

LA BIOLOGIE INTEGRATIVE VEGETALE

(Version 7: Rapport remis à la Direction générale INRA le 21 février 2005)

Groupe de réflexion du Conseil Scientifique INRA :

Noël BOEMARE (INRA)
David BOUCHEZ (INRA)
André CHARRIER (ENSA.M.), coordonnateur
Jean-Christophe GLASZMANN (CIRAD)
Jacques JOYARD (CNRS)
Gilles LEMAIRE (INRA)
Jean-François MOROT-GAUDRY (INRA)
André POUZET (CETIOM)
Bernard SAUGIER (Université Paris XI)

Secrétariat : Pascaline GARNOT (INRA)

SOMMAIRE

	Pages
Introduction	1
1 La Biologie intégrative	3
1.1. Définitions	3
1.2. Biologie intégrative ou Biologie Systémique	6
1.3. Objectifs	8
2 La biologie intégrative végétale en France	9
2.1. Département BV	9
2.2. Département CEPIA	9
2.3. Département EA	10
2.4. Département EFPA	10
2.5. Département GAP	11
2.6. Département MIA	12
2.7. Département SPE	12
2.8. Dans l'Enseignement supérieur et la Recherche	14
2.9. La biologie intégrative au niveau international	15
3 Les approches et les verrous méthodologiques	17
3.1. Les espèces-modèles en biologie végétale	17
3.2. Génomique végétale : l'existant et l'émergent	18
3.3. Génétique et phénotypage	25
3.4. Bioinformatique et génomique comparative	27
3.5. Ecophysiologie, modélisation et biologie intégrative	29
4 Organisation, partenariat et formation	33
4.1. Organisation fonctionnelle	33
4.2. Structuration des groupes de recherche	34
4.3. Compétences et formation	36
Perspectives	39
Bibliographie	42
Annexes (30 pages)	

INTRODUCTION

Au cours des années 90, une approche nouvelle de la Biologie, qualifiée de « Biologie intégrative » s'est développée. Elle est actuellement considérée par certains comme une véritable révolution pour les Sciences de la Vie. En grande partie, elle a pour origine le développement spectaculaire, des nouvelles technologies de la génomique et du séquençage des génomes aux techniques d'analyse fonctionnelle des réseaux de gènes et de protéines. Les masses de données disponibles, sans commune mesure avec celles manipulées auparavant par les biologistes, nécessitent un véritable changement de paradigme pour, au delà de la description mécaniste des réseaux moléculaires en jeu, parvenir à une interprétation pertinente intégrant les différents niveaux d'organisation biologique. Ce changement apparaît d'autant plus nécessaire du fait de la prévalence des approches analytiques et réductionnistes au cours des décennies précédentes, qui atteint actuellement son stade ultime -et ses limites probables- avec le séquençage de génomes complets ou les approches "métabolomiques".

A la demande de la Direction Générale de l'INRA, le Conseil Scientifique du 3 février 2004 a constitué un groupe de réflexion chargé d'élaborer un rapport sur la biologie intégrative végétale, considérée, au delà de la plante, jusqu'au peuplement végétal dans son environnement. Ce groupe a été composé de 5 membres du conseil scientifique et de 4 personnalités scientifiques extérieures à ce conseil. Il a reçu pour mandat :

1. d'établir un état de l'art sur la Biologie intégrative ;
2. d'analyser la place de l'INRA et son organisation dans ce domaine ;
3. de clarifier les enjeux scientifiques et les questions de recherche prioritaires ;
4. d'identifier les compétences et les formations à développer.

Pour remplir cette mission, le groupe a réalisé une recherche bibliographique sur le thème et consulté les documents des évaluations et des schémas stratégiques des 7 départements INRA concernés. Il a aussi organisé des consultations de responsables de départements et d'unités, ainsi que 4 rencontres sur site (Bordeaux, Montpellier, Région parisienne et Toulouse) des scientifiques INRA et leurs partenaires dans les unités mixtes de recherche, sur la base d'un questionnaire en 12 points (*Annexe I6*). Ces différentes réunions ont été l'occasion d'échanges d'expériences, de débats ouverts et ont initié la nécessaire mutualisation de la réflexion.

Le groupe de réflexion a présenté son dossier au cours de son élaboration devant le Conseil Scientifique INRA (24 juin et 25 novembre 2004). Le rapport de synthèse comprend 4 parties :

• La biologie intégrative : définitions, concepts et objectifs

La consultation des scientifiques et les informations bibliographiques conduisent à une diversité de définitions et de conceptions de la Biologie intégrative. En première analyse, on peut distinguer deux démarches convergentes : l'une mobilise les approches réductionnistes de la génomique, et assigne à l'intégration l'objectif de rendre intelligibles les grandes fonctions biologiques au niveau de l'organisme ; l'autre part d'une vision globale des systèmes biologiques (agro-écosystèmes, pathosystèmes...) et analyse la variabilité de leurs composantes en vue d'identifier les fonctions physiologiques sous-jacentes. Le groupe de réflexion a porté son attention sur ces différentes démarches et sur leur contribution à la compréhension de systèmes biologiques complexes.

• La biologie intégrative végétale en France

L'affichage de la Biologie intégrative dans les Universités et les Organismes de recherche en France a été analysé par une recherche documentaire. A l'INRA, la volonté de développer les approches intégratives est mentionnée dans les schémas stratégiques de la plupart des

départements de recherche ; quelques exemples choisis dans le domaine végétal permettent d'illustrer ces projets, en particulier à travers des projets de recherche en partenariat impliquant souvent le CNRS, le CIRAD ou l'IRD... Le contexte international de la recherche et de la formation en biologie intégrative est aussi brièvement abordé.

• **Les approches et les verrous méthodologiques**

Ce chapitre n'a pas pour objet de présenter en détail les différentes facettes de la génomique en biologie végétale. Vu les progrès exceptionnels réalisés en 10 ans en génomique structurale et fonctionnelle, en biotechnologie, en analyse de la diversité génétique, en modélisation de fonctionnement de la plante et du peuplement végétal dans leur environnement, le groupe de réflexion s'est interrogé sur la pertinence des approches, sur le choix des modèles biologiques... Il a surtout cherché à identifier les verrous méthodologiques, les approches émergentes en génomique et en génétique, ainsi que certains nouveaux défis pour l'écophysiologie.

• **Organisation, Partenariats et Formation**

La consultation des scientifiques et l'analyse de l'organisation fonctionnelle des projets de recherche conduisent le groupe de réflexion à poser quelques principes directeurs pour favoriser la prise en compte de la complexité dans l'étude des systèmes biologiques. Il propose également différents moyens incitatifs pour motiver les équipes s'engageant dans la réalisation de projets de Biologie intégrative. L'acquisition des compétences, la valorisation des savoir-faire et la sensibilisation des scientifiques aux approches émergentes à l'interface biologie/mathématiques/physique en sont des aspects essentiels.

En conclusion, ce rapport de synthèse reprend quelques propositions formulées et tente quelques réflexions spéculatives sur la Biologie intégrative du futur.

1 La Biologie Intégrative : définitions, concepts et objectifs

L'information disponible dans la littérature scientifique ou sur le web à propos de la Biologie Intégrative ou Systémique¹ est particulièrement abondante, en particulier dans le domaine biomédical ou en microbiologie. On en retire le sentiment d'un champ d'investigation prometteur et dynamique, où comme souvent pour les disciplines émergentes, les promesses et espoirs sont plus faciles à mettre en avant que les réalisations concrètes. Les scientifiques consultés nous ont aussi donné des définitions et des visions variées de la Biologie intégrative végétale.

1.1 Définitions

Le terme de Biologie Intégrative peut être considéré comme un pléonasme. En effet, la Biologie, science des êtres vivants, a par essence un objectif d'intégration, même si cette vision intégrée du fonctionnement des organismes et des populations vient le plus souvent après une phase de déconstruction analytique. Depuis Claude Bernard, l'approche réductionniste vise à expliquer les organismes par le fonctionnement coordonné "d'organes" remplissant des "fonctions" ; lesquels organes et fonctions peuvent à leur tour être décomposés en composants de plus en plus élémentaires. Si l'on éprouve aujourd'hui le besoin de lui adjoindre le terme « intégrative » c'est sans doute essentiellement pour des raisons historiques, liées aux évolutions de la discipline au cours des dernières décennies. En effet, les études analytiques, portant sur des parties et des mécanismes de plus en plus fins, ont largement dominé cette discipline pendant de nombreuses années et ont abouti à une fragmentation des savoirs en Biologie. D'où un besoin de plus en plus clairement exprimé par les biologistes, de reconsidérer l'organisme vivant dans son intégrité et son fonctionnement global, en interaction avec son environnement.

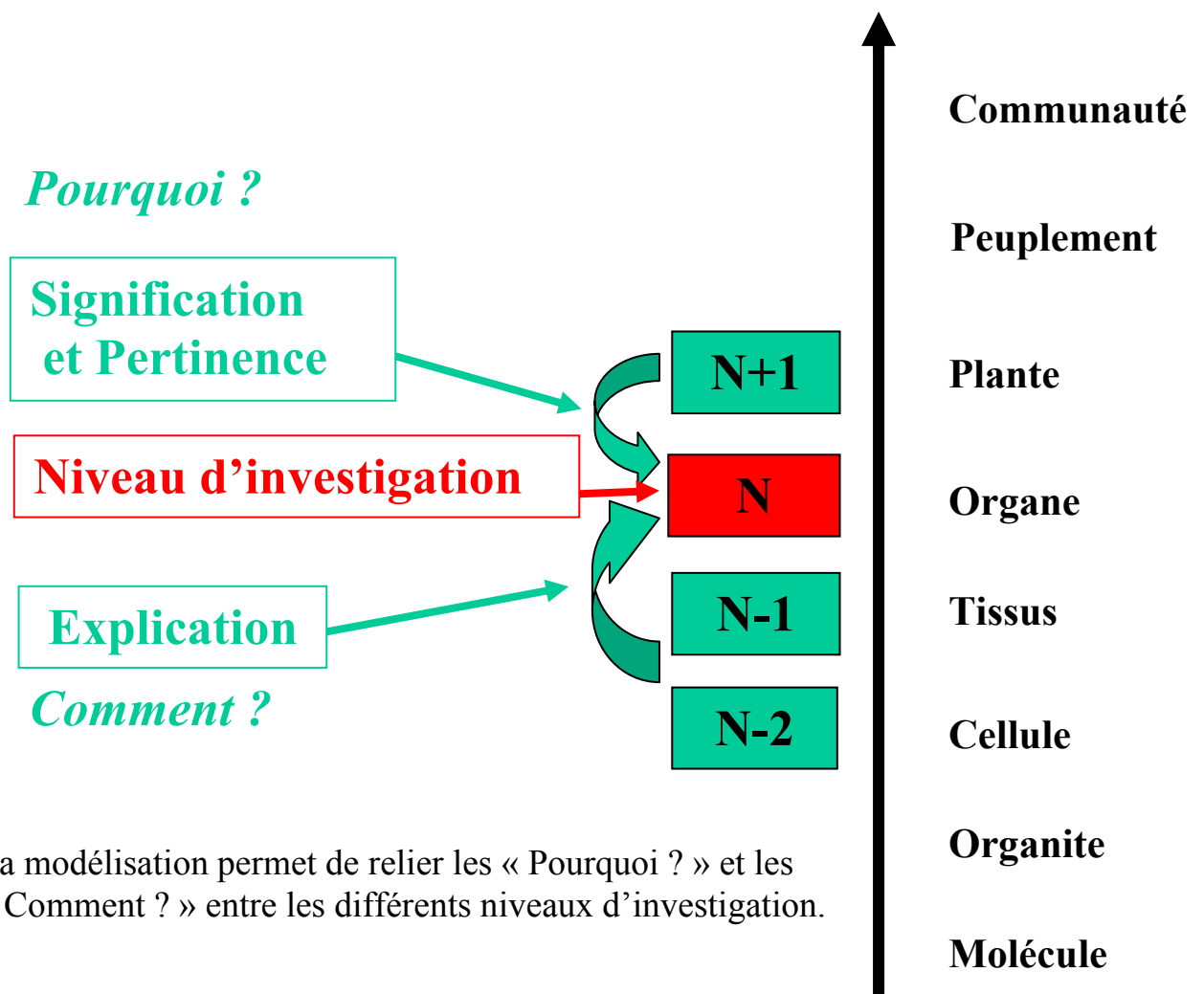
Si la définition du terme « Biologie » ne pose pas de problème, il n'en va pas de même du mot "Intégrative" qui donne lieu à une grande diversité de définitions parmi les biologistes, souvent en fonction de leur discipline et de leur expérience personnelle de recherche. L'intégration est souvent comprise dans une première acception, selon laquelle il s'agit surtout d'intégrer une masse importante de données ou de connaissances résultant des niveaux d'organisation les plus élémentaires, c'est à dire moléculaires (génomique, transcriptomique, protéomique, métabolomique) qui sont progressivement mis à la disposition des chercheurs grâce aux méthodes d'investigation massive. C'est l'option qui prévaut le plus souvent parmi les spécialistes de génomique, de génétique et physiologie moléculaire. Intégrer signifie alors « *donner du sens à l'ensemble d'informations disponibles* ». Se pose alors immédiatement la question du *sens* en biologie, dont la réponse dépendra bien entendu des options et sensibilités "disciplinaires". D'autres questions concernent la quantité et la qualité des données requises pour la mise en place de modèles explicatifs cohérents : l'exhaustivité des données élémentaires est elle toujours nécessaire ? Il s'agit ici d'une intégration de nature "horizontale", concernant des données homogènes à un niveau d'organisation donné, mais dont la masse implique le développement d'outils mathématiques et informatiques spécifiques permettant de traiter, classer et mettre en relation les données afin d'en permettre une interprétation biologique pertinente.

Le terme "intégratif" peut aussi qualifier le *passage d'un niveau d'organisation élémentaire à un niveau d'organisation plus complexe (donc plus « intégré »)*. Il s'agit alors d'une

¹ Voir partie 1.2 : "Biologie Intégrative ou Biologie Systémique ?"

intégration de type "vertical". Dans ce sens, l'intégration est une notion qui découle de la théorie des systèmes complexes. Un système complexe est composé de groupes de composants élémentaires, au sein desquels les interactions sont plus importantes et surtout plus rapides que les interactions entre groupes (Auger et Lett, 2003). Cette grille de lecture permet d'identifier, au sein du système, des niveaux d'organisation associés à des échelles spatiales et temporelles différentes. Cette approche des systèmes complexes semble pouvoir s'appliquer à l'ensemble des systèmes biologiques et écologiques, dès lors que l'on peut identifier clairement des niveaux d'organisation qui sont pertinents pour le fonctionnement du système : molécule, organite cellulaire, cellule, tissu, organe, organisme, population, communauté...

Pour l'analyse d'un tel système hiérarchisé, l'intégration des informations à un niveau d'organisation N dans un modèle explicatif cohérent nécessite une "traduction" dont la pertinence ne peut être jugée qu'au niveau d'intégration supérieur N+1, cette opération pouvant être réitérée pour les niveaux d'ordre supérieur N+2...N+i... Comme l'a indiqué Passioura (1979), un système hiérarchisé possède un langage et des concepts qui sont propres à chacun des niveaux d'organisation, et notre compréhension d'un tel système dépend en fait de notre capacité à mettre en correspondance (traduire) les langages des niveaux d'organisation adjacents. Traduire vers un niveau d'organisation plus élémentaire (N-1) permet de fournir une explication (approche réductionniste), tandis que traduire vers un niveau d'organisation supérieur (N+1) permet de donner une signification aux phénomènes étudiés à un niveau d'organisation donné N (approche holiste). Ainsi notre compréhension d'un phénomène biologique étudié à un niveau d'organisation donné reste incomplète tant qu'on ne peut pas le relier aux processus actifs aux niveaux d'organisation adjacents. Si nous ne pouvons traduire ce phénomène à un niveau d'organisation plus élevé, alors celui-ci risque d'être trivial et déconnecté de son contexte, c'est à dire sans signification fonctionnelle, et si nous ne pouvons le traduire à un niveau d'organisation plus élémentaire alors l'étude sera purement descriptive. Ainsi selon Passioura (1979) la recherche en biologie devrait toujours (i) identifier le(s) niveau(x) d'investigation du processus étudié, le(s) plus pertinent(s) par rapport à la question posée, (ii) chercher les éléments d'explication aux niveaux d'organisation N-1, et (iii) découvrir sa signification (donc sa fonction) aux niveaux N+1...



La modélisation permet de relier les « Pourquoi ? » et les « Comment ? » entre les différents niveaux d'investigation.

Figure 2 : Interrelations entre niveaux d'organisation d'après Passioura (1979).

Un troisième sens peut être donné au mot "intégrative". Il s'agit alors d'une intégration qu'on peut qualifier de "transversale", qui permet la transposition de connaissances acquises sur une espèce particulière à un ensemble plus large d'espèces. Cette dimension prend une importance particulière dans une biologie souvent dominée par les travaux sur organismes modèles. Ceci correspond au champ de la biologie évolutive et comparative qui, en replaçant les connaissances acquises dans un cadre évolutif prenant en compte la diversité des individus et des espèces, permet d'identifier pour chacun des niveaux d'organisation étudiés la valeur générique ou au contraire la spécificité des résultats produits. Au delà de ces aspects de généralisation (qui restent une difficulté majeure en biologie), ce type d'approche peut apporter beaucoup dans la compréhension fine des processus biologiques : voir le succès des approches "évolution-développement" à la croisée de la biologie évolutive et de la biologie du développement. La biologie comparative ne doit pas rester confinée au seul niveau d'organisation moléculaire (arbres de séquences...), mais doit impliquer l'ensemble des niveaux pertinents pour l'évolution des génomes et des organismes. Toutefois, la possibilité de mener des comparaisons évolutives sur des génomes complets comme c'est le cas sur certains microorganismes paraît riche d'enseignements, et il est souhaitable que dans ce domaine, la Biologie Végétale puisse s'inspirer des avancées du domaine. La question du choix judicieux des modèles d'étude et de leur valeur générique pour l'ensemble du monde végétal est alors posée. Au delà de la transposition des connaissances obtenues sur les plantes

modèles vers les plantes d'intérêt agronomique, il s'agit bien ici de développer les approches permettant de révéler la structuration de la régulation des différentes fonctions du vivant au sein du règne végétal et aussi entre les différents règnes animal et végétal. Il importe alors de se doter des outils, des concepts et des méthodes adéquats pour développer cette approche comparative et évolutive qui doit permettre l'approche de l'expression de la diversité des phénotypes et des interactions génotype-environnement.

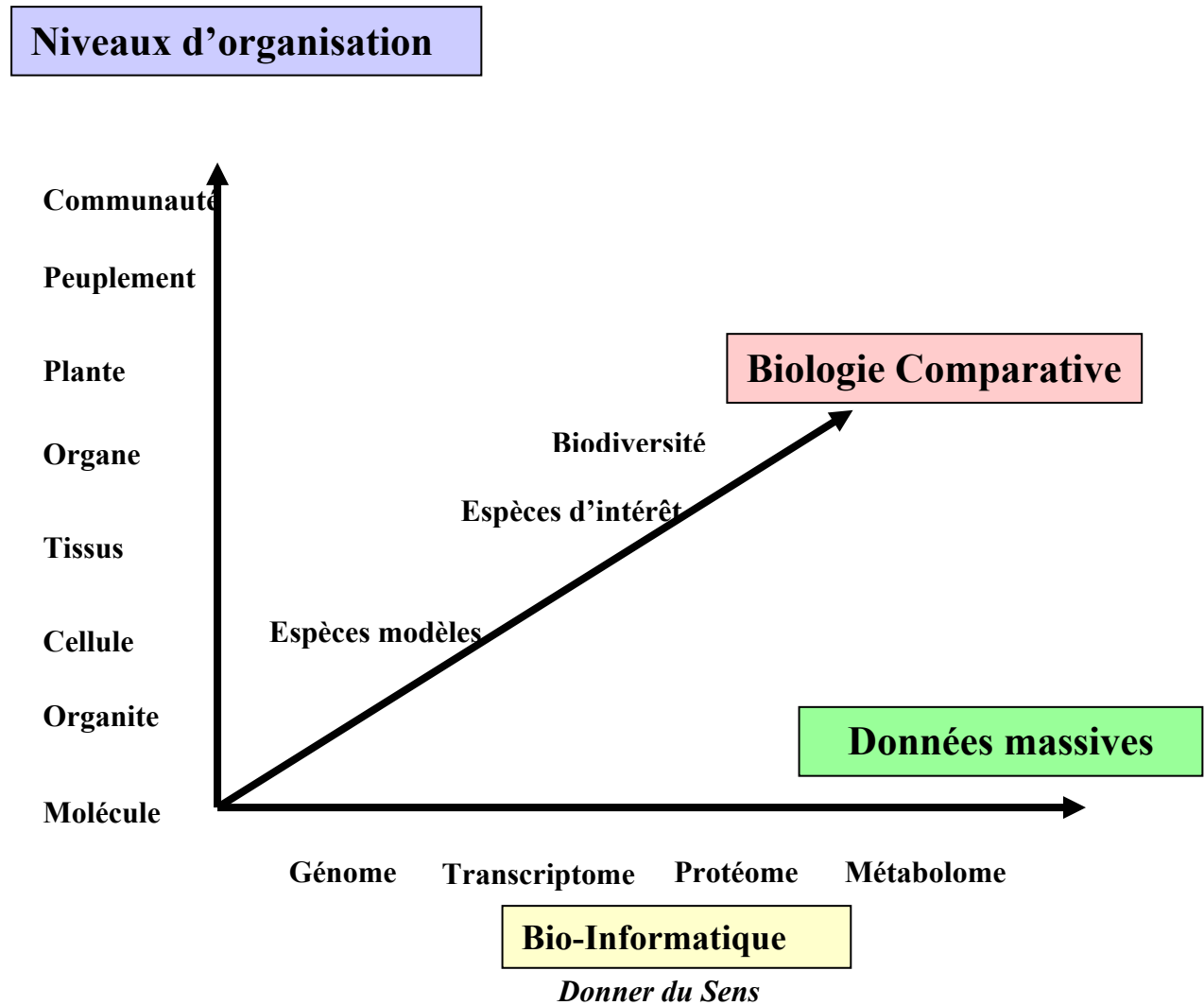


Figure 1 : Les trois dimensions de l'intégration en Biologie

1.2 Biologie Intégrative ou Biologie Systémique ?

Les différents sens associés au terme « intégration » peut faire préférer le terme de "Biologie des Systèmes" (System Biology) beaucoup plus répandu dans le monde anglophone. La Biologie des Systèmes est « *l'étude itérative et intégrative des systèmes biologiques en tant que systèmes, en réponse à des perturbations* » (Auffray et al. 2003). Ces auteurs suggèrent que les systèmes biologiques sont des systèmes auto-organisés autour d'entités de natures différentes : (i) les composantes moléculaires (ADN, ARN, protéines...) directement reliées à une information génétique de nature digitale, et (ii) les réactions biochimiques et métaboliques qui sont soumises aux informations environnementales qui sont elles de nature

analogique. Cette distinction entre deux sources d'information, l'une structurelle d'origine génétique et l'autre contextuelle d'origine environnementale, reprend sous une forme actualisée le dualisme génome-environnement pour la construction du "phénotype" des organismes vivants. En fait le fonctionnement des systèmes biologiques est déterminé par les nombreuses interactions entre ces deux types d'information. F. Capra (1996) dans son livre *The Web of Life*, montre que pour comprendre le fonctionnement d'un système biologique, l'étude "hors-contexte" des objets élémentaires qui composent le tout ne suffit pas, et qu'il s'agit d'identifier les interrelations qui lient ces objets entre eux. Plus encore, il ne s'agit pas seulement d'identifier les relations internes au système étudié, mais aussi les relations qui lient ce système dans son entier aux systèmes englobants dans lesquels il est situé (cellule dans son tissu, organe dans la plante entière, plante dans son environnement naturel...), ce qui constitue le contexte. Ceci implique un changement de vision : il faut passer de la mesure quantitative des éléments du système à l'étude qualitative des relations entre ces éléments et de leurs configurations au sein d'un patron d'organisation.

Les systèmes biologiques sont des systèmes « ouverts » dans ce sens qu'ils échangent de façon continue de la matière et de l'énergie avec le milieu extérieur (von Bertalanffy, 1940). De ce fait, leur fonctionnement peut être approché par l'application des lois de la thermodynamique. A titre d'exemple, c'est ce qui a été fait dans les années 1960-1990 par l'Ecophysiologie qui a assimilé les peuplements végétaux à des surfaces d'échanges de masse et d'énergie. Ceci a permis de modéliser la production primaire des peuplements végétaux sans avoir à expliciter l'ensemble des processus biologiques sous-jacents et leurs régulations génétiques. Cette approche féconde a atteint ses limites car les processus qui contrôlent la répartition des assimilats et donc de la biomasse entre les différentes parties d'une plante ne pouvaient être analysés au niveau d'organisation du peuplement. Ces processus qui découlent de la morphogenèse de la plante ne peuvent être appréhendés que par une approche architecturale de la croissance d'une plante individuelle. Celle-ci doit être replacée dans son environnement, c'est-à-dire en compétition avec les plantes voisines pour les ressources du milieu au sein d'un peuplement. Ainsi d'une approche du peuplement comme une machine thermodynamique échangeant de la masse et de l'énergie avec son environnement, l'Ecophysiologie est passée à une approche du peuplement constitué d'une collection de plantes individuelles qui interagissent entre elles. De ces interactions il est alors possible d'explicitier des propriétés émergentes qui traduisent le fonctionnement collectif de la population.

La Cybernétique a eu un rôle important dans l'émergence de la théorie des systèmes complexes en permettant justement d'analyser et de caractériser les patrons d'organisation des systèmes : boucles et réseaux d'information. Les boucles de rétro-action (feed-back), très abondantes dans les systèmes biologiques permettent à ces systèmes d'avoir des propriétés d'auto-organisation et d'apprentissage qui leur sont caractéristiques. Ces capacités d'auto-organisation permettent à ces systèmes complexes d'avoir des propriétés émergentes à chacun des niveaux d'organisation qui les structurent. C'est l'existence de ces propriétés émergentes qui permettent de dire que le tout n'est pas réductible à la somme des parties. Ainsi la simple addition des éléments constitutifs d'un système en dehors d'un assemblage déterminé par un patron d'organisation des relations entre ces éléments ne permet pas de découvrir les propriétés de l'ensemble du système. Cette vision systémique s'applique quelle que soit la taille et le niveau d'organisation du système (de la cellule à l'organisme jusqu'à la communauté) dès lors qu'un minimum de complexité permet de voir émerger des patrons d'organisation. Ainsi les gènes et leurs produits dérivés (ARN et protéines) ne sont que les composants élémentaires des systèmes biologiques complexes qui sont gouvernés par ailleurs par des propriétés physiques et chimiques non réductibles à ces composantes élémentaires.

1.3 Objectifs

Les trois visions de l'intégration, l'une quantitative et horizontale, l'autre hiérarchisée et verticale, et la troisième transversale peuvent se combiner dans le concept de Biologie des Systèmes Intégrés. Plus on descend dans les niveaux d'organisation les plus élémentaires, plus la quantité d'informations élémentaires disponibles est massive, information qu'il convient d'organiser et qui ne peut être interprétée qu'à un niveau organisationnel plus large. Ainsi lorsque l'on remonte à des niveaux d'organisation hiérarchiquement plus élevés et complexes, seule une partie de l'information sous-jacente reste pertinente. L'intégration consiste alors à déterminer cette part de l'information qui reste pertinente et les interactions qui lui donnent un sens au niveau d'organisation plus élevé. C'est en ce sens que l'on peut parler de propriétés émergentes d'un système. Ces propriétés identifiables à un niveau d'organisation donné ne sont pas décelables, ni déductibles par la somme des informations des niveaux sous-jacents. La détermination des propriétés émergentes d'un système à chacun de ses niveaux d'organisation nécessite le recours à la modélisation qui permet de faire une véritable traduction d'information entre niveaux.

Ainsi on peut émettre le postulat que l'échelle de l'organisme dans son environnement détermine autant l'expression des gènes que l'échelle des gènes et de leur expression détermine le fonctionnement de l'organisme et ses échanges avec le milieu. Les deux visions sont chacune une facette du système biologique complexe que représente un organisme vivant. Elles sont toutes les deux légitimes, mais chacune d'elles reste partielle. L'objectif de la Biologie Intégrative est justement de développer ces deux points de vue de manière convergente, afin que les informations relatives à chacune des approches trouvent son sens vis à vis de l'autre approche. Un organisme vivant ne peut être réduit ni à un ensemble aussi complet soit-il de séquences géniques, ni à un ensemble quantifié de relations avec son environnement ; c'est bien l'interaction entre ces deux ensembles qui constitue l'objet d'étude de la Biologie intégrative.

Pour conclure cette 1^{ère} partie sur la définition et les concepts de la Biologie intégrative, nous retiendrons 4 objectifs :

- Donner du sens aux approches analytiques, en particulier aux données issues de la génomique ;
- Intégrer les informations acquises à différents niveaux d'approche (dialogue entre les niveaux et les disciplines), c'est-à-dire donner explication et signification aux processus étudiés ;
- Construire un cadre théorique et linguistique commun à la Biologie, à la Chimie, aux Mathématiques et à la Physique dans la perspective d'une meilleure compréhension du Vivant ;
- Changer l'état d'esprit et la façon de concevoir les programmes de recherche en Biologie.

2 La Biologie Intégrative Végétale en France

L'affichage de la Biologie intégrative végétale dans les universités et les organismes français de recherche a été analysé par une recherche documentaire extensive et illustré par quelques exemples choisis dans le domaine végétal, principalement dans les départements et les unités mixtes de recherche INRA.

2.1 Biologie Végétale (BV)

Le département BV, par nature pluridisciplinaire, est un des lieux privilégiés de la construction d'une biologie intégrative végétale. Il développe en particulier une physiologie moléculaire dont le but est de comprendre l'intégration des processus physiologiques et développementaux dans la physiologie de la cellule et de la plante.

L'émergence de nouveaux outils pour identifier, cloner les gènes et modifier les génomes rend possible l'exploration des relations fonctionnelles entre les aspects moléculaires et l'expression globale (phénotype) au niveau de la plante entière. Elles sont principalement développées sur des plantes modèles pour lesquelles le séquençage du génome est acquis ou en cours, ainsi que la modélisation de la croissance et de l'architecture.

Le génie génétique offre la possibilité théorique d'introduire un caractère plus rapidement que par la sélection classique. En identifiant les couples gènes/fonction, les approches de physiologie moléculaire permettent de proposer les transformations génétiques capables de répondre à un objectif agronomique donné. Toutefois, les résultats encore limités en terme d'application montrent la nécessité d'acquérir des connaissances sur l'intégration physiologique des étapes élémentaires codées par les gènes que l'on veut modifier. En particulier, il est impératif de définir dans une famille de gènes contribuant à un processus physiologique, ceux dont l'expression est limitante dans le processus. La prise en compte de la variabilité génétique est un puissant outil exploitable dans cette perspective.

En conséquence, le département focalise ses efforts sur la physiologie moléculaire de la nutrition carbonée, azotée et minérale de plantes modèles, avec des études sur les réponses intégrées de la plante à la qualité de l'environnement et aux contraintes environnementales.

Le transfert des connaissances et outils vers des espèces végétales d'intérêt agronomique nécessite souvent le développement d'études spécifiques en collaboration avec d'autres départements².

[Voir Annexe 1: Nutrition azotée – F.Vedele et B. Ney]

2.2 Caractérisation et Elaboration des Produits Issus de l'Agriculture (CEPIA)

La contribution du département CEPIA s'inscrit en partie dans la biologie intégrative, avec une limitation importante des niveaux biologiques considérés aux seuls organes ou tissus d'intérêt pour l'élaboration des produits agro-industriels.

Une condition de succès de ce champ est de déterminer les relations entre les événements biologiques au niveau de l'expression et de la diversité des génomes d'une part, et les

² La création à Montpellier d'un *Institut de Biologie Intégrative* rassemblant deux unités, l'une rattachée au département BV et l'autre au département EA témoigne de la prise en compte de l'importance d'une démarche commune à plusieurs départements de l'INRA analysant le vivant à divers niveaux d'intégration.

événements biochimiques appréhendés en génie de la réaction d'autre part. Cette démarche d'investigation est réalisée dans un continuum avec les généticiens et les physiologistes qui, par leurs nouvelles connaissances en génomique, sont en mesure de produire une diversité raisonnée et systématique de matières premières. L'interdisciplinarité entre physico-chimistes - biochimistes de CEPIA et biologistes des départements d'amont est l'une des conditions du succès.

[Voir Annexe 2 Qualité de la tomate M. Causse]

2.3 Environnement et Agronomie (EA)

Dans le Département EA l'intégration des connaissances obtenues à différentes échelles (peuplement, agroécosystèmes...) est un objectif important, en particulier pour appliquer les sorties de l'écophysiologie à l'Agronomie et à l'Ecologie. La participation du Département à la démarche de Biologie Végétale Intégrative est principalement localisée dans le champ thématique « Modélisation du phénotype et des flux plante-environnement, interaction génotype – environnement ».

La volonté clairement exprimée du Département EA est de s'impliquer dans la mise au point et le partage des outils de la Biologie Intégrative, avec ses compétences, ses méthodes et ses concepts spécifiques. Le rôle du Département ne peut donc être simplement celui d'un prestataire de services pour caractériser des mutants. L'application des modèles validés en écophysiologie et en agronomie pour la caractérisation des phénotypes (de plantes modèles et de plantes cultivées) est certainement l'un des moyens à privilégier pour favoriser le développement de la biologie végétale intégrative. Un exemple de démarche de ce type est fourni par les études communes entre EA-Grignon et BV-Versailles sur le métabolisme azoté.

[Voir Annexe 1 : F. Vedele et B. Ney]

2.4 Ecologie des Forêts, Prairies et Milieux Aquatiques (EFPA)

Dans le cadre de l'ex-département Forêts et Milieux Naturels (FMN), trois champs thématiques -Ecophysiologie et Cycles biogéochimiques ; Interactions forêts et agents biologiques ; Génétique et Physiologie des arbres forestiers- ont participé directement à la problématique Biologie Intégrative. Ces différentes approches sont aujourd'hui réparties et bien identifiées au sein des différents champs thématiques du nouveau département EFPA.

Centrés sur un objet bien identifié -l'arbre et le peuplement forestier-, ces champs thématiques abordent l'intégration de fonctions physiologiques élémentaires à des niveaux d'organisation assez larges (plante entière, station, parcelle) et sur des échelles temporelles très longues. Le terme de « biologie intégrative des arbres » est même explicitement utilisé pour identifier un besoin de renforcement de compétences « *afin de favoriser l'émergence des recherches à la confluence de la génomique, de la physiologie et de l'écophysiologie, avec une priorité accordée (1) à la tolérance au stress hydrique dans une perspective de changement climatique, (2) à la formation du bois en considérant que les propriétés technologiques de ce matériau naturellement composite dépendent de la manière dont il s'est constitué* »

Il est aussi associé à l'approche de la diversité des espèces forestières et de son évaluation (Annexe 4). Le projet de l'UMR BIOGECO (INRA – Université Bordeaux 1) repose sur une analyse plus intégrée de la diversité biologique, considérant les interactions comme les moteurs de son évolution. L'approche intégrative est focalisée sur quelques modèles biologiques (forêts de chêne, pin maritime...) étudiés dans des sites expérimentaux choisis pour représenter des écosystèmes types. Les niveaux d'intégration concernent les échelles

biologiques d'organisation, les échelles de distribution spatiale de la diversité et les étapes de l'évolution de la diversité depuis l'holocène. Les données expérimentales sont confrontées aux approches théoriques associant modèles analytiques et simulation. L'objectif est de dégager des conclusions de portée générale facilitant leur transfert aux gestionnaires des milieux et forêts.

[Annexe 3 - *Efficienc e d'utilisation de l'eau (WUE)*]

[Annexe 4 – *Evolution de la diversité des espèces forestières*]

2.5 Génétique et Amélioration des Plantes (GAP)

Ce Département a joué un rôle pionnier en France dans le développement de méthodes à haut débit pour l'analyse des génomes et de leur expression pour la Biologie végétale. En effet, les recherches en biologie végétale se sont très largement appuyées sur le développement de ressources biologiques (collections de mutants d'insertion, variabilité naturelle...), génomiques (puces CATMA, ORFéome..) et informatiques (par exemple à travers Génoplante) et ont constitué un investissement majeur du Département GAP de l'INRA au cours des dernières années. Très largement développées sur *Arabidopsis thaliana*, ces approches sont aussi mises en œuvre actuellement sur *Medicago truncatula* pour contribuer aux avancées méthodologiques sur ces deux espèces et exploiter les connaissances acquises au profit des espèces cultivées. Cette intégration inter-espèces (transversale) des connaissances est particulièrement importante pour le développement des recherches finalisées de l'INRA.

Ayant ainsi développé, souvent pour l'ensemble de la communauté végétale française, des méthodes d'investigation massives ou à haut débit, les chercheurs du Département GAP tentent actuellement d'intégrer les données ou connaissances résultant de recherches à des niveaux moléculaires d'organisation du vivant pour l'analyse de fonctions bien définies :

- la croissance, le développement, la morphogenèse et l'architecture ;
- la biologie des organes reproducteurs et la recombinaison ;
- la formation et le remplissage des du grain et le développement des fruits afin d'élaborer les composantes de la qualité ;
- la résistance aux bioagresseurs des cultures ;
- la tolérance aux contraintes du milieu physique.

D'autres recherches GAP contribuent au développement de stratégies d'utilisation et de gestion durables de la diversité des plantes cultivées et de leurs apparentées. Ces travaux bénéficient d'un matériel biologique remarquable et diversifié maintenu dans des collections de ressources génétiques et enrichi par des populations en ségrégation, des collections de mutants, d'hybrides entre espèces... Ils s'appuient sur les concepts de la génétique des populations et de la génétique quantitative, avec une intégration rapide des connaissances sur les génomes (cartes génétiques et physique ; gènes candidats...) et l'utilisation des ressources génomiques (marqueurs, EST) ainsi que la modélisation de la variation des caractères quantitatifs.

Ce Département réalise aussi de nombreux travaux qui ont pour objectif de passer d'un niveau d'organisation élémentaire à un niveau d'organisation plus complexe. Il s'agit en effet de replacer le fonctionnement des génomes dans le fonctionnement intégré de la plante, seule ou en peuplement. C'est la raison pour laquelle les collaborations développées avec les physiologistes et les écophysiologistes ont été réalisées sur divers matériels végétaux (variabilité naturelle, mutants, transgéniques) placés dans diverses conditions environnementales (stress environnementaux, interactions avec des agents pathogènes...). Les travaux sur l'interaction *Génotype x Environnement* ou ceux réalisés sur le développement du fruit dans le programme *Qualité de la tomate* témoignent de l'importance de cette stratégie

d'intégration et de renforcement des collaborations interdisciplinaires au sein du Département GAP, en partenariat avec d'autres départements de l'INRA (BV, SPE, EA...). Cette stratégie est essentielle dans la mesure où le processus d'innovation variétale est au cœur des préoccupations du Département qui doit pour cela développer des méthodes et concepts nouveaux basés sur l'intégration des connaissances du gène à l'écosystème.

Enfin, le Département GAP met en œuvre un niveau supplémentaire d'intégration au sein d'un champ thématique ayant pour objectif d'appréhender de façon globale les problèmes liés à l'introduction des innovations variétales. Il s'agit de constructions interdisciplinaires au sein de l'INRA, associant sciences sociales et sciences biotechniques. Ce niveau d'intégration sort du champ de la Biologie intégrative, mais démontre la diversité des situations abordées dans une démarche finalisée. Un exemple de projet intégré de ce type est le projet *Pois d'hiver* qui associe remarquablement l'ensemble des compétences présentes au sein de l'INRA pour introduire de nouvelles variétés sur le marché.

2.6 Mathématiques et Informatique Appliquées

Dans le domaine de la Biologie intégrative, les recherches réalisées par MIA, souvent en partenariat avec les autres départements, concernent différents thèmes comme :

- la recherche d'une structure dans les données « en masse » de la génomique (bioinformatique),
- la modélisation pour tester des hypothèses biologiques (mise en relation des données d'entrée et de sortie d'un processus) ou intégrer les processus à différentes échelles (organisme, espace, temps),
- l'exploitation d'informations dynamiques et spatialisées, etc...

(Voir le rapport Modélisation : contribution B.GOFFINET au CS INRA du 15/02/05)

2.7 Santé des plantes et environnement (SPE)

Le département SPE est aussi par nature impliqué dans une démarche de biologie intégrative, aussi bien dans la démarche « gène - organisme entier » que dans la démarche systémique inverse, puisqu'il s'agit « d'intégrer » les interactions existant entre les plantes, leurs symbiotes et leurs pathogènes. Sont envisagés également les éléments biotiques contrôlant les pathogènes et symbiotes des plantes. Ainsi la démarche « quantitative et horizontale » se reconnaît notamment dans les études consacrées à la réponse aux agresseurs ou « auxiliaires » (microorganismes, insectes, nématodes), tandis que la « démarche hiérarchisée et verticale » apparaît principalement lors des travaux sur les effets anthropiques des xénobiotiques utilisés pour combattre les bio-agresseurs, dans les travaux dévolus à l'épidémiologie des maladies végétales et dans ceux consacrés à la dissémination des invertébrés, prédateurs, vecteurs de maladies ou auxiliaires de la protection des plantes.

Le Département a contribué et contribue à l'effort de génomique sur les groupes d'organismes impliqués dans la santé des plantes, ce qui lui a permis de mobiliser ses équipes dans l'essor de la biologie à haut débit, d'intégrer et, parfois, d'initier des consortiums internationaux. Au-delà de ces projets de séquençage qui constituent un élément structurant, l'enjeu est de mettre l'accent sur les questions biologiques qui permettront d'accéder à une vision intégrative du fonctionnement des interactions plantes/insectes, plantes/micro-organismes.

S'agissant de la biologie des populations, les enjeux actuels concernent une meilleure prise en compte d'une part de la diversité des échelles spatiales, ce qui suppose un renforcement en dynamique des populations, d'autre part des processus évolutifs aux différentes échelles temporelles, ce qui implique une conversion des généticiens de la génétique neutre à l'étude des phénomènes d'adaptation.

Une large gamme de dynamiques de bioagresseurs a été étudiée et modélisée par les dynamiciens des populations et les épidémiologistes de l'INRA. La modélisation des systèmes plantes / bioagresseurs, fondée sur la biologie des bioagresseurs en interaction avec les populations d'hôtes, doit aborder différents aspects dont l'importance varie selon le système considéré. La formalisation des connaissances biologiques et l'amélioration des modèles de développement passent par la résolution de problèmes scientifiques et méthodologiques, parmi lesquels certains (classés par ordre d'importance croissante) nous apparaissent cruciaux :

- La caractérisation des processus dynamiques à différentes échelles spatiales emboîtées (de l'organe à la plante, au peuplement, à la région, au continent): relation avec l'échantillonnage, définition d'échelles spatiales pertinentes pour la quantification ou l'analyse de processus, couplage entre les composants des systèmes (plantes, bioagresseurs, ennemis naturels), définition des risques en protection des cultures.
- La caractérisation des processus dynamiques à différentes échelles temporelles: relation avec les phénomènes d'invasion et de persistance des populations au sein d'un (agro) système, les paramètres pour les décrire et prédire le développement ou non de dynamiques – étude des processus polyétiques et des tailles des populations de bioagresseurs au cours de cycles culturaux précédents.
- L'analyse démo-génétique : couplage entre dynamique et génétique des populations de bioagresseurs afin de prendre en compte les effets réciproques de la démographie et de l'évolution des populations et de répondre aux enjeux de gestion durable.
- L'introduction de la stochasticité dans les scénarios épidémiologiques à l'échelle de la parcelle : prise en compte des événements rares mais à fortes conséquences sur la dynamique ultérieure du système (dispersion à longue distance, introduction d'espèces, extinction locale).
Collaboration avec MIA et EFPA

En termes de biologie intégrative, on relève donc une participation active de SPE à différentes échelles dans les champs thématiques suivants:

- Génomique fonctionnelle et biologie intégrative des interactions biotiques autour de la plante cultivée
- Découverte de gènes – outils pour la génomique fonctionnelle des bioagresseurs et des symbiotes
- Mécanismes du pouvoir pathogène des bioagresseurs et réactions de défense des hôtes (plantes, insectes)
- Bases biochimiques, génétiques et neurobiologiques des interactions entre insectes et avec les plantes
- Bioinformatique – datamining
- Ecologie des communautés et fonctionnement des écosystèmes cultivés

Les champs interdisciplinaires à la frontière entre génomique et génétique, entre génétique et dynamique apparaissent comme des fronts de science particulièrement pertinents à explorer par rapport aux problématiques des relations plantes-bioagresseurs.

2.8 Dans l'Enseignement supérieur et la Recherche

L'objectif est de donner un aperçu général des démarches de Biologie intégrative végétale dans différentes structures.

2.8.1 Les Ecoles doctorales et les IFR concernés

Les Ecoles doctorales et les IFR recensées (23 ED et 14 IFR) abordent la Biologie intégrative dans des domaines variés (*annexes 8 et 9*). Les couples ED/IFR d'un même site couvrent différentes situations :

- des ensembles plus centrés sur la Biologie intégrative (Sciences de la vie et de la santé) concernant surtout l'Homme, le domaine animal ou les micro-organismes, et marginalement le domaine végétal ;
- exceptionnellement quelques ensembles plus nettement ciblés sur la Biologie intégrative végétale (ED 145 Sciences du végétal, Paris XI ; ED 167 Biologie des systèmes intégrés, Montpellier 2 ; ED 223 La logique du vivant, Paris 6 ; ED 423 Des génomes aux organismes, Evry).

2.8.2 Biologie intégrative au CNRS & au CEA

Le CNRS a récemment ajouté l'adjectif « intégrative » au nom de la section « Biologie végétale » qui est devenue la section « Biologie végétale intégrative » (section 28). Ce changement illustre l'idée que l'intégration est centrale en Biologie végétale. Il faut souligner que CNRS et INRA ayant en commun de nombreuses unités mixtes, le développement d'approches intégrées en Biologie végétale a très souvent bénéficié de la synergie entre les deux organismes. C'est en particulier le cas en Ile de France sud (Pôle vert), à Montpellier, Toulouse, Lyon, Nice, Dijon...

La biologie végétale intégrative au CNRS concerne essentiellement une intégration horizontale, cherchant à interpréter les données génétiques, moléculaires et cellulaires issues des approches globales produites dans les domaines abordés dans les équipes et unités dépendant de la section 28 : métabolisme, signalisation, développement, interactions plantes-microorganismes, etc... Au CEA, la même stratégie est mise en œuvre (tant à Cadarache qu'à Grenoble). Dans tous les cas, il s'agit de passer de la production de données brutes à une information biologique pertinente. Le programme interorganismes de *Toxicologie environnementale* (initié par le CEA) illustre un ensemble de projets de recherche intégrée : un exemple en Biologie végétale concerne l'impact des métaux lourds sur le fonctionnement des plantes à divers niveaux d'organisation.

Il existe aussi au CNRS diverses équipes s'intéressant à une approche intégrée plus transversale (passage d'espèces modèles à d'autres espèces). Un exemple particulier concerne la biologie marine où les modèles biologiques sont d'une très grande diversité phylogénique. Enfin, il existe des sites où la diversité des unités de Biologie végétale permet d'intégrer les données recueillies à tous les niveaux d'étude : du gène à la cellule et à la plante jusqu'aux populations et aux écosystèmes. Un exemple remarquable est le site Gif/Orsay qui rassemble 5 unités³ regroupées au sein de l'IFR87 (*La plante et son environnement*, directrice H.

³ L'Institut des Sciences du Végétal (ISV, UPR 2355 CNRS, H. Barbier-Brygoo), le laboratoire Ecologie, Systématique et Evolution (P. H. Gouyon), l'Institut de Biotechnologie des Plantes (IBP, UMR CNRS/UPS, M. Kreis), l'unité de Génétique Végétale du Moulon (UMR CNRS/INRA, D. de Vienne), et l'équipe de Photosynthèse et télédétection (I. Moya), unités dépendant chacune de plusieurs partenaires institutionnels (Université Paris-Sud, CNRS, INRA, INA-PG, ENGREF).

Barbier-Brygoo) menant des recherches complémentaires recouvrant l'étude du développement des plantes et l'analyse de leur génome, l'étude des relations entre les plantes et leur environnement biotique et abiotique, la description des bases génétiques et moléculaires de la variation de caractères à déterminisme complexe et l'exploitation de ces données en sélection végétale, et l'analyse de la dynamique des populations et espèces de leur environnement.

2.8.3 La Biologie intégrative végétale au CIRAD

Le projet Oryzon réunit des chercheurs CIRAD spécialisés en biologie moléculaire, génétique, écophysiologie, sélection variétale, agronomie et modélisation dans un effort pour intégrer les différentes échelles d'étude considérées, de la génomique fonctionnelle du riz à l'étude du comportement de la culture au champ. Il s'agit de développer les connaissances et les outils de prédiction du comportement phénotypique de la plante et des peuplements de riz, en fonction du génome et de l'environnement, en vue de concevoir et de réaliser des idéotypes de riz.

Le riz est un matériel biologique de choix. Outre la séquence complète du génome de plusieurs variétés de riz, on dispose de plusieurs dizaines de milliers de mutants d'insertion (dont ceux produits dans le cadre Génoplante) et de la capacité à tester la fonction de gènes par ARNi. Le riz étant également une culture de première importance mondiale, de nombreux résultats proviennent d'équipes internationales, notamment asiatiques. Enfin, le riz est probablement la céréale qui décline la plus large gamme d'adaptations agro-écologiques. Les aspects génériques du projet faciliteront l'étude et l'amélioration d'autres graminées cultivées, blé, sorgho, maïs, canne à sucre et plantes fourragères.

Le projet est divisé en quatre sous projets complémentaires : 1) Découverte et caractérisation fonctionnelle de gènes ; 2) Analyse de la plasticité phénotypique de la plante ; 3) Etude du comportement des génotypes à l'échelle du peuplement végétal ; 4) Modélisation.

Le projet est mis en œuvre à Montpellier en liaison avec trois terrains aux Philippines, en Camargue et à Madagascar représentatifs respectivement des rizicultures irriguée tropicale, irriguée tempérée et pluviale.

Le projet a aujourd'hui le statut de projet fédérateur du Cirad et bénéficie à ce titre de financements incitatifs pour son animation. Le groupe mobilisé est unique en Europe, mais il a vocation à établir des coordinations avec d'autres groupes français et européens travaillant sur les céréales.

[Voir Annexe 10 –Projet Oryzon, JC GLASZMANN]

2.9 La Biologie intégrative au niveau international

2.9.1 Revues BIV/Colloques

Une dizaine de titres de journaux récents en rapport avec le domaine ont été répertoriés (*Annexe 11*) dans les domaines de la biologie intégrative, de la biologie des systèmes et de la bio-informatique. Le groupe de presse NATURE publiera à compter de mars 2005 une nouvelle revue intitulée « Molecular Systems Biology ». Il n'est donc pas crucial de proposer la création de nouveaux supports pour valoriser les productions scientifiques de ces domaines.

La communauté scientifique organise aussi des séminaires et conférences régulières sur ces thématiques. La liste des manifestations programmées en 2005 en témoigne (*Annexe 12*). En outre, des formations prédoctorales et doctorales sont dévolues à la biologie intégrative et systémique dans tous les domaines d'application, en Europe et aux Etats-Unis (*Annexe 14*)

2.9.2 Equipes internationales et champs d'application (*Annexe 13*)

Le site http://www.sysbio-muenchen.de/sb_int.html fournit une liste importante d'équipes impliquées dans la biologie systémique. L'essentiel des équipes ainsi identifiées travaille sur des modèles animaux, et en général sur la biologie humaine. Le point d'ancrage des équipes est très souvent la bio-informatique. On notera l'implication du US Department of Energy dans des appels à projet pour mieux utiliser les bactéries dans des fonctions environnementales. Le NSF semble plus impliqué dans les approches intégratives. L'approche systémique de la biologie des micro-organismes semble plus présente que pour les végétaux. Une autre caractéristique est que les différentes équipes constituant un groupe de biologie systémique sont localisées le plus souvent dans la même ville, voire au sein de la même Université : la proximité géographique n'est peut-être pas une condition à l'élaboration d'un groupe de biologie systémique, mais visiblement elle en facilite l'émergence.

Parmi les « Challenge Programs » du CGIAR, véritables programmes transversaux de recherche en coopération à l'échelle internationale lancés en 2004 avec les Centres Internationaux de la Recherche Agronomique, on relèvera les thèmes :

- Biofortification des variétés végétales pour leur teneur en micronutriments.
- Génomique et valorisation de la biodiversité dans les agricultures du Sud.

3 Les approches et les verrous méthodologiques

Des verrous méthodologiques à l'approche biologie intégrative végétale ont été identifiés dans les 5 domaines suivants :

- Espèces modèles en biologie végétale.
- Génomique végétale : l'existant et l'émergent.
- Génétique et Phénotypage.
- Bioinformatique et génomique comparative.
- Les nouveaux défis de l'écophysiologie.

3.1 Espèces modèles en biologie végétale

En biologie, un organisme modèle est une espèce vivante représentative d'un groupe taxonomique, qui possède des propriétés facilitant l'expérimentation et la démonstration de principes biologiques généraux. L'ensemble des connaissances qui en sont issues constitue lui-même un modèle de fonctionnement de l'organisme, au sens classique du terme, c'est à dire une représentation simplifiée d'un système, visant à rendre intelligibles ses aspects essentiels sous une forme suffisamment simple pour en permettre une formalisation mathématique.

La notion d'organisme ou espèce-modèle est centrale pour la biologie moderne. Elle a pourtant donné lieu à très peu d'analyses théoriques et critiques, tant de la part des biologistes eux-mêmes que de la part des philosophes et épistémologues. Or il importe de mettre cette notion en question, dans la mesure où une part importante des travaux et des moyens est focalisée sur un petit nombre d'espèces dites "modèles", en particulier pour les travaux de biologie fondamentale. Les connaissances issues de ces travaux représentent donc une part de plus en plus importante de la théorie biologique actuelle, et il importe de bien identifier les conséquences implicites de tels choix.

En ce qui concerne la biologie végétale, le modèle dominant est bien évidemment *A.thaliana*, pour des raisons qui, à l'origine, tiennent à sa biologie, particulièrement propice à l'expérimentation en laboratoire, ainsi qu'à l'organisation de son génome, simple et d'une extrême compacité. Le nombre de publications consacrées à cette espèce croît de façon exponentielle depuis 1990 (plus de 2.000 publications par an depuis l'année 2000). Devant un tel succès, il importe de s'interroger sur les bénéfices et les risques associés à cette approche sur notre vision de la biologie et de la diversité végétales.

L'intérêt scientifique de la concentration de moyens, d'outils, d'informations, de ressources est une évidence qui s'impose d'elle-même en ce qui concerne *A.thaliana* (voir par exemple <http://signal.salk.edu/>). En résumé, l'espèce modèle permet :

- de diversifier et mutualiser les outils, les méthodes et les ressources biologiques : gain de temps et d'efficacité...
- de partager, de croiser et de globaliser l'information : meilleure prise en compte de la complexité, des interactions...
- de favoriser l'inter-disciplinarité en mobilisant les chercheurs sur un même objet biologique : diversité de points de vue, de méthodes, possibilité de synthèses et de synergies novatrices
- de créer une communauté scientifique organisée, capable de mobiliser des moyens importants pour des programmes d'envergure

A l'évidence, beaucoup de ces points concernent directement la biologie intégrative et sont en plein accord avec ses objectifs. Il paraît donc important de préconiser une focalisation des efforts sur un nombre très réduit de systèmes modèles, au moins dans un premier temps, qui servira de période exploratoire pour l'élaboration des concepts et méthodes à mettre en jeu.

C'est la condition de l'émergence de réelles synergies interdisciplinaires. Même si l'exploration de la diversité intra-spécifique est d'ores et déjà possible et nécessaire, il paraît relativement peu raisonnable de se placer d'entrée dans une optique multi-espèces.

En revanche et malgré son intérêt heuristique indéniable, la notion d'espèce modèle porte en elle plusieurs contradictions dont on a généralement peu conscience :

En premier lieu, la notion d'espèce représentative d'un groupe est contradictoire à plus d'un titre : d'abord parce les limites du groupe taxonomique considéré ne sont pas prédéfinies mais à examiner au cas par cas (*A.thaliana* est-il un modèle pour le genre *Arabidopsis*, pour la famille des Brassicaceae, pour les Dicotylédones, pour les Angiospermes, pour les plantes terrestres ?

Dans le couple unicité/diversité du monde vivant, la notion d'espèce modèle considère essentiellement le premier terme. D'une manière générale, quelle valeur de généralité accorder aux théories biologiques dès lors qu'on dépasse le cadre du modèle sur lequel elles ont été établies ? La phylogénie et l'analyse comparée ne sont que rarement prises en compte, alors que la capacité de généralisation dépend fortement de la position de l'espèce dans le groupe et du statut ancestral ou dérivé des caractères étudiés.

Enfin, et c'est probablement l'aspect le moins souvent mis en exergue, les critères de choix des modèles biologiques biaisent de façon importante les informations qu'on en tire. En effet, même s'ils sont changeants en fonction des disciplines concernées (biologie du développement, génétique, écologie) et selon les époques, ils sont relativement constants d'un organisme à l'autre, à tel point qu'on peut parler d'un véritable « syndrome d'appropriation » des espèces modèles par les biologistes.

Parmi les caractéristiques recherchées, on peut citer notamment qu'il s'agit la plupart du temps d'organismes de petite taille, à développement rapide et invariant, à maturation sexuelle rapide et temps de génération court, prolifiques, possédant un génome simple... Or, quasiment tous ces critères de sélection sont susceptibles d'apporter des biais non explicites ; par exemple, les organismes modèles sélectionnés sur ces bases risquent fort de partager un certain nombre d'adaptations au développement rapide, souvent des caractères dérivés absents d'espèces à développement plus lent. Idem pour la durée du temps de génération, critère de choix pour les généticiens. De la même manière, le choix d'organismes pour lesquels les processus de développement sont invariants d'un individu à l'autre et par rapport à l'environnement (on parle de développement canalisé) tend à gommer les effets de l'environnement et à donner une vision trop déterministe du développement.

Les organismes modèles constituent donc une fraction infime de la biodiversité actuelle, et pourtant l'essentiel de nos connaissances en biologie provient de l'étude de quelques organismes emblématiques, qui ont tous été sélectionnés sur les mêmes bases.

Au regard des avantages cités plus haut, il paraît malgré tout évident qu'il faut tirer parti de l'existence des quelques espèces modèles bien établies internationalement chez les plantes (*Arabidopsis*, Riz, *Medicago truncatula*...). Il faut par ailleurs initier une réflexion approfondie sur les dangers d'une trop forte concentration et jeter les bases méthodologiques et épistémologiques pour la généralisation des connaissances des systèmes modèles "de laboratoire" vers les espèces sauvages ou cultivées.

3.2 Génomique végétale : l'existant et l'émergeant

Dans beaucoup de propos exprimant le besoin d'un renouveau de la biologie vers une vision plus intégrée et systémique, le rôle décisif de la génomique et des approches "haut-débit" est

mis en avant. Cependant, les opinions et les propos sont singulièrement contrastés sur les apports de ces méthodologies à notre compréhension du vivant. Certains voient dans la puissance de ces avancées technologiques une chance extraordinaire pour l'avancée des connaissances en biologie, qui accèderait enfin au statut de science "dure". D'autres déplorent l'utilisation aveugle (et dispendieuse) de ces techniques qui ne serviraient qu'à accumuler des masses de données peu fiables et difficiles à exploiter de façon pertinente.

A partir du début des années 90, la génomique connaît en effet un développement extraordinaire grâce à plusieurs avancées conceptuelles et technologiques majeures. A l'origine de ce mouvement, on peut citer notamment les marqueurs moléculaires, les méthodes d'analyse de locus à effets quantitatifs (QTL), ou encore les techniques mégabase (électrophorèse en champ pulsé, clonage de grands fragments) et le séquençage automatique. Ces approches permettent de caractériser les bases moléculaires de caractères repérés grâce aux approches génétiques et cartes génétiques. Ensuite, l'automatisation et la miniaturisation aidant, il est devenu possible d'élaborer de vastes catalogues de séquences de gènes, d'obtenir des séquences complètes de génomes modèles, d'aborder l'étude du fonctionnement des cellules à travers la caractérisation exhaustive de l'expression des gènes...

La génomique s'intéresse aux génomes en tant que tels et peut être définie comme l'étude globale de leur structure et de leur fonctionnement. Au delà des approches de génomique « structurale » (cartographie physique et génétique, inventaire de gènes, séquençage), les outils de génomique « fonctionnelle » sont en pleine expansion et ont investi tous les champs de la biologie. Dans ce cadre, l'acception courante de la biologie « intégrative » sous-entend la compréhension des processus biologiques par la mise en oeuvre de méthodes globales (en – omique) et de niveaux d'analyse variés (du gène à l'écosystème).

Il ne s'agit pas ici de dresser un inventaire complet des technologies disponibles ; de ce point de vue on pourra se référer utilement à une publication récente ("La génomique en biologie végétale", INRA Editions, 2004). Par contre il semble utile de préciser, pour les grandes avancées technologiques récentes qui ont bouleversé les approches biologiques, d'en cerner les contours, les forces et les limitations en ce qui concerne la biologie intégrative végétale. Dans la mesure où la plupart des approches sont récentes et encore dans une phase de développement technologique rapide, un recul critique documenté est difficile dans bien des cas. Sans préjuger des nouveaux développements et améliorations à venir, on peut cependant, en l'état, assez facilement identifier un certain nombre de limitations inhérentes aux approches en question.

3.2.1 Génomique Structurale et séquençage

3.2.1.1 Cartographie et marqueurs moléculaires

Depuis les années 80, la mise au point de marqueurs moléculaires sous leurs différentes formes (RFLP, RAPD, microsatellites, etc) ont permis la construction de cartes génétiques saturées à un coût relativement modeste. Fournissant un canevas génétique fiable et exhaustif, elles permettent le repérage et la localisation précise de locus à effets discrets ou quantitatifs (D. de Vienne, 1998).

Ces nouveaux marqueurs ont permis en quelques années la construction de cartes génétiques d'une précision inégalée sur la plupart des espèces végétales cultivées et plusieurs espèces modèle. Des travaux de caractérisation de locus à effets quantitatifs sont alors menés sur de nombreuses espèces cultivées. Les résultats de tels travaux sont souvent directement utilisables en sélection (sélection assistée par marqueurs : SAM).

De façon concomitante, le développement de techniques de clonage et d'analyse de grands fragments génomique sous forme de clones dans des bactéries (BACs) ou levure (YACs) permet la réalisation de cartes physiques. Ces techniques lourdes sont tout d'abord développées sur *Arabidopsis* ou le riz dans les années 1995 avant d'être étendues par la suite à des espèces cultivées dont les génomes sont plus complexes, en particulier sous l'incitation des programmes "génomes de plantes" financés par la NSF, qui consent des investissements lourds dans le domaine.

Les marqueurs moléculaires, qui peuvent à la fois révéler des polymorphismes génétiques qu'on peut cartographier, et dont la région d'origine est facilement repérable sur la molécule d'ADN elle-même, permettent d'ancrer les deux types de cartes chromosomiques, génétique et physique, dont les échelles (exprimées en centimorgans et kilobases) ne sont pas colinéaires, loin s'en faut.

Ces nouvelles techniques d'analyse, qui ont apporté des gains de puissance considérables, ont permis un renouvellement complet de la discipline et un rapprochement entre génétique formelle et génétique moléculaire, convergence enfin rendue possible par les outils de la génomique. Grâce à la densification des cartes génétiques et aux techniques de manipulation de grands fragments d'ADN, il est devenu possible d'approcher à un même niveau de résolution chromosomique le locus du généticien et le gène du biologiste moléculaire.

3.2.1.2 Séquençage haut-débit (EST et génomique)

Le progrès et l'automatisation (partielle) des techniques de séquençage ont permis dès le début des années 90 d'envisager des approches de séquençage systématique. Au départ, ces dernières se sont attachées à dresser des catalogues de séquences exprimées (expressed sequence tags, EST). L'approche EST, efficace et relativement peu coûteuse, a connu une fortune extraordinaire pour les plantes (notamment les espèces cultivées), puisque sur environ 140 espèces pour lesquelles plus de 100 000 EST sont disponibles en Octobre 2004, près de 40% sont des plantes (consulter dbEST: database of "Expressed Sequence Tags")

Cependant, l'approche EST a plusieurs limitations que ne compensent pas les multiples améliorations techniques développées au fil du temps : en premier lieu, les séquences sont de qualité médiocre et très partielles -quelques centaines de nucléotides- sauf quand un séquençage de transcrits pleine longueur est associé au programme (auquel cas le coût change du tout au tout). D'autre part, seuls les transcrits les plus exprimés et les plus constitutifs sont détectables, les ARNm peu abondants ou exprimés dans des conditions très spécifiques n'étant en général pas représentés. Il en découle une très forte redondance des données relatives aux ARNm les plus abondants qui peuvent représenter la majorité des enregistrements. On estime en général que des approches EST d'une profondeur raisonnable ne permettent pas d'identifier plus de 55-60% des transcrits d'une espèce.

D'où la volonté affirmée dès le début des années 1990 en Europe et aux Etats-Unis d'engager le séquençage du génome d'*A. thaliana*. Un consortium international (Arabidopsis Genome Initiative) a été mis sur pied en 1996 avec un partage du travail entre les divers laboratoires européens, américains et japonais, et avec pour objectif la caractérisation complète des cinq chromosomes pour 2004. L'achèvement de la séquence génomique complète de l'espèce modèle *A. thaliana* est pourtant annoncé dès décembre 2000.

L'obtention de cette séquence d'excellente qualité, au delà du tour de force technologique et organisationnel, a permis la floraison d'une multitude d'analyses et commentaires qu'il est impossible de résumer ici. Rappelons simplement que sur ce génome d'environ 125 mégabases, on dénombre environ 25 à 30 000 gènes. La plupart de ces gènes appartient à des

familles multigéniques, seul un tiers est en copie unique. L'organisation de ce génome est nettement plus complexe qu'attendu et révèle une histoire évolutive riche en rebondissements.

Dans les dernières années, plusieurs autres programmes de séquençage de génomes d'angiospermes ont été lancés : le travail de séquençage est en phase de finition pour le riz et pour le peuplier. Il ne fait que démarrer pour *Medicago truncatula*, la tomate, ou *Physcomitrella*. Des programmes de séquençage génomique partiel (limité aux régions non répétées) sont par ailleurs initiés pour des espèces d'intérêt agronomique comme le maïs ou le soja.

3.2.2 Génomique fonctionnelle : La saga des "-omiques"

Les séquences complètes du génome d'*A. thaliana*, du riz et du peuplier sont donc maintenant disponibles, d'autres plantes sont en cours de séquençage et de vastes collections d'ESTs existent également pour plusieurs dizaines d'espèces de plantes. Or, la plupart des informations fonctionnelles dérivées du traitement bioinformatique des données doivent être considérées comme hypothétiques en l'absence de confirmation expérimentale. Il s'agit donc maintenant de renverser l'approche traditionnelle de la génétique (quels sont les gènes derrière cette fonction ?) en un nouveau paradigme "post-génomique" (quelles sont les fonctions derrière ce(s) gène(s) ?).

Comment faire face à l'ampleur de cette tâche ? Sur plus de 25.000 gènes identifiés chez *A. thaliana*, une information fonctionnelle fiable n'est disponible que pour quelques milliers d'entre eux. Ce catalogue complet a permis d'établir des classifications fonctionnelles et a donné pour la première fois une idée de la structure globale et du contenu en gènes d'un génome de plante supérieure. Cependant, sur les 25 500 gènes repérés dans la publication originale, seulement près de 10% ont une fonction biologique définie expérimentalement. Près de 30% ne présentent aucune similarité de séquence avec un gène de fonction connue, et pour les 60% restant, on peut définir une catégorie fonctionnelle sur la base de comparaison de séquence. Cependant, pour la plupart, ces prédictions restent sujettes à caution tant qu'une validation expérimentale n'a pas été entreprise. C'est dire l'ampleur de la tâche qui attend les biologistes dans les prochaines années...

La génomique fonctionnelle propose des approches à haut débit permettant l'acquisition rapide de connaissances sur la fonction des gènes. C'est en particulier l'objectif ambitieux que s'est fixé la communauté internationale des biologistes d'*A. thaliana* pour les dix prochaines années⁴. Cela nécessite donc de rassembler pour chaque gène et protéine un faisceau d'informations sur la structure, la régulation, l'activité biochimique et la localisation sub-cellulaire, les interactions avec d'autres acteurs moléculaires, la variabilité allélique naturelle, etc. Il est en effet d'ores et déjà clair qu'aucune méthodologie particulière, si précise et exhaustive soit-elle, ne peut prétendre à elle seule donner un sens biologique à la fonction d'un gène ou d'une protéine, même en se limitant à son environnement cellulaire ou tissulaire.

Les dernières années ont vu le développement spectaculaire de ces méthodologies d'analyse fonctionnelle « haut-débit », qui permettent d'analyser simultanément l'activité de vastes ensembles de gènes et de protéines, voire l'ensemble du génome. On peut y distinguer les

⁴ Voir par exemple ce résumé des objectifs du programme NSF "Arabidopsis 2010" : "To determine the function of all genes of a reference species within their cellular, organismal and evolutionary contexts by the year 2010". Même si l'ambition démesurée de ce raccourci n'est pas à prendre au pied de la lettre, cela représente bien l'état d'esprit dans lequel se placent la plupart des intervenants du domaine.

méthodologies d'étude de l'expression, visant à quantifier les variations des produits des gènes (transcrits et protéines) dans le temps, dans l'espace, au cours du développement, ou en réponse à divers stimuli de leur environnement biotique et abiotique.

3.2.2.1 Transcriptomique

Des diverses approches en "-omique", la transcriptomique est la plus ancienne, la seule qui présente un caractère réellement exhaustif, et également la seule qui a atteint un certain niveau de maturité et de stabilisation des techniques. A partir de la caractérisation exhaustive des populations de transcrits propres à 2 ou n états d'un système biologique, on cherche à identifier des transcrits qui présentent des variations significatives d'abondance d'un état l'autre, et à établir des groupes de co-expression. L'appartenance à un même groupe de co-expression peut être interprétée comme l'indication d'une implication dans une même fonction, et/ou d'une régulation commune.

Diverses difficultés d'ordre technique (taille et variabilité des échantillons biologiques, contrôle des sources de variation expérimentales, plans d'expériences...), ainsi que dans le traitement statistique des données, continuent à poser problème, et donnent lieu à de multiples recherches. La qualité des données dépend aussi en grande partie du soin apporté à la préparation et à la caractérisation des échantillons biologiques, avec un nombre de répétitions suffisant pour permettre un traitement statistique correct. D'autre part la quantité de matériel nécessaire et la difficulté d'isoler, en particulier chez les plantes, des populations de cellules homogènes, limitent fortement la possibilité d'aborder des régulations spatiales fines.

Par rapport aux approches expérimentales classiques, basées sur des hypothèses de travail explicites, que l'on cherche à valider ou invalider, la transcriptomique revendique souvent -et assume parfois- une option purement "exploratoire" et corrélative. Une critique fréquente de ce type d'approche "haut-débit" est liée à la difficulté d'interprétation pertinente des données. En effet, le caractère "massif" des données oblige à une extraction sélective de l'information forcément basée sur des critères statistiques, pour permettre, dans un deuxième temps, l'intervention de l'expertise du biologiste. Des variations de faible amplitude, qui peuvent pourtant avoir une grande pertinence vis à vis du phénomène biologique étudié, ne seront en général pas distinguées du "bruit de fond". Enfin et surtout, s'il est bien établi que l'étape transcriptionnelle est souvent un niveau majeur de régulation de l'expression génique, de nombreux résultats indiquent qu'il y a peu de corrélation entre la quantité de transcrit et celle de la protéine correspondante.

Il y a lieu alors de se méfier des tentations de "surinterprétation" des résultats obtenus à travers les approches de transcriptomique lourde. Au delà des interprétations en tendance, par grandes fonctions et groupes de gènes co-exprimés, le "sens" biologique et l'interprétation de ces variations quantitatives d'abondance de transcrit, en particulier leur signification et rapports de causalité par rapport au phénomène étudié, dépend en grande partie d'une validation expérimentale indépendante et d'un affinement ultérieur des résultats.

3.2.2.2 Protéomique et Biologie intégrative

L'analyse protéomique consiste à identifier, caractériser et quantifier les protéines d'une cellule, d'un tissu, d'un organisme placé dans un environnement donné. L'analyse rapide et à grande échelle des protéines est devenue possible grâce à la réalisation de trois conditions essentielles. D'une part, le succès des programmes de séquençage des génomes a permis d'identifier la séquence et la position de tous les gènes (et donc potentiellement d'identifier les protéines correspondantes) d'un nombre croissant d'organismes. D'autre part, les méthodes de séparation des protéines se sont affinées, par exemple grâce aux développements de l'électrophorèse bidimensionnelle. Enfin, les méthodes de spectrométrie de masse ont fait

des progrès spectaculaires en sensibilité et en résolution qui permettent de connaître avec une précision remarquable la masse (et la séquence) des peptides provenant de la digestion de mélanges protéiques. Le développement des bases de données protéiques et des outils bioinformatiques permet alors de déterminer avec une grande sûreté les protéines présentes dans ces mélanges.

Cependant, l'intérêt de l'analyse protéomique ne se restreint pas à l'établissement de l'inventaire des protéines d'un compartiment, d'une cellule, d'un tissu ou d'un organisme. Dans la perspective d'une approche intégrée, la dynamique d'un protéome est particulièrement importante pour comprendre le fonctionnement du vivant. Une telle approche est actuellement réalisable pour des études du transcriptome, mais de portée limitée dans le cas d'analyses du protéome. Les analyses protéomiques ne sont actuellement que des flashes ponctuels et non exhaustifs exploités par des comparaisons d'images de gel. Or le répertoire des protéines d'un compartiment cellulaire ou leurs modifications post-traductionnelles évoluent très largement au cours du temps, dans des conditions de fonctionnement normal ou pathologique, lors des divers phénomènes physiologiques sous l'action des facteurs environnementaux. C'est cette dynamique qu'il faut appréhender. Actuellement, les analyses protéomiques ne bénéficient pas de méthodes d'amplification qui permettent de quantifier au cours du temps l'évolution des protéines présentes à l'état de trace dans un échantillon.

D'importants verrous technologiques limitent donc encore une étude fine de la dynamique des protéomes. C'est en particulier le cas d'analyses protéomiques d'échantillons de taille limitée (groupes de cellules). De telles études nécessitent le développement d'outils de séparation et des progrès de la spectrométrie de masse vers une sensibilité accrue. L'utilisation combinée de spectromètres de masse à résonance cyclotronique par transformée de Fourier couplés aux méthodes d'ionisation modernes, de peptides synthétiques marqués avec des isotopes lourds et d'outils bioinformatiques nouveaux permettra d'envisager des études dynamiques à grande échelle des protéomes dans un avenir proche.

3.2.2.3 Biochimie des protéines et techniques d'imagerie cellulaire

Les approches de biochimie et de cytologie, malgré la difficulté à faire entrer de telles disciplines dans l'ère du "haut débit", ont connu récemment de nombreuses évolutions visant à simplifier, automatiser, et fiabiliser les analyses : clonage systématique des séquences codantes pour cribles d'interaction protéine-protéine, purification de complexes, expression hétérologue, construction de puces protéiques... De plus, des techniques basées sur la complémentation ou le transfert d'énergie inter-moléculaire (BiFc, FRET) permettent d'ores et déjà de tester et de quantifier des interactions moléculaires *in vivo*, ouvrant la voie à une véritable biochimie en temps réel des protéines dans leur contexte cellulaire.

Dans le domaine de l'imagerie cellulaire "*in vivo*", une véritable révolution est à l'œuvre grâce aux progrès de la microscopie photonique et à l'utilisation de marqueurs protéiques fluorescents. La mise au point de la microscopie confocale à balayage laser a en effet apporté la possibilité de travailler sur des échantillons biologiques vivants (une pointe racinaire, un méristème) sans fixation chimique, ni coupe. En particulier, la plupart des échantillons végétaux peuvent survivre plusieurs heures sur la platine du microscope. Grâce à une succession de "coupes optiques", on peut balayer l'échantillon et recueillir le signal selon différents plans optiques séparés par une fraction de micron. On peut ainsi reconstituer la structure tridimensionnelle du signal dans les cellules, et suivre son évolution au cours du temps. Les microscopes confocaux "spectraux" sont capables d'exciter et de recueillir simultanément la fluorescence dans plusieurs longueurs d'onde définies au nanomètre près : on peut ainsi suivre plusieurs marqueurs indépendants, dans le temps et en trois dimensions.

De façon concomitante, on dispose maintenant de nombreuses versions de marqueurs protéiques fluorescents comme la GFP ou la RFP (respectivement green fluorescent protein et red fluorescent protein) avec des spectres d'absorption/émission variés permettant l'utilisation simultanée de plusieurs marqueurs dans une même cellule et de s'affranchir de l'autofluorescence propre à certains tissus ou organites végétaux. L'expression *in vivo* de fusions protéiques entre ces marqueurs et des protéines d'intérêt permet de suivre de façon dynamique et non invasive leur localisation tissulaire, cellulaire et sub-cellulaire avec un degré de finesse sans précédent.

L'importance de ces méthodologies par rapport aux objectifs de la Biologie Intégrative est considérable, puisqu'elles permettent d'explorer l'activité des protéines individuelles et des réseaux d'interaction *dans leur contexte cellulaire, au sein de l'organisme*. Pour ce qui concerne les plantes, il s'agit d'une avancée d'autant plus décisive qu'une cellule végétale isolée, hors de son contexte tissulaire (comme un protoplaste), n'a que peu de rapport avec une cellule de plante, enchâssée dans une paroi pecto-cellulosique semi-rigide, connectée à ses voisines à travers des ponts cytoplasmiques (les plasmodesmes), et dont le devenir en terme de différenciation est déterminé non pas par son lignage cellulaire mais par sa position au sein du tissu.

3.2.2.4 Nanobiotechnologies et Biologie intégrative

Une révolution technologique, basée sur le développement des Nanosciences, arrive à propos pour envisager d'aller encore plus loin dans la compréhension du vivant. Grâce aux nanotechnologies, nous sommes désormais en mesure non seulement de caractériser et manipuler la matière à l'échelle du nanomètre mais aussi de fabriquer de nouvelles structures supramoléculaires et des outils miniaturisés à partir d'atomes ou de molécules. Ces nouvelles structures présentent de nouvelles propriétés qu'il est possible d'exploiter pour mieux comprendre le vivant. Ainsi, on sait fabriquer des nanocristaux aux propriétés semi-conductrices ou des nanocristaux fluorescents qui constituent des sondes particulièrement sensibles pour explorer, à l'échelle du nanomètre, de nombreuses molécules biologiques dans une cellule, un tissu ou un organisme vivant.

L'objectif ultime des approches de génomique à haut débit est d'aborder l'étude dynamique de molécules uniques dans des cellules uniques, au sein d'organismes vivants grâce au développement des nanotechnologies. Actuellement, notre connaissance fine des caractéristiques physiques des constituants cellulaires reste très rudimentaire. Dans ce domaine, les frontières auxquelles les chercheurs sont confrontés peuvent être illustrées par une série de questions : en quelle quantité les constituants cellulaires sont-ils présents et dans quel compartiment ? Quelle est la dynamique de leur expression dans le temps et l'espace ? Avec quelles molécules interagissent-ils, avec quelle affinité ? Quelles sont les forces générées, les mouvements, les changements de conformation ? Ce sont des notions quantitatives et il sera impossible de progresser dans notre compréhension des structures cellulaires ou des réseaux de régulation (par exemple) si nous sommes incapables de mesurer physiquement l'ensemble de ces caractéristiques physiques.

Les biologistes ont besoin d'outils pour collecter des données expérimentales, les stocker, les analyser et les diffuser. Ils ont donc besoin d'une part de nouvelles méthodes physiques, de nouveaux outils pour déterminer ces paramètres et d'autre part de stratégies permettant de manipuler ces énormes quantités de données. C'est la convergence naturelle entre la Biologie, les Nanotechnologies et l'Informatique. Le développement des puces à ADN est un exemple de l'intérêt de pouvoir réaliser des analyses moléculaires parallélisées et à haut débit. En miniaturisant, parallélisant, intégrant et automatisant, les nanotechnologies permettent de concevoir de nouveaux outils d'analyse des molécules du vivant (puces à cellules, par

exemple) permettant ensuite d'intégrer les données obtenues au niveau subcellulaire, cellulaire ou du tissu.

Cependant, la multiplication de données quantitatives précises sur le fonctionnement du vivant n'est pas suffisante pour le comprendre et nécessite alors le passage à un niveau de compréhension supérieur, par exemple être en mesure de cartographier les interactions que l'on a décrites lors d'analyses quantitatives et parallélisées. Le développement des micro- et nanotechnologies pour le vivant devrait permettre à terme de disposer d'outils d'intégration pour analyser les processus biologiques dans leur ensemble. C'est, par exemple, tout l'enjeu du développement des puces à cellules ou de la manipulation des molécules uniques *in vivo*. Ces approches ouvrent la voie à l'analyse fonctionnelle des nombreux modèles biologiques générés par les approches génétiques systématiques (mutants, RNAi, etc...). L'informatique intervient alors pour donner du sens à ces données, l'apport des outils de modélisation et de bioinformatique étant, là encore, essentiel.

A long terme, on peut raisonnablement envisager un va et vient entre Biologie et Nanosciences, ces dernières apportant des outils et des concepts nouveaux pour disséquer le vivant qui en retour peut apporter des modèles pour concevoir des outils nouveaux. En d'autres termes, la Nature a développé au sein des cellules des voies, réseaux, structures, machines d'une complexité et d'une finesse remarquables et qui peuvent servir de modèles pour concevoir des nano-outils, par exemple pour une utilisation biomédicale.

3.3 Génétique et Phénotypage

Les approches de génétique ont ceci de particulier par rapport aux approches moléculaires et « omiques » qu'elles se placent par construction au niveau de la relation génotype-phénotype. Elles permettent d'explorer la complexité de cette relation, de mieux comprendre ce qu'est un phénotype et comment il se construit. C'est pourquoi elles occupent une place particulière et indispensable dans la panoplie des outils du génomicien.

3.3.1 Mutagenèse et Génétique inverse

Un grand volet des approches de génomique fonctionnelle, en particulier celles menées sur *A. thaliana*, concerne la génétique : la disponibilité de collections de mutants et la possibilité d'inactivation de gènes par génétique inverse y sont par exemple un atout majeur pour l'analyse fonctionnelle.

Notons que malgré des années d'efforts de la part de groupes importants, les tentatives de maîtrise de la recombinaison homologue pour la modification ciblée de gènes *in situ* n'ont toujours pas abouti chez les plantes supérieures. Dans le monde végétal, cette approche n'est disponible que chez la mousse *Physcomitrella patens*.

[Voir Annexe 14 La recombinaison homologue, F. Nogue]

3.3.2 Génétique quantitative et génétique d'association

Les approches de génétique quantitative (recherche et caractérisation de locus à effet quantitatif, génétique d'association), s'appuyant sur les informations et outils de la génomique, devraient permettre la dissection fine de caractères à déterminisme complexe, et replacer les caractères biologiques dans une perspective écologique et évolutive.

3.3.3 Phénotypage haut-débit

Une des limitations actuelles de l'analyse fonctionnelle de la variabilité génétique réside dans la difficulté qu'il y a à caractériser le phénotype des plantes en relation avec l'environnement réellement perçu par ces plantes. En effet l'analyse des interactions génotype-environnement qui est à la base de toute tentative d'adaptation des plantes à des milieux contraignants ou à des modes de conduite de culture différents implique d'établir des relations de causalité entre l'expression d'un phénotype et les caractéristiques précises du milieu perçu par la plante. Les simulations sur plantes virtuelles, telles qu'elles ont été présentées ci-dessus, ne sont qu'une approche très incomplète dans la mesure où, si elles permettent effectivement de « balayer » une gamme la plus large possible de milieux, elles ne permettent pas de prendre en compte la gamme de variation génotypique qui est *a priori* inconnue.

Cette approche par simulation doit donc être complétée par une approche expérimentale permettant de mesurer des critères phénotypiques sur un grand nombre de génotypes (plusieurs milliers) dans des milieux parfaitement maîtrisés. La maîtrise de ces installations expérimentales requiert le recours à des instrumentations lourdes et automatisées (robotique) qu'il convient impérativement de développer si l'on veut pouvoir acquérir des données de phénotypage à un rythme élevé, non limitant par rapport aux débits du génotypage et du séquençage.

En particulier, pour les approches de génétique quantitative, la puissance statistique du dispositif expérimental mis en place est le paramètre essentiel. Il s'agit de permettre la détection du maximum de locus, même mineurs, sans inclure de faux positifs, de les positionner sur le génome avec des intervalles de confiance les plus réduits possible, et d'en estimer précisément les effets et les relations d'épistasie : dans cette optique, le paramètre essentiel est l'effectif des génotypes disponibles pour l'analyse phénotypique. L'ensemble des ressources biologiques et génétiques en cours de développement sur certaines espèces comme *Arabidopsis*, où plusieurs dizaines de milliers de lignées recombinantes seront disponibles d'ici peu, devrait fournir à terme un potentiel expérimental unique, qui n'a encore jamais été développé pour aucune espèce vivante.

Afin de tirer parti au mieux de ce potentiel, il faut optimiser la qualité des estimations phénotypiques, la possibilité de répétitions expérimentales, et la réduction des sources de variation non contrôlée, qui jouent fortement sur l'héritabilité des caractères et plus généralement sur la puissance du dispositif. La qualité de l'estimation phénotypique doit donc concorder avec les ressources mises en place.

[Annexe 6 – Utilisation des approches écophysiologies pour le phénotypage à l'échelle de la plante entière : approche comparative et modélisation J. LECOEUR]

3.3.4 Imagerie et fluorescence : vers la fonction par l'image

L'utilisation des méthodes de mutagenèse aléatoire (mutagenèses chimique ou physique ou insertionnelle) est aujourd'hui l'un des outils majeurs pour caractériser la fonction d'un gène. Pour les végétaux supérieurs, l'expérimentateur dispose maintenant de nombreuses banques de mutants comme chez les espèces modèles *A. thaliana* et *O. sativa* de moyens relativement limités pour aborder de manière systématique la mise en évidence d'un phénotype. Mis à part des changements morphologiques ou d'aspects (par ex. couleur, nécroses) qui permettent de caractériser facilement un phénotype mutant, les mutations ont le plus souvent une expression phénotypique subtile et difficilement détectable que seuls des criblages biochimiques ou moléculaires et des tests biologiques, permettent la mise en évidence. L'utilisation de telles approches lors de criblages en série relève d'approches systématiques.

Des méthodes biophysiques basées sur l'analyse de signaux radiatifs (fluorescence, émission IR) permettent d'aborder quelques-unes des grandes fonctions des plantes de façon « non-invasive » et offrent des solutions pertinentes pour une démarche de criblage fonctionnel. Ainsi, la fluorescence rouge, proche infrarouge des chlorophylles des feuilles dont le rendement dépend de l'aptitude de la photosynthèse à utiliser l'énergie lumineuse absorbée permet de quantifier l'activité photosynthétique et de caractériser l'état physiologique foliaire. Dans le domaine spectral infrarouge lointain (IR), les variations de rayonnement IR émis par la feuille peuvent être utilisées pour quantifier les variations de la transpiration et de l'ouverture stomatique. D'autres signaux optiques telles la luminescence, la fluorescence bleue-verte induite sous lumière UV et la réflectance renseignent sur les changements respectivement de l'état de peroxydation des lipides, de composition et de structure des feuilles.

L'imagerie de certains de ces signaux au stade plantules a permis la mise en place de cribles d'altération de fonction ne nécessitant pas d'analyses destructives et peu consommateurs de temps. Grâce à des cribles à grande échelle, de nouveaux gènes codant pour des facteurs moléculaires impliqués dans les processus fondamentaux gouvernant ces signatures optiques ont été identifiées. Le succès de l'application de tels cribles dépend en grande partie du choix des conditions environnementales qui définissent la spécificité de la signature examinée. En général, des transitions des conditions physiques environnementales tels l'éclairement, la composition gazeuse de l'atmosphère (CO₂, H₂O) doivent être utilisées pour révéler un phénotype d'intérêt utilisable.

Par ailleurs, l'imagerie de ces signaux en permettant d'intégrer la dimension spatio-temporelle de la fonction fournit une approche d'avenir pour aborder à l'échelle de la plante les relations entre fonctionnement cellulaire et organisation tissulaire. Cette approche complète ainsi les stratégies de marquage *in vivo* et d'imagerie utilisées en biologie cellulaire.

3.4 Bioinformatique et Génomique Comparative

La biologie des systèmes⁵ étudie, à différents niveaux d'intégration, le fonctionnement des processus biologiques et ses perturbations lors de modifications physiologiques ou en réponse à des modifications de l'environnement. Ce domaine, dont les fondements sont issus de la biologie théorique (avec la publication du modèle de B. Goodwin, 1965), connaît actuellement un engouement important pour deux raisons :

- i) La croissance exponentielle des données génétiques, moléculaires et cellulaires issues des approches globales à gérer et à interpréter,
- ii) Les progrès de l'informatique en rapport avec la possibilité d'implémenter et d'évaluer les formalismes mathématiques qui modélisent les comportements des réseaux d'interactions impliqués dans les voies de régulation et de signalisation.

3.4.1 Annotation des séquences

Une fois la séquence établie (et même avant), un travail bioinformatique d'annotation est entamé, dans le but d'associer aux données brutes de séquence des informations pertinentes d'un point de vue biologique.

Il s'agit dans un premier temps d'identifier les gènes, leur organisation en introns et exons, en s'appuyant tant sur les caractéristiques intrinsèques des séquences à l'aide d'algorithmes

⁵ Rédigé à partir de documents fournis par Y. Vandembrouck (CEA-Grenoble)

d'analyse appropriés, que sur la comparaison avec les séquences disponibles, notamment les EST. Cette phase d'identification est d'ailleurs loin d'être triviale, et seule la séquence d'ARN messagers pleine longueur permet une validation définitive des modèles de gène. Cette ressource n'est disponible chez les plantes que pour *A. thaliana* à l'heure actuelle.

Dans un deuxième temps, les comparaisons de séquences avec gènes, protéines et motifs déjà connus permettent d'enrichir l'annotation fonctionnelle. Cette approche comparative permet dans un nombre important de cas (60% des gènes/protéines pour Arabidopsis) d'associer une information de type fonctionnel aux séquences brutes. Cependant, cette annotation doit être considérée avec précaution, sachant que :

- elle concerne des gènes dont la structure est souvent elle-même prédite
- elle est générée de façon automatique, le plus souvent sans validation d'expert
- elle est de nature très variable d'un cas à l'autre
- elle s'appuie sur des bases de données de séquence où l'annotation fonctionnelle a longtemps été le parent pauvre, et où les cas de "propagation d'erreurs" sont légion
- elle ne donne dans la plupart des cas qu'une information limitée au niveau moléculaire ou biochimique (motif de liaison, activité enzymatique...)

Enfin, pour environ un tiers des gènes et protéines, on ne dispose d'aucune piste. Toute prédiction fonctionnelle n'a donc qu'une valeur indicative et limitée tant qu'une preuve expérimentale n'est pas disponible. Cependant, ce type d'annotation, couplé à d'autres informations comme le profil d'expression de l'ARNm ou la localisation du produit protéique, permet d'émettre des hypothèses fonctionnelles et d'envisager une validation/invalidation expérimentale.

Cette étape d'annotation constitue un enjeu de recherche majeur en bioinformatique dans la mesure où il doit garantir la cohérence sémantique de l'information replacée dans son contexte biologique. Le changement quantitatif et qualitatif que connaît la biologie implique aussi la mise à disposition de ces données au travers de banques et de bases de données hétérogènes et distribuées. Les différentes sources de données se réfèrent à des niveaux d'organisation biologique très divers, que l'on se place au niveau du gène, d'un génome, de l'expression, ou encore de l'interaction entre molécules au sein d'une voie métabolique. En modélisant explicitement les relations existantes entre les entités biologiques, il devient possible de replacer ces objets dans leur contexte en y rattachant d'autres catégories d'informations (par exemple, un ensemble de gènes co-exprimés dont les produits sont impliqués dans une voie métabolique donnée et pour lesquels des orthologues sont connus chez un autre organisme). Il faut ensuite mettre en œuvre le cadre conceptuel conduisant à l'établissement de bases de connaissances et des méthodologies permettant d'interroger ces connaissances moléculaires à l'échelle cellulaire.

En structurant ce processus de conversion, la bioinformatique propose des méthodologies innovantes et développe le cadre conceptuel permettant d'aboutir à l'élaboration de modèles intégratifs de raisonnement comme outils de recherche. Elle peut aider les chercheurs à avoir une vision intégrée à l'échelle cellulaire des entités moléculaires en présence et des relations qui les caractérisent afin d'y effectuer des interrogations complexes. A long terme, la bioinformatique pourrait aider à comprendre les relations entre les entités composant un réseau biologique donné et le flot d'informations produits par les processus biologiques sous-jacents pour faire émerger de nouvelles propriétés caractéristiques de ces systèmes. La bioinformatique intervient aujourd'hui dans le cycle de recherche dans sa capacité à proposer de nouvelles hypothèses de travail issues des prédictions, par confrontation des connaissances, ou encore en élaborant des modèles mathématiques rendant compte des propriétés dynamiques des systèmes biologiques.

3.4.2 Génomique comparative

Au delà de la séquence complète du génome d'un individu d'une espèce choisie en général comme exemplaire d'un groupe taxonomique (voir ci-dessous), il est rapidement apparu intéressant d'exploiter et d'étendre les données génomiques dans le cadre de ce qu'on peut appeler la génomique comparative. L'objectif est de réaliser des comparaisons globales de génomes au sein d'une même population ou espèce, entre espèces apparentées, ou même entre groupes taxonomiques (très) éloignés⁶, de manière à en induire des conclusions sur l'évolution des gènes et des génomes, la spéciation, l'adaptation....

Les données de séquençage et les nouvelles technologies de génotypage haut-débit permettent en particulier l'exploration à très grande échelle de la variabilité génotypique au sein par exemple d'une espèce. Couplées à des données phénotypiques, ces informations ont des retombées particulièrement intéressantes en ce qui concerne l'adaptation aux contraintes environnementales et la résistance aux bio-agresseurs, ainsi que des implications directes en sélection. Il faut noter qu'une fois une première séquence génomique complète obtenue, des programmes de re-séquençage de très grande envergure peuvent être envisagés⁷.

En ce qui concerne les plantes, il est regrettable qu'à de rares exceptions près (*Physcomitrella*), des raisons autres que scientifiques conduisent à focaliser les travaux de génomique sur quelques espèces de grande culture d'intérêt économique, quelquefois très apparentées (céréales) alors que des pans entiers de la biodiversité végétale terrestre restent totalement inexplorés. Les biologistes animaux ne s'y sont pas trompés et même si les mammifères restent au centre des préoccupations, la variété d'organismes étudiés est nettement plus grande.

3.5 Ecophysiologie, modélisation et biologie intégrative

3.5.1 Les nouveaux défis de l'Ecophysiologie.

3.5.1.1 L'Ecophysiologie : une discipline de l'Ecologie

Les fondements de l'Ecophysiologie végétale ont été établis dans le cadre de l'Ecologie fonctionnelle, c'est-à-dire sur l'étude des interactions entre les populations et communautés d'espèces végétales et leur milieu et ceci dans les deux sens (i) effets du milieu sur le fonctionnement des organismes et leur dynamique populationnelle, et (ii) effet du fonctionnement de ces populations et communautés d'organismes sur l'évolution des milieux. Appliquée de manière privilégiée aux peuplements végétaux cultivés dans une optique d'optimisation de la production agricole et des facteurs mis en œuvre, l'Ecophysiologie à l'INRA s'est progressivement coupée de sa discipline fondatrice en concentrant essentiellement ses études sur le premier sens des interactions milieu-peuplement. Les défis de l'Environnement et de la Biodiversité auxquels l'Agriculture est actuellement confrontée impliquent un retour vers les origines de l'Ecophysiologie. Pour cela il s'agit d'interfacer les modèles de fonctionnement milieu-peuplement avec les modèles de Biologie des populations

⁶ Un exemple récent et fascinant de génomique comparative concerne deux articles consacrés à la prédiction des protéines impliquées dans la formation des cils et des flagelles : ces articles sont commentés dans Pazour (2004) *Curr Biol* 14 : R575.

⁷ Un programme d'obtention des séquences génomiques complètes d'une vingtaine d'éco-type d'*Arabidopsis* a démarré récemment.

pour pouvoir aborder les grandes problématiques concernant l'évolution des agro-écosystèmes, leur contribution aux cycles biogéochimiques, leur interaction avec les changements globaux et leur rôle dans l'évolution de la biodiversité.

3.5.1.2 L'Ecophysiologie : Physiologie de la plante entière en interaction avec son milieu

Comme nous l'avons évoqué plus haut (chapitre 1), l'Ecophysiologie s'est essentiellement développée à partir d'une approche thermodynamique du fonctionnement des peuplements végétaux permettant de modéliser les échanges de matière et d'énergie dans le continuum sol-végétation-atmosphère. Cette approche très féconde (1970-1990) s'est traduite par de grandes avancées théoriques permettant de simuler la productivité primaire des différents écosystèmes, leur réponses aux modifications du milieu (forçages anthropiques ou climatiques) et leurs contributions à l'évolution des milieux et donc de l'environnement. Cependant, le niveau d'organisation de la plante entière n'étant pas toujours pleinement explicité dans les modèles, hormis par des paramètres de type « résistance », il est apparu difficile à partir de cette seule approche de rendre compte de la diversité biologique interne des peuplements végétaux. D'autre part, le couplage avec les approches populationnelles était impossible. L'Ecophysiologie a donc dû investir le champ de la Physiologie de la plante entière progressivement abandonné par les physiologistes qui ont investi massivement les niveaux cellulaires puis moléculaires. Ce recentrage de l'Ecophysiologie au niveau de l'organisme a permis d'expliciter les interactions entre la plante et son milieu et d'analyser la réponse des différentes fonctions (photosynthèse, respiration, absorption minérale, assimilation de l'azote...) aux différentes contraintes du milieu abiotique.

3.5.1.3 Une nouvelle approche plus intégrative : les modèles structure-fonction

Jusqu'à présent la modélisation du fonctionnement des plantes au sein des peuplements végétaux a été réalisée au moyen de modèles compartimentaux représentant des flux de matière (assimilats) entre les divers compartiments morphologiques, biochimiques ou métaboliques d'une plante. Ces modèles ont permis d'analyser les diverses fonctions physiologiques essentielles de la plante : photosynthèse, respiration, transpiration, absorption minérale, répartition des assimilats, constitution et remobilisation des réserves C et N etc... Cependant la limite de cette approche purement compartimentale réside dans le fait que les facteurs du milieu sont considérés comme homogènes alors que la plante occupe un espace discret au sein duquel il existe une variation importante des facteurs du milieu, soit dans le sol (humidité, teneurs locale en N et minéraux), soit au niveau des parties aériennes (niveaux d'éclairement et température des organes) et que du fait de l'organisation architecturale de la plante (géométrie 3D et topologie) les relations entre ses différentes parties (organes) sont elles mêmes structurées dans l'espace.

Le problème du fonctionnement de la plante individuelle au sein d'un peuplement implique de développer une approche architecturale de la croissance des plantes (dynamique de croissance dans l'espace 3D et topologie des organes). Les plantes au sein d'un peuplement sont en effet en compétition pour l'accès aux ressources du milieu qui ne sont pas réparties de manière uniforme. D'autre part elles échangent des signaux par l'intermédiaire de modifications de caractéristiques physiques ou chimiques du milieu (photomorphogénèse, allélopathie). L'étude de l'ensemble de ces interactions qui modifient profondément le fonctionnement des plantes implique d'avoir une représentation 3D de la croissance des plantes couplée à une représentation discrétisée des paramètres du milieu dans cet espace. Les possibilités de calculs permettent aujourd'hui de développer de tels modèles structure-fonction de manière dynamique permettant de simuler le fonctionnement de plantes virtuelles en interaction avec leur milieu (Tardieu, 2003). Il devient alors possible de représenter le peuplement végétal comme une collection de plantes individuelles qui interagissent entre elles

et d'aborder ainsi de manière fonctionnelle les phénomènes de compétition qui n'ont jusqu'à présent fait l'objet que d'approches descriptives.

Cette approche structure-fonction doit pouvoir s'appliquer également à des niveaux d'organisation plus élémentaire. Ainsi la croissance d'une plante peut être analysée comme la somme coordonnée de l'activité de ses différents méristèmes. Le fonctionnement des méristèmes peut alors faire l'objet d'une modélisation représentant des compartiments structurés dans l'espace au sein desquels peuvent être analysés des flux de matières (C, N H₂O, minéraux) correspondant à des activités de division et de grandissement cellulaire. Cette approche couplée aux possibilités d'imagerie doit permettre d'établir une passerelle avec les approches moléculaires en proposant un cadre d'analyse et d'interprétation des données.

3.5.1.4 De la plante entière à l'organe et à la cellule : signalétique et information

Le principal défi de la Biologie Intégrative qui est posé à l'Ecophysiologie est celui de la représentation de la diffusion de l'information de nature analogique en provenance de l'environnement externe (biotique et abiotique) au sein de l'architecture végétale, de sa transduction, de sa perception locale et de son interaction avec l'information de nature digitale en provenance du génome. L'Ecophysiologie doit investir ce champ d'investigation grâce à ses outils permettant de caractériser et quantifier les signaux environnementaux perçus par les plantes qui sont de nature physico-chimique et d'en étudier les modes de diffusion au sein de la plante. Les modèles « structure-fonction » tels que définis ci-dessus sont une étape indispensable pour fournir un cadre d'analyse. Mais des investigations plus mécanistes sur les flux d'information ou les signaux transportés par les sèves xylémienne et phloémienne, et leur diffusion dans les tissus des organes sont indispensables pour comprendre les régulations fonctionnelles entre les sources d'assimilats et les puits, entre les feuilles et les racines, entre les parties végétatives et les organes reproducteurs. Avant d'intervenir directement dans le métabolisme en interaction avec les informations de nature « génomique » ces signaux sont souvent de nature physico-chimique et il convient alors d'en déterminer la nature exacte ainsi que les flux. Les bases physiques et thermodynamiques qui ont fondé l'Ecophysiologie doivent pouvoir être « recyclées » pour ré-investir le domaine de la Biophysique. Les percées méthodologiques et conceptuelles qui ont été réalisées récemment dans le domaine de la perception par les plantes du micro-climat lumineux ou des contraintes mécaniques montrent qu'il est possible de caractériser et de quantifier ces signaux et d'en étudier les réponses coordonnées de la plante. Les développements des Nanotechnologies (voir supra) devraient permettre de mettre en œuvre de nouvelles méthodes d'investigation *in situ* et une nouvelle approche expérimentale.

3.5.2 La théorie du contrôle métabolique

Construite sur la théorie des systèmes complexes et les propriétés émergentes, la théorie des contrôles métaboliques (TCM) fondée dans les années 70 apporte des concepts utiles à la Biologie intégrative.

L'état d'un système métabolique, à un moment donné ou à l'état stationnaire, peut être décrit par les valeurs de ses variables systémiques, à savoir les flux (vitesse de la réaction dans une branche du réseau) et les concentrations en métabolites. Les valeurs de ces variables, qui ne sont évidemment pas indépendantes, sont déterminées par celles des paramètres du système, dont les plus classiquement considérés sont les constantes cinétiques des enzymes (constantes de catalyse, de Michaelis, et d'inhibition notamment), et leurs concentrations. Il est important de noter que ces paramètres sont génétiquement déterminés. Si l'on admet qu'*in fine*, tous les caractères phénotypiques, et donc la valeur sélective d'un individu, sont contrôlés par le métabolisme, on voit qu'une théorie biochimique capable de relier la variation des paramètres

du système à la réponse des flux ou des pools de métabolites sera centrale non seulement en biotechnologies, mais aussi en génétique, évolution ou amélioration des espèces.

C'est précisément ce qu'apporte la TCM, en dégageant un certain nombre de propriétés générales non triviales des systèmes, qui permet de répondre à des questions fondamentales que les approches purement réductionnistes sont incapables d'aborder. Pourquoi est-il beaucoup plus facile de diminuer la concentration d'un métabolite que de l'augmenter ? Pourquoi la plupart des mutations défavorables sont-elles récessives ? Pourquoi rencontre-t-on si souvent des interactions épistatiques ? Pourquoi la plupart des polymorphismes sont-ils phénotypiquement et sélectivement neutres ?

Cette démarche exemplaire de biologie intégrative est réalisée par l'équipe *Génétique Quantitative Fondamentale* de l'UMR de Génétique Végétale du Moulon (INRA-UPS-CNRS-INA PG). Elle s'appuie sur la TCM pour modéliser la variation quantitative et son évolution. Ses contributions concernent la distribution « en L » des effets des QTL (qui s'est révélée être une conséquence de la propriété de sommation des coefficients de contrôle), une généralisation de la TCM qui lève l'hypothèse d'indépendance entre les concentrations d'enzymes (ce qui permet d'intégrer les co-régulations entre concentrations d'enzymes dans les modèles de variation quantitative), et une généralisation du modèle de dominance, qui a débouché sur un modèle d'hétérosis métabolique validé par la reconstruction *in vitro* d'une chaîne métabolique. Les recherches actuelles concernent l'évolution expérimentale des régulations d'un réseau, en utilisant la glycolyse de levure comme système modèle.

[Annexe 7 - La théorie du contrôle métabolique : une démarche exemplaire de biologie intégrative, D. de VIENNE]

4 Organisation/Partenariats/Formation

La consultation des chercheurs INRA et l'analyse de l'organisation de la recherche dans les laboratoires conduisent le groupe de réflexion sur la biologie intégrative végétale (BIV) à :

- Identifier des lignes directrices pour développer les approches biologie intégrative et systèmes biologiques complexes ;
- Recommander une panoplie de moyens et de leviers favorisant la réalisation des projets BIV ;
- Analyser les niveaux opérationnels d'organisation.

4.1 Organisation fonctionnelle

Conditions à satisfaire :

Les stratégies de recherche de la biologie intégrative végétale, dont la physiologie moléculaire et la génétique végétale modernes sont certes des composantes majeures, doivent également associer des approches conceptuelles et méthodologiques relevant de la génomique, de la biologie cellulaire, de la biologie du développement, mais aussi de l'écophysiologie, de l'écologie, de la modélisation, de la pathologie, de l'agronomie...

Force est de constater que les projets intégrés naissent souvent de montages informels et limités dans le temps, par exemple à propos d'un projet de thèse à l'interface de 2 disciplines, mais aussi de projets élaborés par un groupe de travail ouvert à différentes approches disciplinaires, ou encore à l'occasion d'un financement incitatif sur des approches complémentaires.

Une étape initiale d'échange de points de vue, de confrontation d'idées est primordiale à la co-construction d'une plate-forme de recherche avec des contributions disciplinaires équilibrées. Elle mobilisera donc les multiples compétences de haut niveau présentes au sein de l'INRA et au delà, qui participeront à la définition des objectifs généraux de recherche. L'appropriation du projet de recherche sera facilitée par la réalisation d'une expérimentation menée en commun, sur un même modèle biologique et par l'exploitation commune des données issues des différentes approches.

Moyens :

Pour encourager ces synergies, l'organisation de la recherche doit susciter le regroupement de compétences et de disciplines diverses à même de faire émerger des projets de recherche multidisciplinaires originaux dans le cadre du programme transversal Biologie Intégrative affiché dans la stratégie de l'INRA.

Au-delà des capacités propres de l'Institut, l'association d'équipes d'autres organismes de recherches est à rechercher pour leur complémentarité et leur expertise. Dans tous les cas (intra ou inter organismes), la construction de partenariats structurants et durables est l'objectif ultime souhaité.

Le devenir de tels projets multidisciplinaires est très dépendant de la dynamique impulsée par le responsable de projet, leader scientifique reconnu, assurant une animation dynamique et une cohérence forte du groupe. L'engagement des équipes dans des projets BIV considérés « à risque » va largement dépendre des facteurs suivants :

- L'attractivité et la durée des moyens mobilisables sur la thématique Biologie intégrative (appels d'offre national, européen ou international), ainsi que les moyens incitatifs et d'accompagnement pour les équipes (attribution de bourses de thèse ou de post-doctorants ; équipements...).
- La valorisation des résultats de leur recherche à moyen terme (publications, brevets) dans des revues internationales bien cotées.
- La prise en compte par les évaluations individuelle et collective des mobilités thématiques (voire géographique) des scientifiques et leur investissement dans l'organisation, l'animation et la gestion de projets intégrés si l'on veut susciter l'adhésion des jeunes chercheurs à un projet pluridisciplinaire.

4.2 Structuration des groupes de recherche

Conditions à remplir :

Le développement de la Biologie intégrative végétale devrait conduire au rapprochement d'équipes abordant la physiologie végétale, la génétique, la pathologie, l'écophysiologie et l'écologie ainsi que les mathématiques et l'informatique appliquées. Nous en avons donné quelques exemples en annexe. Nombreux questionnements de la hiérarchie de l'INRA et des chercheurs rencontrés ont fait l'objet des débats du groupe BIV. Son avis a été requis sur : la recomposition des départements du secteur végétal ou la redéfinition de leurs contours ; le regroupement de compétences confirmées et de nouvelles compétences dans une nouvelle entité multidisciplinaire ; la création de laboratoires pluridisciplinaires inter organismes ?

Quel que soit le changement de structuration envisagé, le groupe de réflexion a évalué son impact vis-à-vis de la multidisciplinarité, de la production scientifique et de la gestion des équipes de recherches engagées dans un projet de biologie intégrative.

Le concept d'équipe de recherche, unité élémentaire en charge d'un programme est largement accepté. Il a été formalisé par la mise en place des évaluations collectives d'unités et la contractualisation quadriennale, ainsi que par la reconnaissance comme équipe d'accueil de doctorants, à la condition d'être animée par un scientifique titulaire d'une Habilitation à Diriger des Recherches. Dans ce cadre, il est donc possible d'organiser une équipe de recherche pluridisciplinaire contractualisée autour d'un thème de biologie intégrative.

De même, les laboratoires et unités de recherche de l'INRA ont largement participé à la réorganisation de la recherche péri-universitaire, à la contractualisation et à la constitution d'unités mixtes de recherche. Dans ce cadre, des compétences pluridisciplinaires issues des différents départements de recherche du même organisme ou de différents organismes de recherche peuvent s'associer de façon contractuelle sur une thématique BIV.

L'organisation de telles équipes et unités mixtes équilibrées faciliterait la cohabitation et la communication ainsi que la gestion des moyens et des équipes regroupées dans un même laboratoire, autour d'outils communs. Mais le fonctionnement d'équipes et de partenaires en réseau garde aussi tout son intérêt quand la dispersion géographique des compétences et des moyens reste le facteur limitant.

La recomposition des départements du secteur végétal n'est pas apparue comme un levier déterminant pour développer la biologie intégrative végétale à l'INRA. En effet, les départements d'essence thématique jouent un rôle essentiel de référence et d'animation disciplinaire pour des chercheurs et des équipes dispersées. C'est dans cet esprit que le CNRS

vient de créer une nouvelle section 28 dénommée Biologie végétale intégrative ; il s'agit essentiellement d'afficher l'importance accordée à l'approche intégrée en physiologie végétale, dans les domaines utilisant des approches globales à haut débit, et en biologie végétale à différentes échelles et niveaux d'étude.

Par contre, les départements INRA peuvent être perçus comme des structures gestionnaires limitant le développement de la multidisciplinarité et entraînant par nature des « rigidités » structurelles à la réalisation des projets intégrés inter départements et inter organismes. Les différents mécanismes de concertation et les outils de gestion élaborés pour la gouvernance des structures INRA permettent-ils de réduire ces frontières ? Le groupe de réflexion BIV recommande :

- d'ajuster si nécessaire le contour des départements du domaine végétal quand la position d'une thématique particulière est reconnue hors champ à l'évaluation du département ;
- de coordonner les projets de BIV au plus haut niveau (*top-down*), le directeur scientifique Plante et produits du végétal, en termes de moyens incitatifs, de ressources humaines, d'équipes de recherche, d'infrastructures et équipements de recherche... ;
- dans le même temps de favoriser la montée (*bottom-up*) des projets de jeunes chercheurs décidés à se constituer en équipe pluridisciplinaire dont on favorisera l'émergence après évaluation de la pertinence ;
- de renforcer significativement l'appui des services au montage et à la gestion des projets de recherche multidisciplinaire (surtout dans le cadre UE et international).

Moyens structurants :

1. L'accès à des plates-formes spécialisées :

Le développement d'approches intégrées en biologie, en particulier dans le domaine végétal, s'appuie sur le réseau des Génopoles qui représente un outil essentiel pour mettre en place ce type de stratégie à l'échelle du territoire national. En effet, la volonté de renforcer la recherche génomique a conduit le ministère de la Recherche à créer dans les régions, avec l'aide des collectivités territoriales, des structures comparables à celle de la génopole d'Evry. Les génopoles sont structurées en réseau et sont implantées à Lille, Lyon-Grenoble, Marseille-Nice, Montpellier, Rennes-Nantes, Strasbourg, Toulouse et en région parisienne. Elles ont développé un ensemble de plates-formes technologiques et de ressources en génomique disponibles pour des projets scientifiques dans les domaines végétaux, animaux ou microbiens proposés par la communauté nationale. Ces outils pour la biologie à grande échelle vont du séquençage et du génotypage, à l'analyse structurale et fonctionnelle des gènes et de leurs produits. Des plates-formes de bioinformatique, d'analyse du transcriptome et du protéome, et des centres d'imagerie de très haute technicité ont ainsi été mis en place dans de nombreuses régions, en cohérence avec les plates-formes RIO.

2. La spécialisation des dispositifs expérimentaux

L'approche génétique implique le développement de plateaux technologiques pour le phénotypage à haut débit. Mais au-delà de cette approche en conditions contrôlées, il est indispensable d'aborder l'analyse des interactions GxE en conditions agronomiques, c'est-à-dire « au champ ». Il importe alors de se donner les moyens de caractériser « au champ » les conditions environnementales biotiques et abiotiques réellement perçues par les peuplements végétaux étudiés en instrumentalisant un petit nombre de sites expérimentaux au sein desquels

on pourra faire varier expérimentalement une ou quelques composantes de l'environnement. Il s'agit alors de remplacer la variation environnementale « aveugle ou aléatoire » classiquement obtenue par un certain nombre de « sites-années », par une variation environnementale contrôlée et mesurée permettant d'approcher la nature des interactions mises en jeu grâce à une approche de modélisation.

Des plateaux techniques et des moyens expérimentaux appropriés sont donc indispensables pour des phénotypages et des évaluations de plantes en conditions plus ou moins contrôlées (serres, chambres de culture, essais en champ...). L'adaptation des unités et installations expérimentales à ces nouveaux besoins et leur requalification pour répondre aux exigences des approches de post-génomique imposent une réflexion approfondie et des choix drastiques. Elles doivent pouvoir fournir un cadre expérimental unique et exceptionnel, au prix d'un effort de structuration, d'équipement et de formation au niveau national permettant de constituer un véritable réseau.

3. Centres de ressources biologiques (dossier INRA/BRG)

L'INRA assure la gestion de collections végétales en rapport avec les programmes d'amélioration de nombreuses espèces de plantes cultivées, ainsi que de souches pathogènes de ces cultures. Dans le cadre de la Charte nationale du BRG, la gestion de ces ressources génétiques a été confiée par espèce ou groupe d'espèces apparentées à 11 unités de recherche impliquées. La perspective de leur promotion au statut de CRB contribuerait à leur reconnaissance internationale, sur la base d'un cahier des charges précis pour la conservation, la multiplication, la caractérisation et la constitution de bases de données. Outre les plantes cultivées, le DGAP a créé 2 centres de ressources pour les 2 espèces modèles *A. thaliana* et *M. truncatula*, et du matériel biologique expérimental diversifié (RILs, HD, mutants, populations...) indispensable aux études de génomique et de génétique.

Dans le même temps, les approches génomiques à haut débit ont généré des ressources génomiques (banques BAC ...) et des banques de données variées de séquences (EST), de cartographie (cartes, marqueurs, QTL), d'expression (puces à ADN, protéome...). L'ensemble des ressources génomiques INRA sont conservées par le CNRGV de Toulouse.

L'organisation mise en place au niveau national est remarquable et opérationnelle. Elle doit répondre aux besoins en matériel biologique et en informations nécessaires aux travaux de biologie intégrative.

4. L'organisation de pôles de recherche et de campus universitaires :

Elle répond à l'objectif à moyen terme de structuration de l'espace de recherche européen. L'enjeu est la construction d'un nombre très limité de pôles européens sur des grandes thématiques -la biologie intégrative végétale par exemple, avec des projets de recherche mis en commun. Leur réalisation s'appuie d'une part sur le développement de plates-formes technologiques ouvrant un accès partagé à des outils coûteux, d'autre part sur une construction de pôles associant différents départements (BV/EA/EFPA/GAP/SPE et MIA pour l'INRA) et différents organismes (CIRAD, CNRS, INRA, IRD...)

4.3 Compétences et Formation

Fondée sur des approches pluridisciplinaires, la recherche en BIV pose d'emblée les questions suivantes : quelles sont les différentes spécialités disciplinaires requises? Comment

faciliter la compréhension et la communication entre spécialistes de disciplines différentes ? Dans quels domaines faut-il développer la formation des scientifiques ?

Le groupe de réflexion propose quelques pistes pour répondre à ces questions :

1- La spécialisation est le 1^{er} sujet de débat : est-il préférable d'associer des chercheurs de différentes spécialités scientifiques ou d'identifier des profils à double compétence ? Il nous paraît souhaitable de réunir les meilleurs spécialistes dans leur domaine de recherche pourvu qu'ils répondent aux 2 exigences suivantes.

2- Quel que soit leur formation, il nous paraît indispensable de favoriser le débat scientifique et de développer les capacités de compréhension et de collaboration sur les bases suivantes:

- contrebalancer l'hyperspécialisation et s'ouvrir à d'autres disciplines ;
- s'appropriier les thématiques et les langages d'autres domaines scientifiques ;
- développer les approches de la complexité et de la modélisation des systèmes.

3- Quelles formations supérieures développer en Biologie pour former les futures générations de chercheurs répondant aux critères et compétences souhaitables en BIV ? Une forte remise en cause des spécialisations sur le modèle DEA/doctorat des années 90 est indispensable dans le cadre de la réforme en cours des Masters et formations doctorales. Des formations complémentaires aux interfaces de la biologie, des mathématiques, des géosciences, de la physique (thermodynamique) et de la biochimie seront accessibles grâce aux modules optionnels des formations M et D.

Globalement, il est souhaitable de s'appuyer sur les meilleures expertises disciplinaires possibles, à condition de les faire évoluer dans 3 voies complémentaires :

- i) depuis 20 ans, les formations doctorales en Biologie se sont souvent hyperspécialisées en biologie et génétique moléculaires au détriment de la biochimie, de la physiologie et de l'écophysiologie. L'enseignement de Biologie végétale est souvent resté descriptif et analytique sur la base d'ouvrages en partie obsolètes (saluer l'initiative JF Morot Gaudry/ JF Briat). Il est indispensable de repenser la formation disciplinaire des chercheurs pour leur donner les concepts et outils leur donnant une vision intégrative et une culture plus globale. La formation complémentaire donnée aux doctorants en offre une possibilité timide. La réforme Master est une première ouverture possible pour construire des parcours individuels diversifiés.
- ii) les biologistes avec une culture mathématique et informatique (formation ingénieur en particulier) sont très recherchés en BIV pour leur capacité à créer et gérer les bases de données, analyser les données de génomique, modéliser des processus et pour développer une recherche en collaboration avec des spécialistes en mathématique et informatique appliquées.
- iii) la modélisation du fonctionnement des entités biologiques étudiées (cellule, organe, plante, peuplement...) est le point de passage obligé de toute tentative d'approche intégrative en biologie. De plus la modélisation est par définition un « médiateur » interdisciplinaire essentiel. La formation de modélisateurs « biologistes » doit donc être plus présente au sein de « masters » fusionnant des compétences en mathématiques appliquées et en biologie.

Nous tenons à alerter nos tutelles sur le désintérêt actuel des ingénieurs agronomes spécialisés en Biologie possédant une culture scientifique large recherchée dans les recherches de BIV, et

d'une façon plus générale des étudiants universitaires, résultant du peu d'attractivité pour des études longues peu rémunérées et de l'altération de l'image de la recherche en biologie végétale.

Pour mobiliser et valoriser le potentiel « chercheur » de l'INRA de disciplines variées dans les projets du BIV, il faut aussi soutenir les différentes initiatives et opportunités de formation continue

- écoles-chercheurs, séminaires- en adéquation avec les principes généraux rappelés précédemment.

Enfin, on ne peut ignorer le rôle majeur joué par les ingénieurs et les personnels techniques des équipes et unités de recherche engagées dans la BIV. Pour les plus jeunes recrutés, ce sont leurs compétences en génomique sur des technologies de pointe qui sont recherchées, et qu'il leur faudra régulièrement actualiser. Pour d'autres, leur expertise et leur qualification font partie des acquis qui peuvent être valorisés ou remis en cause pour s'adapter à l'évolution des démarches expérimentales et des outils. Des efforts importants devront être faits pour la formation de ces agents et leur offrir des possibilités d'évolutions attractives d'activité.

Perspectives

1. La Biologie intégrative s'impose comme un nouveau paradigme dont l'objectif est de « donner du sens » aux approches analytiques de la génomique ou d'intégrer les informations acquises à différents niveaux d'approche pour donner une explication et une signification aux processus étudiés. La démarche Biologie intégrative est encore jeune, souvent exploratoire et dans une phase de construction. L'effort de recherche en génomique sur plantes modèles reste prioritaire dans le cadre européen et international pour la décennie en cours, avec une priorité pour l'INRA au transfert des connaissances et ressources génomiques aux plantes d'intérêt agronomique. A ce stade, il paraît souhaitable

- i) de prolonger le projet Génome par l'affichage d'un grand programme transversal de BIV placé sous la tutelle du directeur scientifique PPV, ouvert à l'interdisciplinarité et aux autres organismes de recherche
- ii) de favoriser les débats dans les équipes de recherche, la confrontation des idées, l'apprentissage des échanges, des travaux et des projets interdisciplinaires en Biologie intégrative
- iii) de mettre en œuvre les moyens « facilitateurs » et motivants recommandés par le groupe de réflexion en terme d'organisation et de compétence (voir p.33-38).

Les projets à développer en BIV construits par des équipes des départements BV, CEPIA, EA, EFPA, GAP, MIA et SPE se rattachent essentiellement aux thèmes suivants :

- la génomique fonctionnelle de l'ensemble des gènes des espèces modèles ;
- le développement des plantes et leur reproduction ;
- les déterminants et les modèles de la qualité des produits végétaux ;
- les mécanismes et les gènes en jeu dans les interactions plantes/microorganismes pathogènes ou bioagresseurs de plantes;
- les mécanismes et les fonctions physiologiques assurant l'adaptation et la tolérance des plantes aux contraintes abiotiques... ;
- la reconnaissance des réseaux de gènes responsables et une identification des éléments clefs déterminants de ces adaptations ;
- les signaux en provenance du milieu abiotique et biotique déclencheurs de la cascade des réponses moléculaires adaptatives.

2. Le temps est aussi venu de « spéculer » sur le futur de la Biologie intégrative après une phase analytique intense fondée sur le développement de la génomique par les pôles de recherche qui ont investi dans les plates-formes technologiques permettant la production massive de données, et leur intégration aux différents niveaux hiérarchiques considérés. Nous ne pourrions nous satisfaire de l'approche des grandes fonctions physiologiques et des interactions biologiques et environnementales sur la base de données expérimentales partielles, acquises par sondage/observation externes et construites sur des corrélations. Cette démarche s'apparente selon J. Samarut à celle des « archéologues du système reconstruisant le passé à partir de carottages/fouilles ». On espère au contraire déboucher sur de nouveaux modes de raisonnement et découvrir des voies épistémologiques originales.

De nouvelles méthodes sont à rechercher et de nouveaux concepts sont à privilégier en Biologie intégrative, en particulier la modélisation des systèmes biologiques complexes (voir rapport Modélisation, B. Goffinet) et l'expérimentation directe, en temps réel, au sein des systèmes biologiques. Nous tenterons de les illustrer par quelques pistes originales, parfois spéculatives.

Grâce à l'important investissement réalisé en bioinformatique, l'analyse des données « en masse » produites par la génomique est maintenant prise en charge efficacement. Par contre, il est nécessaire d'élaborer des méthodologies fondées sur l'intelligence artificielle et les

statistiques pour générer de nouvelles hypothèses sur les données de génomique, à soumettre à des validations.

3. L'étude des objets biologiques vivants en milieu réel sous l'influence de leurs environnements abiotique et biotique, dans un suivi dynamique et/ou global, est réalisable (ou le deviendra) grâce à des évolutions technologiques comme :

- Le suivi des métabolismes *in vivo* par les isotopes naturels C^{13}/N^{15} : Le fractionnement ou la discrimination isotopique des molécules lors des processus biologiques, comme l'assimilation du carbone à la lumière, aboutit à une différence infime de l'abondance naturelle des isotopes entre le substrat et le produit, mesurable à l'aide du spectromètre de masse isotopique. La signature isotopique naturelle de la matière organique et des métabolites intègre tout le passé métabolique de l'objet biologique étudié, ainsi que l'effet de son environnement.

- Le recours aux nanobiotechnologies (voir p 24-25): L'échelle nanométrique est celle des composés biochimiques et des ultrastructures fonctionnelles de la cellule vivante. L'émergence des nanosciences en physique, en chimie et en médecine a fait l'objet en 2003/2004 de rapports d'évaluation de l'Assemblée nationale et de l'Académie des sciences (JL Lorrain et D Raoul, 2004). L'apport de ces approches à la biologie n'est que partiellement abordé (voir JC Mounolou, NSS à paraître) et concerne :

- l'élaboration d'objets inédits, complexes, à fonctions nouvelles résultant des interactions moléculaires et utilisables dans l'exploration fonctionnelle intracellulaire et l'analyse *in vivo* (capteurs moléculaires, navettes émettant des images électroniques, puces à ADN et protéines, imagerie avec des molécules marquées..) ;
- le nanomagnétisme exploitant les propriétés magnétiques de certains composés moléculaires dans la compréhension de leurs fonctions ;
- la technologie des verres nanostructurés en microscopie réalisant le codage et le transfert rapide d'information.

4. Dans le développement de la Biologie intégrative, la génétique joue un rôle charnière pour déterminer les fonctions des gènes et l'expression du génotype dans son environnement (phénotype). Les mutants monogéniques des espèces modèles sont exploités intensivement dans ce but. Une large marge de progression est attendue dans plusieurs domaines comme:

- la caractérisation et la dissection des zones génomiques impliquées dans les caractères quantitatifs d'intérêt agronomique et d'adaptation afin d'identifier les gènes en jeu, leurs relations et leurs effets;
- la génétique d'association entre des polymorphismes moléculaires et les caractères ;
- la conception de génotypes virtuels par une modélisation des combinaisons d'allèles favorables et leur création par des croisements contrôlés en vue de leur évaluation ;
- le contrôle du niveau d'appariement et de recombinaison homéologue dans le cadre de l'exploitation et de la maîtrise de l'introgession des gènes d'espèces sauvages au sein des complexes d'espèces polyploïdes nombreux chez les plantes cultivées.

De nouvelles découvertes peuvent aussi réorienter les recherches dans ce domaine, comme la régulation épigénétique des gènes. Elle résulterait de l'invasion des zones codantes et non codantes des génomes par des éléments transposables ou des acides nucléiques étrangers (virus, transgènes..). Son lien avec la méthylation des séquences et des histones a été démontré. L'importance des phénomènes de régulation dans l'expression des gènes conduit à approfondir cette orientation de recherche.

Mais le principal verrou réside dans la non-maîtrise de la recombinaison homologue ciblée chez les plantes. Actuellement, elle est réalisée chez les virus, les bactéries et des cellules de

souris, alors que la transgénése végétale produit aléatoirement et à très faible fréquence quelques rares recombinants homologues. L'amélioration de ce système expérimental est indispensable pour la BIV. Deux voies sont à l'étude : le système RNAi (mutants d'inactivation par interférence ARN) sur riz ; la compréhension du système de recombinaison homologue de la mousse *Physcomitrella*.

5. Un autre volet Biologie intégrative en plein développement concerne l'exploitation des données de séquençage dans les approches comparative et évolutive :

- phylogénie moléculaire de groupes d'organismes ;
- évolution moléculaire de certains organes (fleur des Angiospermes) ;
- comparaison sur des temps longs des stratégies d'adaptation des végétaux aux changements climatiques et des mécanismes moléculaires dans les interactions biologiques (symbiose, pathogène) ;
- diversité allélique en relation avec la biodiversité.

L'avenir de la diversité biologique exploitée (les ressources génétiques) paraît incertain et très dépendant des changements globaux et des activités humaines. Comme nous venons de le voir, nos connaissances sur l'origine, la construction et les fonctions des gènes conduisent à s'interroger sur les possibilités d'enrichir les ressources génétiques par la construction raisonnée de gènes. Ce nouveau champ de recherche repose sur une approche dite « déductive », fondée sur les connaissances les plus récentes de la génomique et sur les processus d'évolution de la structure des gènes (principalement bactériens mais ce processus semble de portée générale). Certaines constructions fonctionnelles pourraient être sélectionnées pour leur adaptation et leur intérêt. Une telle perspective doit être encadrée sur les plans éthique et juridique.

6. Ecophysiologie/modélisation (voir p.29-32)

Deux approches complémentaires permettent d'aborder l'interaction GxE : l'évaluation phénotypique dans des milieux très bien caractérisés et instrumentalisés de génotypes originaux construits *a priori* ; la caractérisation de génotypes réels dans des milieux diversifiés +/- bien caractérisés. L'exploitation *a posteriori* de ces données de caractérisation s'appuie sur la modélisation statistique de l'interaction GxE et sur la modélisation par les ecophysiologistes du fonctionnement des peuplements végétaux en réponse aux facteurs du milieu.

L'analyse et la représentation des données spatialisées sont d'utilisation courante. Par contre, l'exploitation des informations issues de l'étude de la dynamique des processus que nous préconisons requiert des compétences rares à l'INRA. D'une façon générale, quelques équipes devraient se doter des spécialistes nécessaires pour développer une approche systèmes dynamiques complexes.

BIBLIOGRAPHIE

1. Ouvrages/Articles/Rapports d'Intérêt Général

ACADEMIE DES SCIENCES, 2000 : Le monde végétal, du génome à la plante entière. Rapport sur la science et la technologie n° 10. Tech et doc. 220 p.

CNRS, 2002 : Contribution de la section 27 du Comité national au rapport de conjoncture et prospective dans le domaine des Sciences de la vie.

De VIENNE, 2002 : Molecular markers in plant genetics and biotechnology. Science Publishers, Inc.

E. FOX KELLER, 2004 : Génomique, postgénome : Quel avenir pour la biologie ?
La Recherche, 376, 30-37

E. FOX KELLER, 2003 : Le siècle du gène. Gallimard.

J.L. LORRAIN et D. RAOUL, 2004. Nanosciences et progrès médical. Rapport de l'Office parlementaire d'évaluation des choix scientifiques et technologiques. N° 1588. Editions du Sénat. 308p

J-F. MOROT-GAUDRY et J-F. BRIAT, Coord. 2004. La génomique en biologie végétale, 582 p. INRA Editions.

PLANT PHYSIOLOGY, 6th special issue devoted to Arabidopsis (June 2005): Genomics and the virtual plant : new approaches to integrating Arabidopsis genome and function

2. Systems Biology

F. CAPRA, 1996. The web of life: a new scientific understanding of living systems. Anchor Books, Double day (NY). 368 p.

H. KITANO, 2002. Systems biology: a brief overview. Science. 295, 1662.

P.A. MIQUEL. Qu'est-ce-que l'émergence en biologie ? Feux croisés. Vivant editions.

S.A. NEWMAN, 2003. The fall and rise of systems biology. Gene watch. 8-12.

3. Biologie Intégrative

D. DELNERI, F.L. BRANCIA, S.G. OLIVER, 2001. Towards a truly integrative biology through the functional genomics of yeast. **Current opinion in Biotechnology** 12, 87-91.

J. ENGELBERG, 1995. Integrative study in physiology and medicine : obstacles on the road to integration. **Integrative physiological and behavioral science**, 30 (4), 265-272.

- JB. PASSIOURA, 1979. Accountability, philosophy and plant physiology. **Search**. 10 (10), 347-350
- J.S. PEARSE, 2003: The promise of integrative biology : resurrection of the naturalist. **Integr. comp.biol.** 43, 276-277.
- M. PIGLIUCCI, 2003. From molecules to phenotypes ? The promise and limits of integrative biology. **Basic Appl. Ecol.** 4, 297-306.
- J. REES, 2001. Post-genome integrative biology : so that's what they call clinical science. **Clinical medecine**, 1(5), 393-400.
- M. Van BAALEN, 2003. Integrating across levels : what is the relevant level of organisation? Virulence evolution suggests it may not always be so simple. **Biology International** n° 44, 62-68.
- M.H. WAKE, 1998. Integrative biology in biodiversité : an approach to questions, answers and training. *In* **Frontiers in Biology** CH. CHOU and KT SHAO ed. Academia sinica Taïpei. 35-40.
- M.H. WAKE, 2003. What is Integrative biology ? **Integr. comp. biol.** 43, 239-241
- 4. Biologie théorique (modélisation)**
- C. AUFRAY, S. IMBEAUD, M. ROUX-ROUQUIE, L. HOOD, 2003. From functional genomics to systems biology : concepts and practices. **CR Biologies**. 326, 879-892.
- P. AUGER, C. LETT, 2003. Integrative biology : linking levels of organisation. **C.R. Biologies**. 326, 517-522.
- G.L. HAMMER, MS KROPFF, TR SINCLAIR, JR. PORTER, 2002. Future contributions of crop modelling – from heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement. **Europ. J. Agronomy**, 18, 15-31.
- J.D. MURRAY, 2000. Pattern formation in integrative biology – a marriage of theory and experiment. **C.R. Acad. Sci. Paris, Sciences de la vie**, 323, 5-14.
- R. PATON, 2002. Process, structure and context in relation to integrative biology. **Biosystems**. 64, 63-72.
- A. PAVE et C. SCHMIDT-LAINE, 2004. Integrative biology : modelling and simulation of the complexity of natural systems. **Biology international**. 44, 13-24.
- F. TARDIEU, 2003. Virtual plants : modelling as a tool for the genomics of tolerance to water deficit. **Trends in plantscience**. 8 (1), 9-14.