientifiques abordées au travers d'une démarche systémique. Un environnement scientifique favorable au développement de la biologie systémique me semble pouvoir être présenté autour de 3 thèmes :

- L'enseignement et la formation
- la transdisciplinarité
- Les applications de la recherche

2.1 enseignement et formation :

Le rôle croissant de la systémique en biologie et l'impact grandissant de la modélisation mathématique a nécessairement un impact fort sur le métier de biologiste. Cette nécessité pour le biologiste d'une maîtrise raisonnable des mathématiques résulte aussi plus simplement de la nature même de l'expérimentation de plus en plus éloignée d'une observation « directe » d'un phénomène. La physique, la chimie, la bioinformatique qui s'intercalent entre le plan expérimental et les résultats induisent des biais qui sont loin d'être intuitifs. On ne peut plus ignorer que le résultat, dit brut, d'une expérience est le résultat d'une modélisation. Par ailleurs le coût (en temps, en moyen humain et en argent) d'une expérience biologique est également en croissance et faire une expérience « pour voir » sans planification expérimentale rigoureuse est toujours une perte sans profit. Il est donc évident qu'il faut adapter les formations à ces nouvelles évolutions de la biologie, que ce soit au niveau des formations initiales que continues.

Comme décrit dans le paragraphe 1) un effort important a déjà été réalisé dans la région en faveur d'une forte interaction au niveau de l'enseignement entre biologie et mathématique (toujours avec le sens étendu à l'informatique) au niveau de la formation initiale universitaire et ingénieural mais aussi au niveau de la formation continue. Il me semble cependant que trois aspects pourraient être amélioré. D'abord la visibilité au niveau master est bonne mais souffre de complexité liée aux structures administratives : plusieurs masters ou options insérées dans d'autre masters, un aspect apparemment disciplinaire ce qui paraît un peu contradictoire avec la transdisciplinarité de la démarche systémique, des noms pas toujours très clairs. Il me semble qu'un unique master « biologie systémique » ouvert sur toute la biologie au travers de modules partagés avec d'autres masters (en écologie, biologie cellulaire et moléculaire, santé voire SHS) et situé au niveau de l'université de Lyon (et donc partagé par l'ensemble des établissements) pourrait plus facilement obtenir une reconnaissance nationale et internationale. Le niveau master est évidemment essentiel dans la construction d'un environnement scientifique puisque porte d'entrée vers les doctorats dans nos laboratoire il est aussi un fort argument pour faire venir des personnalités scientifiques

8 Voir le site du CNRS pour la définition des DIPEE (http://www.cnrs.fr/inee/recherche/dipee.htm)

« fortes » dans la région. Le deuxième aspect, également déjà présent mais qui pourrait être renforcé, est la participation d'enseignements utilisant le web. Le développement de cours Web paraît bien adapté à une démarche d'étude de cas avec des interventions de plusieurs « enseignants » de spécialité différentes, il serait également très favorable à une relation plus efficace entre formation initiale et continue. Les enseignements en présentiel pourraient alors être plus consacré aux liens modélisation / expérimentation. Par ailleurs un tel enseignement peut avoir une visibilité très grande. Le dernier point dont je voudrais souligner l'importance est la poursuite d'un enseignement intégré, dès la première année de licence. La modélisation ne doit pas être uniquement présentée dans les enseignement de mathématique, c'est bien entendu déjà le cas sur Lyon, mais cette perméabilité entre les disciplines me semble pouvoir être encore améliorée. Une dernière remarque, la plus importante sans doute, l'évolution de l'enseignement ne doit en aucun cas ne reposer que sur la bonne volonté des enseignants chercheurs. Cela demande des moyens très importants, au moins autant que la recherche. La qualité de l'enseignement dépend aussi (c'est une évidence parfois oubliée) de la qualité de la recherche réalisée par les enseignants/chercheurs. Ils doivent donc disposer de temps pour s'impliquer dans des enseignements nouveaux, temps qui ne peut pas être pris sur leur temps de recherche.

2.2 transdisciplinarité :

Le développement de la démarche de systémique en biologie a clairement montré que les mathématiques, l'informatique et la biologie pouvaient s'enrichir mutuellement. La biologie apporte des problèmes nouveaux aux disciplines formelles qui peuvent y trouver de questions de recherche propre à leur discipline. Ceci étant ce dialogue n'est en pratique pas toujours simple à mettre en place. Il faut bien admettre qu'il existe peu d'exemples de résultats majeurs en informatique ou en mathématique qui soit issus d'une question biologique. Si on élimine l'informatique dite « inspirée de la biologie » qui n'a eu qu'un lointain rapport avec la biologie et relève souvent de processus d'optimisation stochastique (les algorithmes génétiques en est un exemple typique), on trouvera principalement des résultats de théorie des graphes et de complexité algorithmique. Par contre la biologie a largement bénéficié de la recherche en informatique et en mathématiques et a fourni aux informaticiens et mathématiciens des applications motivante (enfin je le crois!) à leur recherche. Il faut donc admettre qu'il reste une certaine dissymétrie entre disciplines. C'est probablement la raison pour laquelle je crois important que la pluridisciplinarité s'organise autour de relations fortes et durables entre mathématiciens (au sens large) et biologistes. En terme de structures je plaide pour deux stratégies complémentaires et bien représentées dans la région. La première est la constitution de laboratoires mixtes dont la thématique est fixée par la biologie. Les « théoriciens » peuvent s'approprier la thématique biologique, participer ainsi de manière efficace à la modélisation, aux planification expérimentales et, y compris, à la définition des questions biologiques. Le prix à payer est une très faible probabilité pour eux d'obtenir une Medal Fields ou un Turing Award ! Une autre stratégie complémentaire est celle pratiquée par l'INRIA d'équipes projets associant des biologistes et des informaticiens. Ces équipes permettent une collaboration étroites entre disciplines tout en laissant les chercheurs dans leur environnement disciplinaires. Cinq telles équipes existent en Rhône Alpes (environ 20 % des équipes INRIA au niveau national sont transdisciplinaires avec la biologie) montrant l'importance de la biologie systémique pour cet EPST. Enfin il faut reconnaître que de nombreuses questions biologiques restent (heureusement) relativement classiques en terme de modélisation et peuvent rapidement devenir très consommatrice en temps pour les collègues mathématiciens qui acceptent de se prêter au jeu des consultations. Si l'objectif est de rendre la démarche systémique la démarche naturelle en biologie il faut que l'accès à la modélisation soit le plus simple et rapide possible pour les biologistes. L'existence de structures ingénieurales intermédiaires entre la recherche en biologie et la recherche en mathématique et informatique me semble nécessaire. De telles structures existent (par exemple dans certains sites du PRABI) mais nécessiterais un renforcement et une plus grande diversité de compétences biologiques.

Une des conséquences de la biologie systémique est de modifier le point de vue sur l'expérimentation en l'intégrant au sein de la démarche de modélisation. Une nouvelle fois il faut insister sur le fait que le point de départ n'est pas constitué des données mais de la question posée et de la délimitation du système qui en découle. Une conséquence d'une conception expérimentale résultant d'une démarche de modélisation peut être l'émergence de plans expérimentaux moins classiques que lorsqu'ils se fondent principalement sur la disponibilité des techniques. Cela me semble plaider pour le renforcement de plates-formes complémentaires des grandes plates-formes « massives », des plates-formes plus versatiles pilotés par des ingénieurs de haut niveau capable de mettre en place des plans expérimentaux originaux. Parmi ces démarches devraient se trouver des possibilités de développements de types biologie synthétiques.

Une dernière conséquence de la transdisciplinarité est la nécessité de pouvoir consulter rapidement des connaissances, y compris méthodologique dans des domaines autres que son propres domaine de compétence. La disponibilité des bases de connaissances adaptée est nécessaire, leur développement reste à mon avis du domaine de la recherche en informatique. Il faudrait en discuter avec François Sillion (nous étions d'accord pour nous rencontrer mais cela ne s'est pas encore fait) et également avec François Rechenmann.

2.3 dialogue recherche questionnement sociétal :

Comme signalé dans le paragraphe précédent la région s'est dotée d'une interface solide entre la recherche et ses applications en particulier dans le domaine de la santé. Il faut cependant remarquer que l'effort porte majoritairement sur l'aspect infectiologie et que les autres relation de la recherche et des ses applications reposent très majoritairement sur l'université de Lyon et ses partenaires. Ceci étant Bioaster, dans son domaine de compétence sera une expérimentation qui pourra éclairer sur les difficultés de l'introduction de la biologie systémique dans les entreprises biologiques, probablement les deux difficultés majeures seront une formation continue à améliorer et la preuve que les modélisateurs ne sont pas là pour traiter des données mais pour aider à concevoir des plans expérimentaux voire à aider à formuler les questions.

Conclusion (on m'a demandé de faire rêver … j'y arrive!)

Depuis l'aube de l'humanité notre espèce à modifié le monde vivant à notre avantage. La naissance de l'agriculture est une marque historique essentielle dans notre histoire. Cependant jusqu'à récemment la biologie est restée majoritairement une science de l'observation et de la compréhension. La conception même des médicaments reste pour une part importante un criblage d'ensemble de molécules parfois plus choisies sur le coût d'une éventuelle mise sur le marché que sur des potentialités scientifiquement fondées. Cette situation me semble pouvoir évoluer rapidement vers une biologie ou la conception deviendra un troisième axe aussi important que l'observation et la compréhension. Déjà un écosystème, comme un sol ou un microbiome animal (y compris humain), peut être considéré comme un système que l'on peut observer au niveau de ses capacités (résistance à des maladies des plantes, capacité de tétoxification pour un sol, résistances aux pathogènes ou aux maladies non infectieuses pour le microbiome …), de ses fonctions biologiques (étude du métatranscriptome) ou enfin de sa biodiversité (métagénome). La conception d'un sol ayant des propriété spécifique et sa réalisation peuvent sans doute constituer des projets scientifiques crédibles. De même on sait concevoir et « fabriquer » une machinerie de traduction bactérienne capable d'associer à une codon particulier un acide aminé non conventionnel et coder cette machinerie dans le génome d'une bactérie qui fabriquera des protéines jamais fabriquées auparavant par aucun être vivant. Les progrès de nos connaissances sur l'évolution permettent

maintenant de disposer d'arbres phylogénétiques d'espèces relativement fiable de même que d'arbres phylogénétiques de gènes et sans doute rapidement de domaines protéiques. Ces approches ouvrent de plus en plus les portes à la compréhension des mécanismes d'émergences de nouvelles fonctions, ainsi on connaît de plus en plus de fonctions propres aux eucaryotes qui utilisent des protéines ayant des homologues bactériens. L'association de la combinatoire de réutilisation et des mécanismes d'adaptations et leur impact sur les innovations évolutives seront de mieux en mieux comprise et pourront sans doute servir de modèle à la conception de nouvelles fonctions. Dans la plus part des cas il faudra prendre en compte des interactions qui nécessiteront la manipulation de systèmes complexes. La réduction de ses systèmes (je ne crois pas aux « gros » systèmes !) leur modélisation mathématiques seront nécessaire à leur compréhension et à la conception d'expériences efficaces. Il n'est pas pensable que la biologie systémique puisse rester étrangère à cette révolution. Par ailleurs un lien constant avec les sciences de l'homme et des sociétés est indispensable à la fois pour porter un regard extérieur sur les aspect éthiques, sur les risques potentiel et aussi sur la capacité des sociétés à accepter ou pas une biologie « créatrice ». Enfin il ne faut pas confondre la demande naturelle de la science à progresser avec les demandes sociétales. Je ne suis pas certain qu'il existe des exemples ou ces deux demandes aient été synchrones.

Bibliographie :

[1] JM Legay L'expérience et le modèle (1997) 111p Coll. Sciences en questions, INRA édition

[2] JM Legay, A-F Schmid Philosophie de l'interdisciplinarité (2004) édition PETRA

[3] JM Legay La méthode des modèles, état actuel de la méthode expérimentale (1973) Informatique et Biosphère, p7-73