



HAL
open science

MAIAGE - Mathématiques et informatique appliquée du génome à l'environnement

Rapport Hcéres

► **To cite this version:**

Rapport d'évaluation d'une entité de recherche. MAIAGE - Mathématiques et informatique appliquée du génome à l'environnement. 2014, Institut national de la recherche agronomique - INRA. hceres-02032981

HAL Id: hceres-02032981

<https://hal-hceres.archives-ouvertes.fr/hceres-02032981v1>

Submitted on 20 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Évaluation de l'AERES sur l'unité :

Laboratoires Mathématiques et Informatique
appliquées - Jouy en Josas

et

Mathématique, Informatique et Génome

MIAj - MIG

sous tutelle des

établissements et organismes :

Institut National de la Recherche Agronomique - INRA



Décembre 2013



agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

*Pour l'AERES, en vertu du décret du 3
novembre 2006¹,*

- M. Didier HOUSSIN, président
- M. Pierre GLAUDES, directeur de la section
des unités de recherche

Au nom du comité d'experts,

- M. Daniel COMMENGES, président du
comité

¹ Le président de l'AERES « signe [...], les rapports d'évaluation, [...] contresignés pour chaque section par le directeur concerné » (Article 9, alinea 3 du décret n°2006-1334 du 3 novembre 2006, modifié).



Rapport d'évaluation

Ce rapport est le résultat de l'évaluation du comité d'experts dont la composition est précisée ci-dessous.

Les appréciations qu'il contient sont l'expression de la délibération indépendante et collégiale de ce comité.

| | |
|---|---|
| Nom de l'unité : | Mathématiques et Informatique Appliquées - Jouy en Josas et Mathématique, Informatique et Génome |
| Acronyme de l'unité : | MIAj - MIG |
| Label demandé : | UR |
| N° actuel : | 0341 et 1077 |
| Nom du directeur (2013-2014) : | M. Hervé MONOD - M ^{me} Sophie SCHBATH |
| Nom du porteur de projet (2015-2019) : | M. Hervé MONOD - M ^{me} Sophie SCHBATH |

Membres du comité d'experts

Président : M. Daniel COMMENGES, Inserm, Bordeaux

Experts : M^{me} Anne-Laure BOULESTEIX, Ludwig-Maximilians-Universität
Munich, Allemagne
M. Gilles CELEUX (représentant CSS INRA)
M. Hidde DE JONG, Inria, Grenoble - Rhône-Alpes
M. Jean-Philippe VERT, École des Mines de Paris

Délégué scientifique représentant de l'AERES :

M. François COQUET

Représentant(s) des établissements et organismes tutelles de l'unité :

M. Frédérick GARCIA, MIA INRA
M. Cyril KAO (directeur de l'École Doctorale n°435 ABIES)
M^{me} Emmanuelle MAGUIN, MICA INRA
M^{me} Muriel MAMBRINI, Centre Inra de Jouy-en-Josas
M. Pascal MASSART (directeur adjoint de l'École Doctorale n°142)



1 • Introduction

Historique et localisation géographique de l'unité :

MIAj et MIG résultent de la restructuration des deux unités INRA de biométrie de Jouy-en-Josas et Versailles en 2000. Elles sont appelées à fusionner dans le cadre du prochain contrat, et ont demandé une évaluation conjointe dans ce cadre.

Les deux unités sont localisées sur le centre INRA de Jouy-en-Josas, 78352 Jouy-en-Josas cedex, où sera située également la future unité réunifiée.

Équipe de direction

L'équipe de direction est composée de M. Hervé MONOD et M^{me} Sophie SCHBATH.

Nomenclature AERES

ST1 Mathématiques.

Effectifs de l'unité

| Effectifs de l'unité | Nombre au 30/06/2013 | Nombre au 01/01/2015 |
|--|----------------------|----------------------|
| N1 : Enseignants-chercheurs titulaires et assimilés | 2 | 1 |
| N2 : Chercheurs des EPST ou EPIC titulaires et assimilés | 20 | 22 |
| N3 : Autres personnels titulaires (n'ayant pas d'obligation de recherche) | 24 | 23 |
| N4 : Autres enseignants-chercheurs (PREM, ECC, etc.) | | |
| N5 : Autres chercheurs des EPST ou EPIC (DREM, Post-doctorants, visiteurs etc.) | 4 | 3 |
| N6 : Autres personnels contractuels (n'ayant pas d'obligation de recherche) | 2 | |
| TOTAL N1 à N6 | 52 | 49 |



| Effectifs de l'unité | Nombre au 30/06/2013 | Nombre au 01/01/2015 |
|---|----------------------|----------------------|
| Doctorants | 23 | |
| Thèses soutenues | 21 | |
| Post-doctorants ayant passé au moins 12 mois dans l'unité | | |
| Nombre d'HDR soutenues | | |
| Personnes habilitées à diriger des recherches ou assimilées | 10 | 11 |



2 • Appréciation sur l'unité

Avis global sur l'unité

La future unité résulte de la fusion de deux unités MIAj et MIG, appelées ici respectivement équipe 1 et équipe 2. Elle représente une force de recherche importante dans le domaine des mathématiques, statistiques et informatique et leurs applications en biologie. L'enjeu majeur est l'organisation de cette fusion, à la fois sur le plan scientifique et logistique.

Points forts et possibilités liées au contexte

Les points forts sont : la quantité (plus de 200 articles dans la période) et la qualité des publications (plusieurs articles dans Science et dans de très bonnes revues théoriques et appliquées) ; la vie d'équipe qui fonctionne bien avec de nombreux projets, ainsi que de nombreuses activités de recherche et de formation ; la force de la structure informatique avec le projet de « data center » ; la complémentarité des compétences entre les deux équipes fondatrices de la nouvelle unité.

Points faibles et risques liés au contexte

Il y a un risque lié à une certaine disparité des thèmes de recherche, entre recherches très théoriques et très appliquées. L'augmentation des collaborations entre les équipes risque, de plus d'être freinée par la séparation des locaux qu'elles occupent.

Recommandations

Il est recommandé de renforcer les recherches sur la méthodologie statistique pour consolider le pont entre recherche théorique et applications. Un développement des interactions entre les statisticiens de l'unité et les chercheurs ayant des thèmes plus appliqués, notamment en bio-informatique, est souhaitable. Une politique plus volontariste d'échanges internationaux serait profitable scientifiquement.

La mise en place future d'un « data center » sur le site de Jouy pourrait être l'occasion de repenser le rôle et le dimensionnement de la plate-forme MIGALE, si celle-ci est amenée à en prendre la responsabilité.

Une augmentation du nombre d' HDR et du nombre de thèses dans la prochaine période d'évaluation serait souhaitable. Un recentrage sur un plus petit nombre d'écoles doctorales est également souhaitable. Finalement, l'enjeu immédiat est l'organisation de la nouvelle unité. Il convient d'accélérer la réflexion sur ce sujet et de commencer à mettre en place des mutualisations et des collaborations. A terme le regroupement spatial des deux équipes est souhaitable.



3 • Appréciations détaillées

Appréciation sur la production et la qualité scientifiques

La production scientifique apparaît très bonne en quantité et en qualité : 129 articles ont été publiés dans des revues à comité de lecture entre janvier 2008 et juin 2013 par l'équipe 1 (MIAj), 104 par l'équipe 2 (MIG), dont une bonne moitié dans de très bonnes revues. Le panel est très large allant de revues théoriques comme *Annals of Statistics* ou *Bernoulli*, à des journaux généralistes tels que *Science* et des revues appliquées en biologie, épidémiologie et écologie notamment. Il y a également des revues de méthodologie statistique (au sens large) tels que *PLOS Computational Biology*, *Biometrics*, *Computational Statistics and Data Analysis*. Les deux équipes sont également présentes dans de nombreuses conférences nationales et internationales.

Appréciation sur le rayonnement et l'attractivité académiques

L'unité entretient de nombreuses collaborations dont beaucoup ont débouché sur des co-publications. Les collaborations se font avec des UMR, en particulier MICALIS (UMR associant l'INRA et Agro-ParisTech), et des unités CNRS. L'unité a eu de nombreux projets financés, participe à plusieurs réseaux de recherche : GDR, réseau National des Systèmes Complexes, réseaux INRA ; elle a organisé des séminaires, des conférences et des formations et participé à des comités de programmes de congrès. Durant la période d'évaluation elle a reçu la visite de plusieurs professeurs étrangers et accueilli des post-docs. Des membres de l'unité font partie de comités éditoriaux (en particulier *Biometrics*, *Statistical Science*, *Scandinavian Journal of Statistics*, *BMC Structural Biology*, *Biophysical reviews*).

Appréciation sur l'interaction avec l'environnement social, économique et culturel

L'unité a réalisé plusieurs rapports d'étude et documents de diffusion, et participation à des comités ou expositions. En particulier, pour MIG : comité éditorial de la collection de l'INRA « Sciences en question », exposition « Des maths partout... », forums « Emploi Math » et « Vitae », article dans la *Gazette des mathématiciens*. Pour MIAj : comité d'experts de l'ANSES, rapport sur la contamination de coquillages marins, chapitre sur la modélisation pour une commission interministérielle, participation à la rédaction du code d'éthique de l'European Mathematical Society, participation au comité d'éthique en expérimentation animale du Centre de Jouy, témoignage sur la vie de doctorant dans une vidéo du Centre. Les deux équipes ont accueilli des stagiaires pré-baccalauréat.

Particulièrement par le biais de la plateforme MIGALE, l'unité interagit fortement avec l'environnement de la recherche. L'ensemble de ces activités, sans être particulièrement important au vu de la taille de l'unité conjointe, montre l'existence d'une interaction satisfaisante avec l'environnement.

Appréciation sur l'organisation et la vie de l'unité

L'unité est composée de deux UPR (ci-dessous désignées par le terme équipes) jusqu'à maintenant séparées et appelées à fusionner pendant le prochain contrat. Chaque équipe a donc fonctionné jusqu'ici de manière autonome. Dans chaque équipe des réunions régulières ont été organisées. Il y a également des séminaires et des groupes de lecture. Les ingénieurs, techniciens et administratifs ont leur place dans la vie des équipes. Les locaux sont satisfaisants, mais les locaux des deux équipes sont séparés, quoique d'une courte distance. L'équipe 1 (MIAj) est organisée en trois sous-équipes et l'équipe 2 (MIG) est constituée de quatre sous-équipes. La gestion financière et l'informatique sont mutualisées au sein de chacune des équipes et une bibliothécaire est mutualisée pour les deux équipes. L'organisation de la vie de la nouvelle unité reste à construire.

Appréciation sur l'implication dans la formation par la recherche

L'unité est impliquée dans l'enseignement et l'encadrement au niveau master et doctoral. Les membres de l'unité ont principalement délivré des cours de statistique, de mathématique appliquées, de bioinformatique et de biologie des systèmes dans des écoles (AgroParisTech, EHESP, ENS Cachan, Institut Pasteur, ENSAI, INSEE, Supélec) et des universités (AgroParisTech, Université d'Evry-Val d'Essonne). La plateforme Migale organise un cycle de formation « Bioinformatique par la pratique » proposant une douzaine de modules et suivi en moyenne par 145 personnes chaque année. Des membres de l'unité ont participé à l'organisation de plusieurs écoles-chercheurs (Epicasa, Mexico, BioBayes, Aspen).



Vingt-et-une thèses ont été soutenues sur la période et il y a vingt-trois doctorants. Les thèses se passent bien et sont financées. De nombreuses thèses sont en co-direction. On remarque qu'il y a dix chercheurs habilités à diriger des recherches dans l'unité et qu'aucune HDR n'a été passée pendant la dernière période. Cependant une chercheuse vient d'obtenir son HDR (après rédaction du rapport). La principale école doctorale est ABIES (Agriculture, Alimentation, Biologie, Santé, ED 435), portée par AgroParisTech ; un membre de l'unité fait partie du conseil de cette ED. Une autre école bien représentée est GAO (des Génomes Aux Organismes, ED 423, commune entre Versailles et Evry). L'unité a aussi des liens avec des ED de mathématiques appliquées, en particulier l'ED142 Mathématique de la région Paris Sud-Paris 11 (au conseil de laquelle laquelle participe un membre de l'unité). On relève en plus d'ABIES, de GAO et de l'ED142 un grand nombre d'autres écoles doctorales impliquées dans les thèses passées et en cours. Cette diversité d'écoles doctorales peut s'expliquer par la diversité des thèmes et des collaborations de l'unité, mais risque de nuire à son identité propre. L'unité pourrait avantageusement profiter de la création de nouvelles écoles doctorales sur Paris-Saclay en 2015, en particulier ED STIC qui regroupera EDIPS et EDITE, pour se recentrer sur quelques ED emblématiques de ses activités. La fondation mathématique Jacques Hadamard pourrait également aider à regrouper les recherches dans ce secteur.

Appréciation sur la stratégie et le projet à cinq ans

Sur le plan scientifique, l'unité relève que le développement de techniques produisant une masse de données de plus en plus importante en imagerie et dans les différentes « omiques » (génomique, transcriptomique, protéomique, métabolomique...) offre un champ de plus en plus large aux statistiques et aux mathématiques appliquées. Il n'y a pas eu de présentation détaillée de la stratégie scientifique sur cinq ans en dehors de la continuation des recherches actuelles. Le seul nouveau projet clair est le projet de Data Center porté principalement par la plate-forme MIGALE. Sur le plan de l'organisation de l'unité, l'enjeu principal est la fusion des équipes (anciennement unités) MIAj et MIG dans une nouvelle unité. Un comité de fusion a été créé. Le projet relève la complémentarité des compétences, la possibilité de mutualiser les services administratifs et techniques, et l'effet de taille qui donnera une visibilité accrue de la nouvelle unité. Elle a la potentialité de devenir un pôle fort dans le domaine des mathématiques, statistiques et de l'informatique et de leurs applications en biologie au sein du centre INRA de Jouy et de l'ensemble Paris Saclay qui se structure. Les membres des deux équipes travaillent à cette fusion dans un esprit positif. Cependant on ne sait pas encore comment sera organisée la nouvelle unité. On peut penser que la fusion se fera progressivement. Le fait que les deux équipes occupent des locaux dans des bâtiments différents risque de ralentir ce processus.



4 • Analyse équipe par équipe

Équipe 1 : Mathématiques et Informatiques Appliquées - Jouy-en-Josas

Nom du responsable : M. Hervé MONOD

Effectifs

| Effectifs de l'équipe | Nombre au 30/06/2013 | Nombre au 01/01/2015 |
|--|----------------------|----------------------|
| N1 : Enseignants-chercheurs titulaires et assimilés | 1 | |
| N2 : Chercheurs des EPST ou EPIC titulaires et assimilés | 11 | |
| N3 : Autres personnels titulaires (n'ayant pas d'obligation de recherche) | 10,8 | |
| N4 : Autres enseignants-chercheurs (PREM, ECC, etc.) | | |
| N5 : Autres chercheurs des EPST ou EPIC (DREM, Post-doctorants, visiteurs etc.) | 2 | |
| N6 : Autres personnels contractuels (n'ayant pas d'obligation de recherche) | | |
| TOTAL N1 à N6 | 24,8 | |

| Effectifs de l'équipe | Nombre au 30/06/2013 | Nombre au 01/01/2015 |
|---|----------------------|----------------------|
| Doctorants | 11 | |
| Thèses soutenues | 11 | |
| Post-doctorants ayant passé au moins 12 mois dans l'unité | | |
| Nombre d'HDR soutenues | | |
| Personnes habilitées à diriger des recherches ou assimilées | 6 | |



• Appréciations détaillées

Appréciation sur la production et la qualité scientifiques

Cette équipe est en fait l'UPR MIAj, appelée à se fondre dans la nouvelle structure. La production scientifique apparaît très bonne en quantité et en qualité : 129 articles ont été publiés dans des revues à comité de lecture entre janvier 2008 et juin 2013, dont 68 dans des revues classées excellentes ou exceptionnelles selon la grille NORIA. On remarque des articles publiés dans des très bonnes revues théoriques comme Annals of Statistics ou Bernoulli, ou encore Statistical Science. Il y a également des revues de méthodologie statistique (au sens large) tels que PLOS Computational Biology, Biometrics, Computational Statistics and Data Analysis, et des revues d'écologie et de biologie. L'équipe est également présente dans de nombreuses conférences nationales et internationales.

Appréciation sur le rayonnement et l'attractivité académiques

L'équipe entretient de nombreuses collaborations dont beaucoup ont débouché sur des co-publications. Les collaborations se font avec l'équipe MIG et avec des UMR, en particulier MICALIS (UMR associant l'INRA et AgroParisTech), et des unités CNRS. L'équipe a eu de nombreux projets financés, quoique pour des budgets relativement faibles, et participe à plusieurs réseaux de recherche : GDR, réseau National des Systèmes Complexes, réseaux INRA ; elle a organisé des séminaires, des conférences et des formations, et participé à des comités de programme de congrès. Durant la période d'évaluation elle a reçu la visite de professeurs étrangers et accueilli des post-docs. Des membres de l'équipe font partie de comités éditoriaux (en particulier Biometrics et Statistical Science).

Appréciation sur l'interaction avec l'environnement social, économique et culturel

L'équipe a réalisé trois rapports d'étude et trois documents de diffusion. Elle a accueilli 14 stagiaires pré-baccalauréat. Par ailleurs, des membres de l'équipe sont ou ont été membres de comités d'experts (ANSES) ou d'éthique (Société Européenne de Mathématiques, COMité d'ETHique en Expérimentation Animale -COMETHEA-). Il semble toutefois que ces interactions, non négligeables, reposent davantage sur des initiatives individuelles que sur une stratégie d'ensemble de l'équipe.

Appréciation sur l'organisation et la vie de l'équipe

Des réunions mensuelles sont organisées. Il y a également des séminaires et des groupes de lecture. Les ingénieurs, techniciens et administratifs ont leur place dans la vie de l'équipe. Les locaux sont satisfaisants. L'équipe est organisée en trois sous-équipes, mais a souhaité se présenter sous la forme de deux « thèmes » pour l'évaluation présente. La gestion financière et l'informatique sont mutualisées entre ces thèmes.

Appréciation sur l'implication dans la formation par la recherche

L'équipe est impliquée dans l'enseignement et l'encadrement au niveau master. Les membres de l'équipe ont principalement délivré des cours de statistique et de mathématique appliquée dans des écoles d'ingénieurs (AgroParisTech, EHESP, ENS Cachan, Institut Pasteur, ENSAI, INSEE) et des universités. La principale école doctorale est ABIES, portée par AgroParisTech, mais l'équipe a aussi des liens avec des ED de Mathématiques appliquées. Onze thèses ont été soutenues sur la période.

Appréciation sur la stratégie et le projet à cinq ans

Sur le plan scientifique, l'équipe relève que le développement de techniques produisant une masse de données de plus en plus importante en imagerie et dans les différentes « omiques » offre un champ de plus en plus large aux statistiques et aux mathématiques appliquées. Sur le plan de la stratégie, l'équipe envisage la fusion avec l'équipe MIG d'une manière positive. Elle relève la complémentarité des compétences, la possibilité de mutualiser les services administratifs et techniques, et l'effet de taille qui donnera une visibilité accrue à la nouvelle unité.



Conclusion

- *Points forts et possibilités liées au contexte :*

Quantité et qualité des publications. Vie d'équipe qui fonctionne bien, nombreuses activités de recherche et de formation.

- *Points à améliorer et risques liés au contexte :*

Les activités d'expertise ne semblent pas très importantes. Une certaine hétérogénéité des publications entre des travaux très théoriques et les applications. Le spectre est large mais il est souhaitable de renforcer la continuité entre ces domaines.

- *Recommandations :*

Vu l'importance des activités, nous encourageons une politique d'engagement de nouvelles thèses plus active que celle en cours ces dernières années. Il ne fait aucun doute que l'équipe MIAj a en réserve de nombreux et intéressants sujets de thèse en statistique méthodologique et appliquée à proposer. L'orientation de la recherche par des sujets finalisés donne de bons résultats et mérite d'être poursuivie résolument. Il est souhaitable de poursuivre les recherches sur la méthodologie statistique pour renforcer le pont entre recherche théorique et applications.



Équipe 1 : Thème 1 Statistique et extraction d'information

Nom du responsable : M^{me} Sylvie HUET

Effectifs

Les effectifs seront donnés en ETP.

| Effectifs du thème en Équivalents Temps Plein | Au 30/06/2013 | Au 01/01/2015 |
|---|------------------|------------------|
| ETP d'enseignants-chercheurs titulaires MLT | 1 | |
| ETP de chercheurs des EPST ou EPIC titulaires | 6,5 | |
| ETP d'autres personnels titulaires n'ayant pas d'obligation de recherche (IR, IE, PRAG, etc.) | 3 | |
| ETP d'autres enseignants-chercheurs (PREM, ECC, etc.) | | |
| ETP de post-doctorants ayant passé au moins 12 mois dans l'unité | | |
| ETP d'autres chercheurs des EPST ou EPIC (DREM, etc.) hors post-doctorants | | |
| ETP d'autres personnels contractuels n'ayant pas d'obligation de recherche | | |
| ETP de doctorants | 5 | |
| TOTAL | 15,5 | |

• **Appréciations détaillées**

Ce thème est ancien dans l'équipe MIAJ dont les (sous-)équipes Dynenvie et Megadim qui y sont impliquées comprennent un nombre important de statisticiens. Ces équipes bénéficient d'une bonne notoriété et possèdent un dynamisme important qui se traduit par des collaborations avec d'autres groupes de statisticiens de la région parisienne et une activité notable dans l'animation de groupes de travail. Ainsi, les derniers recrutements de chercheurs réalisés sont de qualité et répondent bien à une orientation heureuse consistant à privilégier des travaux visant des solutions algorithmiques performantes, des réalisations logicielles et des mises en application efficaces de modèles statistiques complexes en génomique et en épidémiologie.

Les chantiers ouverts sur ces thèmes sont nombreux et, en règle générale, bien guidés par des buts biologiques clairement identifiés. Ainsi les publications des deux équipes Dynenvie et Megadim sont nombreuses et présentent un bon équilibre entre articles dans des journaux de biologie (environ la moitié), des journaux de statistique (environ un quart) et des journaux à l'interface des deux disciplines (un quart). Autant que nous avons pu en juger, ils donnent lieu à un nombre respectable de logiciels de nature essentiellement expérimentale. Parmi eux, on peut mettre en exergue le logiciel FILTEREX de filtrage particulière pour la microbiologie alimentaire qui apparaît comme un effort judicieux pour la diffusion de ces techniques modernes d'inférence bayésienne.

De même, parmi les nombreux travaux réalisés sur les thèmes de la statistique et de l'extraction d'information, nous indiquons quelques caractéristiques intéressantes sur des contributions d'envergure et porteuses.

1. Les contributions en analyse de sensibilité ont trait aux plans d'expérience et prennent bien en compte les particularités des applications agronomiques visées.
2. Les travaux sur la sélection de modèles, avant tout théoriques et méthodologiques, sont de grande portée pour la modélisation de données de grande dimension, notamment pour la génomique.
3. Les travaux pour l'estimation de modèles complexes à structure cachée, utiles en épidémiologie et en analyse d'images, sont algorithmiques et utilisent massivement des simulations. Le souci de veiller à une rapidité raisonnable de ces algorithmes coûteux est appréciable.

La situation apparaît donc très saine et les recherches de MIAJ en statistique continueront sans nul doute d'être ambitieuses et utiles. Nous indiquons juste deux recommandations pour améliorer encore l'impact des recherches en statistique.

1. Vu l'importance des activités, nous encourageons une politique d'engagement de nouvelles thèses plus active que celle en cours ces dernières années. Il ne fait aucun doute que ce thème a en réserve de nombreux et intéressants sujets de thèses en statistique méthodologique et appliquée, à proposer.
2. Enfin, l'orientation de la recherche par des sujets finalisés donne de bons résultats et mérite d'être poursuivie résolument.



Équipe 1 : Thème 2 Modélisation pour la biologie et l'évaluation du risque

Nom du responsable : M^{me} Béatrice LAROCHE

Effectifs

Les effectifs seront donnés en ETP.

| Effectifs du thème en Équivalents Temps Plein | Au 30/06/2013 | Au 01/01/2015 |
|---|------------------|------------------|
| ETP d'enseignants-chercheurs titulaires | | |
| ETP de chercheurs des EPST ou EPIC titulaires | 4,5 | |
| ETP d'autres personnels titulaires n'ayant pas d'obligation de recherche (IR, IE, PRAG, etc.) | 3 | |
| ETP d'autres enseignants-chercheurs (PREM, ECC, etc.) | | |
| ETP de post-doctorants ayant passé au moins 12 mois dans l'unité | | |
| ETP d'autres chercheurs des EPST ou EPIC (DREM, etc.) hors post-doctorants | | |
| ETP d'autres personnels contractuels n'ayant pas d'obligation de recherche | | |
| ETP de doctorants | 6 | |
| TOTAL | 13,5 | |

• **Appréciations détaillées**

Les activités décrites dans la section « Modélisation de phénomènes en biologie et évaluation du risque » s'orientent autour de quatre axes principaux : la biologie du développement, la physiologie digestive, l'épidémiologie et la sécurité alimentaire.

Le premier axe « Biologie du développement », dont l'objectif est la conception, le développement et l'application de modèles mathématiques pour le processus de biogénèse des corps lipidiques, l'embryogénèse d'A. Thaliana ou la morphogénèse du trophoblaste, se prête par définition tout particulièrement aux coopérations avec des biologistes de l'INRA. Les résultats présentés pour cette thématique se révèlent toutefois assez modestes autant en termes de quantité que de qualité (peu d'articles dans des revues). L'impact de ces activités pourra être amélioré par une politique de publication plus dynamique.

Au sein du deuxième axe « Physiologie digestive », les activités autour de la modélisation du microbiote intestinal humain ont fait l'objet de bonnes publications dans des revues et conférences de biologie et d'interface au début de la période d'évaluation dans le cadre d'une thèse, tandis que les activités autour de la modélisation de la digestion des protéines ont débouché sur des publications biologiques (au sens large) en coopération avec différents laboratoires. La collaboration avec l'équipe MEGADIM ainsi qu'une stratégie de publication orientée également vers les revues d'interface pourront permettre de continuer à mettre en valeur avantageusement ces activités fructueuses.



L'axe « épidémiologie », qui s'intéresse à la modélisation de phénomènes dynamiques aux propriétés particulières dans le contexte de l'épidémiologie animale ou végétale, fait preuve d'un réel dynamisme qu'illustre le nombre important de collaborations, de thèses (co-)encadrées et de publications issues entre autres de ces thèses aussi bien dans des revues d'interface telles que *Journal of Theoretical Biology* que dans des revues appliquées telles que *Veterinary Science*. Ces résultats très positifs devraient permettre à certains membres permanents de l'équipe de passer leur HDR. On notera ici un fort potentiel en termes d'interdisciplinarité (mathématiques, interface, biologie) et d'application de la « recherche méthodologique en train de se faire » dans des projets en collaboration avec des biologistes. On pourra continuer à mettre l'accent sur ce transfert des outils méthodologiques en pratique dans la prochaine période.

Enfin, l'axe « sécurité alimentaire », qui s'orientait surtout autour de la coopération avec l'unité Mét@risk destinée à être dissoute, est appelé à évoluer. L'expertise acquise par l'unité autour de cette thématique et illustrée par des publications de haut niveau impliquant l'application d'outils mathématiques complexes à des problèmes pratiques pourra être mise en valeur dans le cadre d'autres projets, en particulier ceux concernant l'inférence de graphes au sens large ou l'épidémiologie.

De façon plus générale, on notera aussi pour cette thématique une potentielle synergie avec les autres activités de MIAj plus axées sur les statistiques et, dans le cadre de la fusion MIAj-MIG, avec l'équipe de biologie systémique de MIG travaillant également sur des problèmes de modélisation de systèmes dynamiques (au sens large) et avec l'équipe GenEvol de MIG sur les aspects moléculaires de l'épidémiologie.



Équipe 2 : Mathématique, Informatique et Génome

Nom du responsable : M^{me} Sophie SCHBATH

Effectifs

| Effectifs de l'équipe | Nombre au 30/06/2013 | Nombre au 01/01/2015 |
|--|----------------------|----------------------|
| N1 : Enseignants-chercheurs titulaires et assimilés | 1 | |
| N2 : Chercheurs des EPST ou EPIC titulaires et assimilés | 9 | |
| N3 : Autres personnels titulaires (n'ayant pas d'obligation de recherche) | 13,2 | |
| N4 : Autres enseignants-chercheurs (PREM, ECC, etc.) | | |
| N5 : Autres chercheurs des EPST ou EPIC (DREM, Post-doctorants, visiteurs etc.) | 2 | |
| N6 : Autres personnels contractuels (n'ayant pas d'obligation de recherche) | 2 | |
| TOTAL N1 à N6 | 27,2 | |

| Effectifs de l'équipe | Nombre au 30/06/2013 | Nombre au 01/01/2015 |
|---|----------------------|----------------------|
| Doctorants | 12 | |
| Thèses soutenues | 10 | |
| Post-doctorants ayant passé au moins 12 mois dans l'unité | | |
| Nombre d'HDR soutenues | | |
| Personnes habilitées à diriger des recherches ou assimilées | 4 | |

• Appréciations détaillées

Appréciation sur la production et la qualité scientifiques

Cette équipe est en fait l'UPR MIG, appelée à se fondre dans la nouvelle structure. La production scientifique apparaît très bonne en quantité et en qualité : 104 articles ont été publiés dans des revues à comité de lecture entre janvier 2008 et juin 2013, dont 64 dans des revues classées excellentes ou exceptionnelles selon la grille NORIA. On remarque trois articles publiés dans Science. Il y a également des revues de méthodologie statistique ou computationnelle (au sens large) tels que BMC Bioinformatics et des revues de biologie. L'équipe est également présente dans de nombreuses conférences nationales et internationales.



Appréciation sur le rayonnement et l'attractivité académiques

L'équipe entretient de nombreuses collaborations, locales, nationales ou internationales, dont beaucoup ont débouché sur des co-publications. Les collaborations se font avec l'équipe MIAj et avec des UMR, en particulier MICALIS (UMR associant l'INRA et Agro-ParisTech), et des unités CNRS. L'équipe MIG participe en particulier au groupe Statistics for System Biology (SSB) qui structure les activités de bioinformatique du département MIA de l'INRA. L'équipe a eu de nombreux projets financés, avec des budgets parfois importants, et participe à plusieurs réseaux de recherche : GDR, réseau National des Systèmes Complexes, réseaux INRA ; elle a organisé des séminaires, des conférences et des formations et a participé à des comités de programmes de congrès. Durant la période d'évaluation elle a reçu la visite de professeurs étrangers et a accueilli des post-docs. Des membres de l'équipe font partie de comités éditoriaux : Scandinavian Journal of Statistics, BMC Structural Biology, Biophysical reviews.

Appréciation sur l'interaction avec l'environnement social, économique et culturel

Particulièrement par le biais de la plateforme MIGALE, l'équipe interagit fortement avec l'environnement de la recherche. Elle a également été impliquée dans plusieurs actions de diffusion et de vulgarisation, notamment via le groupe « Science en questions ». Ces actions sont éminemment souhaitables dans un laboratoire dont les thématiques sont plus directement en prise sur la société que celles portées par la majorité des laboratoires de mathématiques : l'équipe doit donc veiller à maintenir ce type d'activité sur le moyen terme.

Appréciation sur l'organisation et la vie de l'équipe

L'équipe est organisée en quatre sous-équipes (qui avaient le statut d'équipes de l'UPR lors du contrat précédent) : GenEvol (structure et évolution du génome), Bibliome (exploitation du bibliome), Biosys (biologie des systèmes) et la plateforme MIGALE. La vie de l'équipe est bien organisée avec un « conseil de service » qui se réunit tous les deux mois. Il y a également quatre assemblées générales par an. Un séminaire d'équipe est organisé tous les deux mois. Un intranet permet la diffusion d'informations. La gestion et l'informatique sont mutualisées entre les thèmes et la plateforme MIGALE fournit un fort soutien informatique.

Appréciation sur l'implication dans la formation par la recherche

L'équipe est impliquée dans l'enseignement et l'encadrement au niveau master. Les membres de l'équipe ont principalement délivré des cours de statistique, de bioinformatique et de biologie des systèmes dans des écoles d'ingénieurs et des universités. Les doctorants sont rattachés à diverses écoles doctorales dont ABIES et GAO (Génomés Aux Organismes, ED commune entre Versailles et Evry). Dix thèses ont été soutenues sur la période. L'équipe Migale organise un cycle de formation «Bioinformatique par la pratique» proposant une douzaine de modules et suivi en moyenne par 145 personnes chaque année.

Appréciation sur la stratégie et le projet à cinq ans

Sur le plan scientifique l'équipe relève que le développement de techniques produisant une masse de données de plus en plus importante en imagerie et dans les différentes « omiques » offre un champ de plus en plus large aux statistiques et aux mathématiques appliquées. Il existe un projet de « Data center » dans lequel l'équipe MIGALE aurait un rôle important.

Sur le plan de la stratégie, l'équipe envisage la fusion avec l'équipe MIAj d'une manière positive. Elle relève la complémentarité des compétences, la possibilité de mutualiser les services administratifs et techniques, et l'effet de taille qui donnera une visibilité accrue à la nouvelle unité.

Conclusion

- *Points forts et possibilités liées au contexte :*

Quantité et qualité des publications ; on remarque en particulier plusieurs articles dans Science ; vie d'équipe qui fonctionne bien ; nombreuses activités de recherche et de formation.



- *Points à améliorer et risques liés au contexte :*

Le thème MIGALE est sous-dimensionné si l'on considère qu'il a vocation à jouer un rôle de plate-forme bio-informatique destinée à l'ensemble du site, et à répondre aux sollicitations de l'ensemble de ses chercheurs.

Le thème BioSys devra veiller à ne pas perdre son focus et disperser ses forces. En particulier, la vie à moyen terme d'une équipe suppose la présence de plusieurs chercheurs habilités à diriger des recherches. Le thème Bibliome a intérêt à un recentrage des efforts de l'outil informatique vers des questions biologiques qui est difficile à réaliser dans une petite équipe ; mais le recrutement récent d'un nouveau chercheur et la fusion de MIG et MIAj en cours peuvent ouvrir de nouvelles opportunités.

- *Recommandations :*

Un développement des interactions avec les statisticiens de l'unité permettrait de renforcer les activités du thème GenEvol.

Une politique claire de priorisation et de financement des projets auxquels l'équipe MIGALE participe devrait être établie. La mise en place future d'un « data center » sur le site de Jouy pourrait être l'occasion de repenser le rôle et le dimensionnement de la plate-forme, si elle est amenée à en prendre la responsabilité.

Une augmentation du nombre d' HDR dans la prochaine période d'évaluation serait souhaitable, en particulier pour le thème BioSys.



Équipe 2 : Thème 1 Structure et Evolution des Génomes + Plateforme Migale

Nom du responsable : M. Philippe BESSIERES et M. Franck SAMSON

Effectifs

Les effectifs seront donnés en ETP.

| Effectifs du thème en Équivalents Temps Plein | Au 30/06/2013 | Au 01/01/2015 |
|---|------------------|------------------|
| ETP d'enseignants-chercheurs titulaires | 1 | |
| ETP de chercheurs des EPST ou EPIC titulaires | 5,5 | |
| ETP d'autres personnels titulaires n'ayant pas d'obligation de recherche (IR, IE, PRAG, etc.) | 8,5 | |
| ETP d'autres enseignants-chercheurs (PREM, ECC, etc.) | | |
| ETP de post-doctorants ayant passé au moins 12 mois dans l'unité | | |
| ETP d'autres chercheurs des EPST ou EPIC (DREM, etc.) hors post-doctorants | 2 | |
| ETP d'autres personnels contractuels n'ayant pas d'obligation de recherche | 1 | |
| ETP de doctorants | 4 | |
| TOTAL | 22 | |

• **Appréciations détaillées**

Ce thème correspond en fait à la réunion de l'équipe GenEvol et de la plateforme MIGALE de l'UPR MIG. Il a une bonne production scientifique, composée à la fois de contributions méthodologiques originales publiées dans des journaux et conférences de statistique et bio-informatique, et de participations à des projets plus appliqués en biologie (notamment en annotation de génomes bactériens) donnant lieu à des publications dans des journaux parfois à très fort impact (Science, Cell, PLoS Genetics...).

GenEvol a une bonne visibilité au sein de l'INRA et au niveau national, via notamment son implication dans le groupe « Statistics for Systems Biology » (SSB) et sa participation aux sociétés savantes comme la Société Française de Bio-informatique (SFBI) dont elle a la présidence depuis 2010. Sa visibilité internationale est comparativement moins établie, et pourrait faire l'objet d'une politique plus volontariste pour favoriser les échanges de chercheurs et doctorants avec des laboratoires reconnus au niveau international, et pour augmenter la participation à des projets de recherche internationaux.

Les principales contributions méthodologiques originales sont de bonne qualité. Les travaux sur l'analyse des motifs dans les séquences biologiques, une des forces historiques de l'équipe, ont été étendus à l'analyse de motifs dans des réseaux, et constituent une voie d'investigation intéressante pour l'analyse future de masses de données représentées sous forme de graphes. Les travaux sur les structures 3D de protéines sont de bonne qualité mais GenEvol n'a peut-être pas une masse critique suffisante sur cette thématique bien établie pour rivaliser avec les principaux compétiteurs. Les travaux sur la transcriptomique et la comparaison de génomes microbiens sont nouveaux et intéressants, même si l'on peut regretter la quantité encore relativement faible de contributions méthodologiques originales concernant l'analyse de données issues de séquençage à haut débit (NGS), qui a véritablement révolutionné la production de données « omiques » depuis 5 ans. Compte tenu de l'expertise de l'équipe, il s'agit d'une thématique de recherche qui devrait être favorisée à l'avenir.



GénEvol a participé à différentes actions de vulgarisation de la science au niveau local. Bien qu'elle travaille sur des questions qui intéressent certainement de nombreuses entreprises agro-alimentaires ou de la santé, elle n'a aucune collaboration industrielle qui pourrait offrir une source intéressante de questions scientifiques et de financement, dans un contexte où les financements de l'équipe par projets publics ont fortement diminué ces dernières années.

L'environnement et les conditions de travail sont agréables, et l'encadrement scientifique et humain semble efficace.

GénEvol a un volume satisfaisant d'encadrement de stages (environ 5 par an), mais seulement 5 thèses y ont été soutenues au cours des 5 dernières années, et 4 sont actuellement en cours, la plupart sous forme de co-encadrement avec un chercheur extérieur à l'unité. Ce sont des volumes assez faibles, compte tenu de la capacité d'encadrement de GénEvol, qui n'accueille par ailleurs plus de chercheur post-doctorant entre janvier et juin 2013.

En résumé, GénEvol possède une très bonne expertise en statistique et algorithmique pour la bio-informatique, et participe à de nombreux projets appliqués à forte visibilité. Elle a trouvé un bon équilibre entre production méthodologique et contributions appliquées, que nous ne pouvons qu'encourager. Elle a un potentiel d'encadrement scientifique actuellement sous-exploité, qui offre des opportunités intéressantes pour l'avenir. La recherche de financements pourrait passer par une participation accrue à des projets internationaux, qui passe elle-même par une politique plus volontariste de promotion d'échanges internationaux, et par la recherche de projets avec des partenaires industriels. Elle a tous les atouts pour devenir un acteur majeur de l'analyse de données issues du séquençage à haut débit en micro-biologie et métagénomique. Un développement des interactions avec les statisticiens de l'unité MIAj dans la future unité aurait tout son sens dans ce cadre.

MIGALE est une plate-forme de service bio-informatique, dont les missions comprennent la gestion d'infrastructure de calcul pour la bio-informatique, la diffusion de savoirs, le développement logiciel et l'hébergement de projets. Ses membres sont scientifiquement rattachés à l'équipe GenEvol.

MIGALE a donc un rôle essentiellement de service. La plate-forme joue un rôle important dans la plupart des projets, en nombre croissant, impliquant la manipulation et l'analyse de grandes quantités de données. Son activité de formation aux outils de bio-informatique est un succès auprès des chercheurs de l'INRA, et a toute les chances de rencontrer un public toujours plus nombreux à l'avenir.

Le rôle de valorisation des développements méthodologiques effectués dans l'unité par leur implémentation dans des systèmes ergonomiques est très important, et représente une force pour l'unité.

MIGALE est sous-dimensionnée si l'on considère qu'elle a vocation à jouer un rôle de plate-forme bio-informatique à l'ensemble du site, et à répondre aux sollicitations de l'ensemble de ses chercheurs. Une politique claire de priorisation et de financement des projets auxquels elle participe devrait être établie. La mise en place future d'un « data center » sur le site de Jouy pourrait être l'occasion de repenser le rôle et le dimensionnement de la plate-forme, si elle est amenée à en prendre la responsabilité.



Équipe 2 : Thème 2 Biologie des systèmes

Nom du responsable : M. Vincent FROMION

Effectifs

Les effectifs seront donnés en ETP.

| Effectifs du thème en Équivalents Temps Plein | Au 30/06/2013 | Au 01/01/2015 |
|---|------------------|------------------|
| ETP d'enseignants-chercheurs titulaires | | |
| ETP de chercheurs des EPST ou EPIC titulaires | 2 | |
| ETP d'autres personnels titulaires n'ayant pas d'obligation de recherche (IR, IE, PRAG, etc.) | 1,5 | |
| ETP d'autres enseignants-chercheurs (PREM, ECC, etc.) | | |
| ETP de post-doctorants ayant passé au moins 12 mois dans l'unité | | |
| ETP d'autres chercheurs des EPST ou EPIC (DREM, etc.) hors post-doctorants | | |
| ETP d'autres personnels contractuels n'ayant pas d'obligation de recherche | | |
| ETP de doctorants | 5 | |
| TOTAL | 8,5 | |

• **Appréciations détaillées**

L'activité de BioSys au sein de MIG a démarré en 2005. L'objectif de BioSys est le développement de modèles mathématiques de systèmes biologiques, ainsi que les outils mathématiques et algorithmiques pour leur analyse. Les modèles et les méthodes développés, souvent inspirés par les approches en automatique, ont la vocation d'être génériques, mais l'équipe a également beaucoup investi dans leur application à des questions biologiques concrètes, en étroite collaboration avec des biologistes de l'INRA et d'ailleurs. Le système biologique sur lequel se sont concentrés une grande partie des efforts est la bactérie modèle *Bacillus subtilis*, en collaboration avec le département MICA.

La production de BioSys est très bonne et équilibrée entre publications en mathématiques appliquées et en automatique, publications en (micro)biologie et publications à l'interface de ces domaines. La proximité des problématiques biologiques est démontrée par le fait que BioSys a contribué à plusieurs publications à très haut impact en biologie (*Science*), dans le cadre du consortium européen BaSysBio. Des résultats notables au niveau méthodologique sont la méthode Resource Balance Analysis (RBA), qui reprend de façon originale les approches de Flux Balance Analysis (FBA) pour y rajouter des contraintes liées à la répartition des ressources entre les différentes fonctions de la cellule, et les approches discrètes pour analyser la structure et la modularité de réseaux de régulation à l'aide de critères donnés par la dynamique de ces systèmes, un travail effectué en collaboration avec l'INRIA Sophia-Antipolis. BioSys a un excellent taux de participation à des projets nationaux et internationaux, dont notamment les consortia européens BaSysBio et BaSynthec qui ont contribué au rayonnement national et international de l'équipe. Les membres de BioSys se sont également fortement investis dans l'encadrement de thèses (5 soutenues, et 5 en cours sur la période d'évaluation) et ont participé aux enseignements dans des écoles d'ingénieurs et un master en région parisienne. Les méthodes proposées ont été accompagnées par des logiciels, qui ont donné lieu à une publication (Basylica dans *Bioinformatics*). L'impact du logiciel sur la communauté d'utilisateurs potentiels pourrait être amélioré par le développement d'un site web de bonne qualité.

BioSys a jeté des bases très solides pour son activité en biologie des systèmes, à l'interface de l'automatique et de la biologie, dans cette période d'évaluation. Le grand nombre de collaborations nationales et internationales mises en place, notamment celles avec des biologistes de MICALIS à Jouy, sont un atout certain pour BioSys et plus largement pour le nouveau laboratoire en cours de réalisation. BioSys est une petite équipe de trois permanents active dans un domaine extrêmement compétitif. L'investissement dans l'Institut de Modélisation des Sciences du Vivant est un excellent moyen de développer les activités de BioSys dans le contexte du plateau de Saclay, mais BioSys devra veiller à ne pas perdre son focus et disperser ses forces. Plusieurs de ses membres ont atteint une indépendance et une maturité scientifique, entre autres par leur implication dans l'encadrement de thèses, qui devrait leur permettre de passer une HDR dans la prochaine période d'évaluation.



Équipe 2 : Thème 3 Extraction de connaissances à partir de textes

Nom du responsable : M^{me} Claire NEDELLEC

Effectifs

Les effectifs seront donnés en ETP.

| Effectifs du thème en Équivalents Temps Plein | Au 30/06/2013 | Au 01/01/2015 |
|---|------------------|------------------|
| ETP d'enseignants-chercheurs titulaires | | |
| ETP de chercheurs des EPST ou EPIC titulaires | 1,5 | |
| ETP d'autres personnels titulaires n'ayant pas d'obligation de recherche (IR, IE, PRAG, etc.) | 1 | |
| ETP d'autres enseignants-chercheurs (PREM, ECC, etc.) | | |
| ETP de post-doctorants ayant passé au moins 12 mois dans l'unité | | |
| ETP d'autres chercheurs des EPST ou EPIC (DREM, etc.) hors post-doctorants | | |
| ETP d'autres personnels contractuels n'ayant pas d'obligation de recherche | 1 | |
| ETP de doctorants | 3 | |
| TOTAL | 6,5 | |

• **Appréciations détaillées**

L'activité de Bibliome au sein de MIG s'inscrit dans le domaine de l'extraction d'information de textes (Natural Language Processing, NLP) appliquée, en particulier, à la biologie (bioNLP). L'objectif de Bibliome est l'extraction et l'interprétation des connaissances dispersées dans la littérature biologique, en les formalisant dans des modèles logiques basés sur des ontologies du domaine des connaissances. Les techniques mises en œuvre par Bibliome proviennent de la linguistique computationnelle, l'apprentissage automatique et l'ingénierie des connaissances, une base méthodologique assez atypique dans l'ensemble des activités de l'INRA. L'intérêt de Bibliome pour l'INRA vient de l'application de ces méthodes, et des logiciels qui les intègrent, à la construction de réseaux de régulation ou d'autres relations entre objets biologiques à partir de la littérature scientifique primaire. Des exemples concrets de cette démarche sont l'extraction des interactions entre gènes de sporulation chez *Bacillus subtilis* (en collaboration avec l'équipe GenEvol), l'extraction de relations entre génotype et phénotype pour des variétés de blé (en collaboration avec des semenciers français, des instituts techniques et l'unité GDEC de l'INRA), ainsi que les relations entre biotopes et microorganismes (dans le contexte du méta-programme Méta-omiques des Ecosystèmes Microbiens).



Bibliome a réalisé une activité de structuration importante au niveau de l'institut et au niveau national et international, qui lui a donné une visibilité certaine. Par exemple, Bibliome a été moteur dans l'organisation de la compétition internationale de comparaison de méthodes d'extraction de connaissances à partir de textes, BioNLP Shared Task. La production scientifique de Bibliome est bonne pour une équipe de petite taille, mais essentiellement focalisée sur des journaux spécialisés en bioinformatique (*BMC Bioinformatics*) et des journaux et des conférences spécialisés en linguistique et en intelligence artificielle. Certes, l'équipe s'investit dans des collaborations avec des biologistes de l'INRA, mais les publications dans des revues de biologie qui témoignent de l'impact de l'application des méthodes manquent pour l'instant. Bibliome a une importante activité de développement de logiciels, notamment à travers la suite Alvis, qui regroupe les différentes méthodes de NLP développées par l'équipe. De la documentation disponible, il n'était pas facile d'avoir une vision globale des différents modules et de leurs interactions. L'impact du logiciel sur la communauté d'utilisateurs potentiels pourrait être amélioré par le développement d'un site web de haute qualité. Cependant, l'intérêt d'Alvis apparaît dans le rôle central qui lui est confié dans la section NLP de Quaero, un projet d'innovation et de collaboration industrielle financé à une hauteur importante par Oséo. Bibliome n'intervient pas dans des formations académiques régulières, mais a contribué à la formation permanente de l'INRA. Actuellement, trois thèses sont en cours.

L'activité de recherche de Bibliome à l'interface de la linguistique computationnelle, l'extraction d'information et de la biologie est originale et lui a assuré un rôle central dans le monde de bioNLP en France et, plus largement, en Europe. Les méthodes développées dans l'équipe ont un grand potentiel applicatif, mais il existe une marge de progression dans la validation des résultats obtenus par ces méthodes, à travers la participation directe à la production de nouvelles connaissances en biologie. Ceci demanderait un recentrage des efforts de l'outil informatique vers des questions biologiques, qui est difficile à réaliser dans une petite équipe, mais le recrutement récent d'un nouveau chercheur et la fusion de MIG et MIAj en cours peuvent ouvrir de nouvelles opportunités.



5 • Déroulement de la visite

Date de la visite

Début : Jeudi 18 décembre 2013 à 08h45

Fin : Jeudi 18 décembre 2013 à 18h30

Lieu de la visite

Institution : INRA, Centre de Jouy-en-Josas

Adresse : Domaine de Vilvert
78352 Jouy-en-Josas Cedex

Locaux spécifiques visités : Locaux de l'équipe MIAj et de l'équipe MIG

Déroulement ou programme de visite

08h45 : Présentation unifiée des deux unités

09h15 : Présentations scientifiques

09h45 : Discussion avec l'équipe de direction

10h15 : Rencontre avec les gestionnaires et administrateurs systèmes

10h30 : Visite des locaux MIAj

10h45 : Rencontre avec l'équipe BioSys

11h00 : Rencontre avec l'équipe GenEvol + Plateforme Migale

11h30 : Rencontre avec l'équipe Bibliome

11h45 : Pause-déjeuner

13h15 : Rencontre sur les thématiques 2 et 3 du MIAj : statistique et extraction d'informations

13h45 : Rencontre sur les thématiques 1 et 4 du MIAj : modélisation pour la biologie et l'évaluation du risque

14h15 : Rencontre avec les personnels scientifiques de rang B

14h45 : Échange avec le « groupe fusion » à propos du projet de fusion

15h15 : Rencontre avec les tutelles

15h45 : Visite des locaux MIG

16h00 : Rencontre avec les responsables d'écoles doctorales

16h30 : Rencontre avec les doctorants et post-doctorants

17h00 : Huis-clos terminal du comité d'experts

18h30 : Fin de la visite



Lieux :

08h30-12h00 : MIAj (accueil, présentations, rencontres, visite des locaux)

12h00-13h15 : MIG (repas)

13h15-15h15 : MIAj (rencontres, échange)

15h15-18h30 : MIG (visite des locaux, rencontres, huis-clos, fin de la
visite)



6 • Observations générales des tutelles

Unités de recherche UR 341 MIA-Jouy et UR 1077 MIG
INRA, Centre de Jouy-en-Josas
Domaine de Vilvert
78326 Jouy-en-Josas

Réf.: Rapport d'évaluation AERES (2014) **S2PUR150008033 - MIAJ + MIG - 0755361V**

Jouy-en-Josas, le 10 mars 2014

Réponse de la direction des unités MIA-Jouy et MIG sur le rapport AERES S2PUR150008033 - MIAJ + MIG - 0755361V, établie en accord avec les tutelles

Nous remercions très sincèrement les membres du comité AERES pour l'attention et le temps qu'ils ont consacré à l'évaluation de nos deux unités de recherche et du projet de l'unité qui résultera de notre fusion au 1^{er} janvier 2015. Nous retenons avec satisfaction l'appréciation très positive de nos activités de recherche, de l'organisation des deux unités et de notre motivation pour l'opération de fusion. Nous sommes très attentifs aux recommandations et aux points de vigilance mentionnés par la Commission.

Nous reprenons ci-dessous les points sur lesquels il nous paraît le plus important d'apporter des commentaires ou des précisions.

Préciser la politique vis-à-vis des écoles doctorales. La commission nous invite à nous recentrer sur un petit nombre d'écoles doctorales adaptées à nos thèmes de recherches. Comme expliqué dans le rapport et pendant l'audition, nous sommes déjà résolument engagés dans une stratégie centrée sur trois écoles doctorales de l'Université Paris-Saclay: Abies (agriculture, alimentation, biologie, santé), EDMH (mathématiques), STIC (informatique, traitement du signal et automatique). Cette politique sera confortée par une augmentation significative de nos HDRs dans les deux années qui viennent.

Une politique d'engagement de nouvelles thèses plus active. La commission nous encourage à une politique plus active d'engagement de nouvelles thèses (sont concernés, le thème « statistique et extraction d'information » de MIAj et l'équipe GenEvol de MIG). Cependant le nombre de thèses dans nos deux unités est loin d'être négligeable, dans un contexte le plus souvent pluridisciplinaire qui nécessite un investissement important en terme d'encadrement. Nous prendrons donc en considération cette recommandation mais en privilégiant le maintien de la qualité d'encadrement.

Présentation peu détaillée de la stratégie scientifique sur cinq ans. La structure de l'unité fusionnée et sa stratégie sont en construction et il était trop tôt pour les présenter de façon détaillée

lors de l'audition. Cependant nous sommes surpris que la commission cite le « data centre » comme « seul nouveau projet clair » et qu'elle ne mentionne pas l'IMSV qui impliquera plusieurs chercheurs des deux unités, dont le porteur du projet.

Des interactions avec les statisticiens de l'unité renforceraient les activités de GenEvol. Nous voulons clarifier ce point. La commission écrit que « *Un développement des interactions avec les statisticiens de l'unité permettrait de renforcer les activités du thème GenEvol* ». Des chercheurs et ingénieurs statisticiens sont de fait présents dans les trois équipes de MIAj et dans l'équipe GenEvol. Ils interagissent déjà et sont tous impliqués dans des recherches en méthodologie statistique et dans des projets appliqués. Nous cherchons bien entendu à favoriser de nouvelles interactions et synergies, mais la fusion ne résoudra pas le besoin de recrutements sur ce thème très sollicité et très compétitif, qui va être affecté par plusieurs départs à la retraite.

L'axe « Biologie du développement » présente des résultats assez modestes. Ce thème, d'une grande importance pour l'unité fusionnée, mérite d'être renforcé. Pour cela, nous partageons l'avis de la commission qu'une politique active de publications est nécessaire. Celle-ci a été vigoureusement engagée depuis que le responsable du thème a transmis ses responsabilités de DU de MIAj fin 2012. Ces efforts ont porté leurs fruits peu après la visite de la commission : un article a été accepté dans *Plant Physiology*, un second a été soumis à *Science Report*, et une thèse vient d'être soutenue le 5 mars. D'autres publications sont en cours de rédaction.

Le projet de « data centre » est porté principalement par la plate-forme MIGALE. Il nous faut préciser que le projet de data centre n'est pas porté par la plate-forme MIGALE, mais par les centres INRA de Jouy et Versailles. Néanmoins nos deux unités sont bien concernées par ce projet et impliquées dans son portage. Le data centre a pour objectif principal la mutualisation des infrastructures informatiques des unités de la région parisienne sur deux plans, la gestion matérielle, (par la mise à disposition de locaux adaptés en termes de puissance électrique, de réseaux et de climatisation), et le fonctionnement en assurant une continuité de service de l'administration des machines). Au moins au début, les services offerts à la communauté des sciences de la vie seront ceux déjà proposés par les plates-formes existantes. Si certains aspects de ces offres de service sont communs, ils seront mutualisés. MIGALE jouera un rôle clé dans ces évolutions. Par ailleurs, l'unité fusionnée a vocation à s'appuyer sur ce data centre pour diffuser des méthodes d'exploitation des données et développer ses activités de modélisation sur le centre et au-delà.

Le thème BioSys devra veiller à ne pas perdre son focus et disperser ses forces. Nous comprenons les interrogations de la commission et nous regrettons de ne pas avoir eu le temps d'aborder cette question avec celle-ci, étant donné le calendrier chargé de la journée. Au regard des développements méthodologiques que nous avons réalisés autour des bactéries, ainsi que des résultats obtenus, il nous a semblé impératif de transférer « au plus vite » ces méthodologies et approches vers les systèmes biologiques multicellulaires, et spécifiquement les plantes. Ceci implique évidemment qu'à effectif constant, les efforts de recherche sur les bactéries soient réduits (et leur continuation dépendra en grande partie des moyens que voudra y consacrer le département MICA). Cette orientation vers les systèmes eucaryotes et multicellulaires nécessite bien

évidemment une stratégie non seulement au regard de l'extrême complexité de ces systèmes, mais aussi de l'existence d'une communauté travaillant dans le champ de la modélisation des plantes.

Cette nécessité scientifique et les opportunités ouvertes par la constitution de la future Université Paris-Saclay, nous ont permis de construire avec les communautés concernées, des projets autour de la modélisation des bactéries et des plantes (Institut de Modélisation des Systèmes Vivants, porté par V. Fromion). Ce projet interdisciplinaire, financé et soutenu par l'Idex, nous permet déjà, au regard de la qualité du partenariat, de déployer dès aujourd'hui nos méthodes et nos outils sur des projets ciblés et ambitieux appelant les méthodes de prédiction d'allocation des ressources pour le vivant (RBA) et la modélisation des réseaux de régulations.

Le thème Bibliome a intérêt à un recentrage des efforts de l'outil informatique vers des questions biologiques qui est difficile à réaliser dans une petite équipe Les activités de recherche méthodologiques de Bibliome sont toutes dirigées par des questions biologiques précises qui sont formalisées et analysées dans le cadre de collaborations durables, telles que l'analyse des régulations génétiques chez les bactéries et les plantes, la découverte de marqueurs génétiques du blé, l'étude de l'adaptation des microorganismes à leur biotope. Les résultats méthodologiques implémentés dans des outils informatiques génériques ont bien vocation à contribuer à la production de nouvelles connaissances en biologie.

L'ouverture de ces outils à une communauté plus large d'utilisateurs que les partenaires de Bibliome relève d'une activité de service hors de la mission de recherche de Bibliome. Leur mise à disposition grâce à des plateformes bioinformatiques telle que la plateforme Migale est à l'étude. Il est à noter que leur exploitation dans des applications spécifiques nécessitera des compétences informatiques additionnelles en ingénierie des connaissances.



Hervé Monod,
directeur de l'unité MIAj



Sophie Schbath,
directrice de l'unité MIG

Frédéric GARCIA
Chef de Département MIA

