MODÉLISATION EN BIOLOGIE VÉGÉTALE

INTRODUCTION

par François **Houllier**[1]

Pour cette session consacrée à la modélisation en biologie végétale, nous avons retenu une acception large de la biologie végétale en considérant, de la graine à la graine et à différents niveaux d'organisation allant des molécules et des cellules à la plante entière et au couvert végétal, l'ensemble des mécanismes et processus impliqués dans le fonctionnement, le développement et la reproduction des plantes placées dans des conditions environnementales fluctuantes ou limitantes. Selon cette acception, les mécanismes et facteurs pris en compte peuvent être écophysiologiques, physiologiques, génétiques ou biochimiques. Ce choix est notamment fondé sur deux considérations : d'une part, la modélisation s'est historiquement propagée des niveaux d'organisation les plus macroscopiques (*cf.* les modèles de culture et les modèles de processus écophysiologiques qui existent depuis une quarantaine d'années [voir l'exposé de Brisson et Launay]) vers des niveaux d'organisation plus fins (avec notamment le développement de la biologie systémique qui s'appuie sur la production à haut débit de données moléculaires au niveau d'organes ou de cellules [Lejay]) ; d'autre part, les enjeux d'intégration des connaissances acquises à des niveaux d'organisation imbriquées ont conduit au développement d'approches de biologie intégrative qui reposent sur un dialogue étroit entre expérimentation et modélisation.

Les modèles sont des représentations nécessairement simplifiées ou stylisées de la réalité. Les méthodes et formalismes mathématiques ou informatiques utilisés — équations différentielles ordinaires ou équations récurrentes en temps discret, équations aux dérivées partielles, processus stochastiques, graphes, langages formels de simulation, algorithmique, méthodes de visualisation, *etc* — varient selon les mécanismes, phénomènes et niveaux d'organisation étudiés, et selon l'avancement même des connaissances. Bien qu'ils puissent justifier des détours et développements méthodologiques longs et techniques, les modèles ne sont pas une fin en soi : leur utilité ne peut être définie que par référence à leur capacité à comprendre des phénomènes, à faire des prédictions, à générer de nouvelles hypothèses ou à susciter de nouvelles expérimentations.

Les modèles sont donc avant tout des outils qui peuvent avoir différentes finalités. Le recours à la modélisation est ainsi motivé par des objectifs et des causes variés. Au rang des objectifs scientifiques on peut citer la réduction et la maîtrise de la complexité des phénomènes dans un but de compréhension : il s'agit alors d'utiliser les modèles pour analyser et disséquer le jeu des interactions en complément de, et en appui à, l'expérimentation, puis pour générer de nouvelles hypothèses, de nouvelles expériences, et ainsi guider l'acquisition des données [J. Traas]. Un autre objectif est de synthétiser et de résumer un ensemble de connaissances afin d'en analyser la cohérence et les lacunes et d'en apprécier la capacité prédictive ; la modélisation peut ainsi servir à

[1] Correspondant de l'Académie d'Agriculture de France, directeur scientifique plante et produits du végétal à l'INRA, DS PPV, 147, rue de l'Université, 75007 Paris.
C.R.Acad. Agric. Fr., 2008, 94, n°1. Séance du 28 janvier 2009.

Copyright – Académie d'Agriculture de France – 2009.

articuler et intégrer des approches disciplinaires jusque là disjointes (génétique, physiologie, écologie, biologie moléculaire, *etc*.), à transmettre les connaissances qui sont exprimées dans un langage non ambigu et qui peuvent ainsi être organisées et « compressées ». Quand ils possèdent un pouvoir prédictif avéré, les modèles et les méthodes d'inversion de modèles servent aussi à accéder à des mesures ou à des paramètres qu'il est difficile de mesurer *in situ* [de Reffye] ; ils peuvent aussi servir à contrôler et piloter des phénomènes ou à faire des prévisions en fonction de scénarios environnementaux [de Reffye ; Brisson et Launay].

L'essor de la modélisation répond aussi à des nécessités méthodologiques liées à la complexité des phénomènes étudiés. Comment pourrait-on rendre compte des synchronisations, coordinations et couplages entre des processus qui se passent à différents endroits dans un tissu (interactions biochimiques et physiques entre cellules [Traas]), dans une plante ou dans un couvert (développement racinaire et aérien, fonctionnement carboné, azoté et hydrique, compétition entre plantes voisines, *etc.* [de Reffye ; Brisson et Launay]) sans modèles ? Comment pourrait-on analyser le jeu des interactions entre des centaines, voire des milliers, de gènes, et tirer parti des masses de données générées par la biologie à haut débit sans formalismes et sans algorithmes ? Qu'il s'agisse d'imagerie cellulaire (et de la capacité de localiser et suivre des mouvements infracellulaires) ou de phénotypage moléculaire ou macroscopique (par exemple par analyse d'images), les perspectives ouvertes par les nouvelles techniques expérimentales renforcent ces besoins de modélisation.

Enfin, l'essor de la modélisation est rendu possible par l'évolution des technologies informatiques — non seulement en termes de capacité de stockage des données ou de puissance de calcul mais aussi dans le domaine de la visualisation — comme par le développement de formalismes adéquats pour représenter les mécanismes du vivant, par exemple pour simuler les processus morphogénétiques [J. Traas ; P. de Reffye] ou pour représenter des réseaux de gènes en interaction [L. Lejay]. La modélisation ne se résume cependant pas à l'importation dans le champ de la biologie végétale de méthodes mathématiques et informatiques : d'une part, la biologie des plantes pose des questions méthodologiques originales, par exemple, la représentation des processus morphogénétiques ; d'autre part, comme dans les sciences de l'environnement, dans la physique ou dans les sciences de l'ingénieur, les questions relatives aux données, à leur qualité, à leur reproductibilité et à leur accessibilité, sont centrales.

Les approches de modélisation se développent donc en biologie végétale, comme dans beaucoup d'autres domaines. Historiquement, ces approches sont d'abord apparues dans d'autres champs de la biologie : en génétique quantitative et des populations, en agronomie avec la théorie des plans d'expérience, en écologie et dynamique des populations. Elles se sont ensuite répandues à partir des années 1970 en bioclimatologie et en écophysiologie végétale [Brisson et Launay] — pour comprendre et prédire la croissance des plantes, la production de biomasse et le fonctionnement (hydrique, carboné, azoté, *etc.*) des couverts végétaux —, puis dans l'étude de la morphogénèse [de Reffye] — pour comprendre les relations entre la structure topologique et géométrique, le fonctionnement et le développement des plantes. Ce n'est que plus récemment qu'elles ont commencé d'être mises en œuvre à des niveaux d'organisation plus fins [Lejay ; Traas] dans des approches dites, selon les circonstances, de biologie intégrative, de biologie des systèmes ou de biologie computationnelle.

Il est ainsi frappant d'observer ce qui s'est passé en Grande-Bretagne : pour développer la biologie des systèmes, la communauté scientifique qui s'intéresse à la plante modèle *Arabidopsis thaliana* a initié en 2006 un dialogue approfondi avec les écophysiologistes et les agronomes qui avaient plus de compétences en modélisation (http://garnet.arabidopsis.info/Modellers_Meeting_Report_final_v2.pdf). En France, l'INRIA

Copyright – Académie d'Agriculture de France – 2009.

(Institut national de la recherche en informatique et automatique) affiche dans son plan stratégique 2008-2012 : « Un autre objectif est celui du modèle agronomique et biologique de la plante numérique qui conjugue à l'agronomie et à la biologie, les mathématiques appliquées, l'automatique, l'informatique graphique, la géométrie et la combinatoire » (http://www.inria.fr/inria/strategie/priorites_6.fr.html). Le développement de la biologie intégrative végétale, mais aussi animale et microbienne, est aussi une priorité de l'INRA (http://www.inra.fr/agroBI). Dans le même ordre d'idées l'ANR soutient depuis 2006 des programmes de biologie des systèmes et de modélisation des systèmes complexes qui sont ouverts à des projets de biologie végétale (http://www.agence-nationale-recherche.fr/AAP-185-Syscomm.html) et le réseau thématique de recherche avancée Agropolis Fondation soutient des projets relevant du thème « Plantes et écosystèmes numériques » (http://www.agropolis-fondation.fr/).

Face à l'ampleur et l'expansion du sujet, le choix a donc été fait d'illustrer, sans évidemment l'épuiser, la diversité des facettes de la modélisation en biologie végétale par quatre exposés qui traitent de questions scientifiques et de niveaux d'organisation différents : les deux premiers relèvent de la biologie végétale au sens strict, ils portent sur les réseaux de gènes impliqués dans le métabolisme azoté et carboné de la plante [Lejay] et sur les mécanismes biophysiques et biochimiques à l'œuvre au niveau d'un tissu de cellules dans la morphogénèse du méristème [Traas] ; les deux derniers relèvent de l'écophysiologie et confinent à l'agronomie, ils abordent le lien entre morphogénèse de la plante entière, architecture et production de biomasse [de Reffye] et l'intégration de connaissances écophysiologiques, bioclimatologiques et biotechniques dans des modèles de culture [Brisson et Launay].

Copyright – Académie d'Agriculture de France – 2009.