

Compte-rendu du Séminaire de restitution de l'étude sur l'utilisation des données de séquençage de ressources génétiques pour l'agriculture et l'alimentation

Lundi 8 octobre 2018 14h00-18h00 à la Maison des Océans



Le séminaire a réuni une quarantaine de personnes et 9 intervenants.

Toutes les présentations sont disponibles.

Ouverture : Jean Lanotte au ministère de l'Agriculture et de l'alimentation a souligné le caractère conflictuel des négociations autour de la DSI (« *Digital Sequence Information* ») au sein des instances internationales (Convention sur la diversité biologique, Traité international sur les ressources phytogénétiques, Convention sur le droit de la mer). Ces instances débattent souvent des DSI de façon abstraite ou théorique, l'idée est que cette étude fasse remonter des pratiques du terrain après avoir clarifié ce que recouvre le terme.

Présentation de l'étude par Charlotte Navarro : Le contexte de la commande de l'étude a été rappelé. L'étude a débuté en juin et se terminera en novembre 2018. L'objectif est de réaliser un état des lieux sur l'utilisation des données de séquençage afin de donner un éclairage sur ce concept. L'étude repose sur une enquête auprès des acteurs/utilisateurs de données de séquençage dans le domaine public et privé. Des cas d'utilisation de données de séquençage de ressources phytogénétiques, animales, forestières microbiennes, d'invertébrés et aquatiques ont été illustrés à travers des exemples.

Des questions ont émergé sur la terminologie à adopter au niveau international et dans le cadre des négociations sur la haute mer, sur la dépendance entre la ressource biologique et les données de séquençage associées, et sur la propriété intellectuelle (PI) qui peut découler de travaux réalisés sur les données de séquençage. A la suite du séminaire, le Comité de pilotage a été sollicité pour émettre un avis sur la terminologie à adopter.

Des interventions ont été l'occasion d'illustrer des cas d'utilisation de données de séquençage :

Jean Christophe Glaszmann, chercheur au Cirad dans l'unité AGAP, a présenté une initiative internationale sur le riz autour du Groupe consultatif pour la recherche agricole internationale (GCRAI). Le riz est la première plante cultivée à avoir eu son génome séquencé, en 2004, car il est simple. Comme première étape, 3000 variétés ont été reséquencées. Plusieurs dizaines de millions de marqueurs moléculaires de type SNP (*Single Nucleotide Polymorphism*) permettent une caractérisation très fine de la structure de la diversité le long du génome. On observe également une forte variation du contenu en gènes. Seules 60% des familles de gènes connues chez le riz sont communes à toutes les variétés ; les autres sont présentes, ou absentes, selon les variétés et constituent des sources potentielles d'adaptation. Les caractérisations phénotypiques sont en cours afin de repérer des variations au niveau de gènes importants sur le plan fonctionnel. Ce type d'approche est en voie de généralisation avec l'initiative DIVSEEK, mais une très faible participation des pays en développement est constatée, soulignant l'importance de la formation dans ce domaine.

Nicolas Langlade, directeur de recherche au Laboratoire Interactions Plantes Microorganismes (UMR CNRS/Inra), a souligné la place essentielle des bioinformaticiens dans le projet SUNRISE pour le traitement des données de séquençage. La première étape consiste à générer une séquence de référence pour le tournesol cultivé qui présente un génome complexe de 3,6 Gb, grâce à la technologie très récente PAC Bio. La deuxième étape est l'identification des gènes et la réalisation d'un atlas d'expression de gènes. Enfin il est possible de reconstituer les voies métaboliques pour produire des caractères d'intérêt (exemple : production d'acides gras). La description de la diversité est permise par le reséquençage de lignées pour y découvrir du polymorphisme par rapport à une séquence de référence. Un autre aspect du projet est l'étude d'espèces sauvages pour générer des séquences de référence nouvelles et distinctes. Les problèmes qui se posent : l'annotation, le mapping (positionnement des gènes sur un génome de

référence), l'obsolescence des références. Le projet Phénomène a pour but de caractériser les phénotypes associés aux caractères d'intérêt économique.

Frédéric Gaucheron est le responsable du service « Science et Technologie du Lait » au sein du Centre national de l'interprofession et de l'économie laitière (CNIEL). Les données de séquençage sont mobilisées dans des projets précompétitifs pour l'amélioration des produits issus de la transformation laitière. Le but est d'améliorer la compréhension des écosystèmes microbiens, leur diversité et leurs fonctionnalités pour différents intérêts (connaissance, transfert d'outils, suivi d'écosystèmes, constituer des banques de référence, etc.). Les caractéristiques du fromage (couleur, acidification, texture, flaveur, etc.) sont générées par ces écosystèmes et contribuent à leur qualité. Un projet abordé a pour objectif d'établir un catalogue des communautés microbiennes présentes dans l'ensemble des fromages AOP français.

Michèle Tixier-Boichard, directrice de recherche à l'Inra en génétique animale, spécialiste en génétique avicole, a présenté le projet Biodiva qui s'intéresse à la diversité génétique des races avicoles locales en vue de leur protection et de leur valorisation pour l'évolution de l'aviculture. Il a permis de caractériser la diversité génétique des races françaises à l'aide de marqueurs haute densité et de la comparer à celle des races commerciales. Des accords de transfert de matériel ont été signés entre l'Inra et chaque fournisseur d'échantillons de races anciennes pour, entre autres, définir les conditions d'échanges, de publications, d'utilisation des données, d'utilisation ultérieure des échantillons, notamment pour du séquençage.

Philippe Vandenkoornhuysse, professeur à l'Université de Rennes et docteur en écologie moléculaire et en écologie de l'évolution, a présenté le concept d'holobionte qui repose sur une perception écosystémique. Le concept d'holobionte correspond à la plante et son cortège symbiotique (champignons, bactéries, etc.). Il souligne l'importance de la symbiose mycorhizienne pour la croissance et la survie des plantes. Les évolutions en génomique, métagénomique, etc. permettent de développer les connaissances autour de ces symbiontes.

[Des interventions ont porté sur les questions de gestion des données :](#)

Hadi Quesneville, directeur de l'unité Inra URGI de bioinformatique, a présenté son laboratoire permettant la gestion et le stockage des données de recherches agronomiques portant sur les plantes. L'URGI s'intéresse aux données de génomique, de génétique et de phénotype. Elles sont stockées et associées entre elles et ensuite intégrées dans un système d'information permettant d'apporter de la valeur à ces données en les faisant « parler ». Le défi scientifique, dans lequel s'inscrivent les activités de l'URGI, est de comprendre le fonctionnement des plantes (génomique dans son environnement avec des pratiques culturales et association avec les traits phénotypiques), notamment par un travail d'identification de gènes responsables de traits via des marqueurs communs. Ces marqueurs sont représentés sous la forme de variations nucléotidiques « SNP ». Les indications de polymorphismes « SNP » sont stockées dans les bases de données. L'URGI offre des outils d'analyse des séquences génomiques (BLAST, Browse, etc.).

Anne-Françoise Adam-Blondon, chercheuse Inra responsable scientifique de la plateforme URGI, présente « l'écosystème de systèmes d'information » pour l'échange de données qui tend à s'inscrire dans l'Open science. Elle présente les infrastructures de production de données (France Génomique, CRB, etc.), les infrastructures d'archives internationales (GenBank, etc.) et les systèmes d'information permettant de traiter ces données (GnplS, Crop Ontology, etc.) pour une meilleure analyse. Le déluge des données provoqué par l'amélioration des techniques de séquençage déstabilisent les infrastructures internationales d'archivage. Aujourd'hui, on assiste à l'émergence de fédérations de systèmes d'informations pour améliorer l'accès au foisonnement de systèmes d'information distribués. Ces fédérations reposent sur les principes de FAIR, « *Findable, Accessible, Interoperable, Reusable* » (« trouvable, accessible, interopérable, réutilisable »).

Samuel Mondy, Ingénieur de Recherche Inra en Bioinformatique et responsable de la plateforme GenoSol, présente les travaux sur le sol qualifié d'environnement complexe. Il rappelle le contexte de

nombreuses applications de l'utilisation de données de séquençage (étude de biodiversité, d'impact), de l'évolution rapide des outils d'analyse d'interprétation et le contexte de l'Open science (impact sur les relations avec les partenaires). La plateforme GenoSol a pour objectif la mesure de l'impact de l'homme (agriculture, réchauffement, etc.) sur la capacité des sols à rendre un service écosystémique. Le devenir des données de séquençage issues de la plateforme est questionné. Comment les mettre à disposition pour leur réutilisation ?

Table ronde :

La table ronde a réuni Jean-Louis Pham (IRD), Edwige Pilard (Genoplante-Valor) Odile Hologne (Délégation Information scientifique et technique, Inra), Anne-Francoise Adam-Blondon (URGV Inra), Michèle Tixier-Boichard (Génétique animale et biologie intégrative Inra), Thierry de Meeûs (IRD), Olivier Lucas (RAGT), Jean-Pierre Lebrun (Confédération paysanne) et Alexandrine Rey (Cirad).

Les intervenants se sont présentés et ont été invités à s'exprimer sur une série de questions. De nouveaux thèmes de débat ont émergé et ont laissé place à des discussions avec le public.

Quels sont les impacts de l'utilisation des données de séquençage (techniques, politiques, économiques, etc.) ? De nombreux points ont été abordés :

- Il a été rappelé que l'accès aux séquences a été une révolution dans les années 1980¹ pour accroître les connaissances sur le vivant de manière extraordinaire et dans une logique d'universalisme de la connaissance de la biodiversité. Les données de séquençage ont d'abord permis l'étude descriptive de la diversité, sans finalités « appliquées » (phylogénie, biologie de l'évolution) puis la génomique a permis l'analyse de caractères d'intérêt et d'affiner les outils de sélection.
- Sur le plan politique, les tensions observées dans les arènes internationales traduisent que l'analyse de données de séquençage apporte beaucoup de connaissances nouvelles.
- Sur le plan technique, les technologies évoluent très vite et les coûts d'analyse, de gestion et de stockage de la donnée deviennent plus importants que le coût de production. Il n'existe pas de modèle économique pour envisager de répondre à ce problème.
- Sur les mentalités, notamment, celles des acteurs en sélection animale, l'évolution des technologies a permis un changement de culture, où l'intérêt privé prime sur la mutualisation. Auparavant, la mutualisation était nécessaire car il était difficile de produire des données. A l'inverse, la sélection avicole a toujours été, elle, confidentielle car il n'existe pas de réglementation nationale. La production facilitée de données génomiques entraîne une plus grande réticence des acteurs en sélection bovine à partager leurs données pour un souci de concurrence.
- Sur le plan juridique, la gestion des contrats de collaboration de recherche est de plus en plus difficile car elle nécessite de canaliser la propriété sur les grandes bases de données et de prendre en compte les droits de producteurs de données.
- Sur le plan de la brevetabilité, le développement de la recherche en génomique permise par l'utilisation des données de séquençage a un impact fort. Par ailleurs, même si ce n'est pas l'objet du débat, le droit des brevets issu de la directive sur la protection des inventions biotechnologiques semble contribuer au malaise car il est encore très accueillant vis-à-vis des gènes natifs en Europe. Des inquiétudes ont émergé concernant la brevetabilité du vivant et la propriété intellectuelle (PI) permises par l'accès aux bases de données et à tout un travail fait par des générations de paysans.
- Pour limiter ces craintes, il a été proposé de travailler sur le périmètre de la brevetabilité du vivant au niveau européen, à l'image de la France qui exclut la brevetabilité de l'information génétique

¹ En 1980: accès à des portions de séquence, depuis 1990 et surtout les années 2000, accès à des génomes entiers.

native². Des questions ont émergé sur les interactions entre les dispositions spécifiques sur la PI au niveau de la CDB, du protocole de Nagoya et l'OMPI³.

- Sur le statut des données de séquençage, un parallèle a été fait avec un colloque ayant eu lieu en septembre 2018 à l'Institut de recherche en propriété intellectuelle sur la circulation des données numérique et la PI. Il a été dit que la donnée de séquençage ne devait pas être considérée différemment des autres données.
- Il a été suggéré de raisonner en termes d'accès en utilisant les systèmes qui existent et les faire fonctionner entre eux. La réglementation autour des accès doit veiller à l'interopérabilité des systèmes juridiques qui existent au plan national, international, et les différents droits : droits des brevets, des bases de données, la transposition de la directive sur les droits des affaires (des entreprises privées qui produisent des données sur leurs fonds propres peuvent protéger leurs données par le secret).
- Il a été dit qu'il est délicat de réfléchir aux mécanismes d'accès et de partage des avantages indépendamment du périmètre des droits de propriété intellectuelle.

Quelles seraient les conséquences d'une réglementation sur l'accès aux données de séquençage dans vos activités (spécifiques/générales) ?

- Il a été dit que l'accès serait complexifié, qu'on risquerait de glisser dans une période de « glaciation » dans le domaine de la recherche, qui résulterait des raidissements des positions au niveau international.
- Ce raidissement pourrait également résulter de la non-prise en compte des revendications sur l'APA pour les DSI.
- Au niveau juridique, les contrats seraient plus nombreux, mais il a été dit qu'il était nécessaire de réglementer pour ne pas menacer les partenariats avec les établissements de recherche locaux dans les pays du Sud. Dans le cadre de partenariats publics-privés impliquant des pays du Sud, les partenariats que tissent certains instituts de recherche avec des pays du sud sont de rares exemples où un partage des avantages est réellement mis en place.
- Il me semble important, au niveau international, d'apporter du soutien aux fournisseurs de ressources génétiques en reconnaissant le besoin d'une réflexion commune sur les données, notamment dans le but d'éviter l'adoption de lois qui en limiteraient trop fortement l'accès. L'obligation d'obtenir un accord sur l'utilisation des données au moment de l'accès à la ressource pourrait constituer un premier pas. En tout cas, une position trop radicale sur l'exclusion des DSI du champ de l'APA ou se retranchant sur l'absence de définition de la notion, pourrait aviver les tensions et fragiliser encore l'accès aux ressources et aux informations.
- Il a été rappelé qu'il existe déjà un encadrement de l'accès aux données. Des modalités différentes d'accès peuvent exister au sein d'un même projet, impliquant l'accord du fournisseur de la ressource ou d'un consortium précédent.

Comment imaginer un partage des avantages qui tienne compte des différents investissements et travaux fournis ?

- Il a été rappelé que la donnée brute a très peu de valeur, qu'elle en acquiert à partir du moment où elle est curée, analysée et accompagnée de métadonnées.
- Il a été proposé que le partage des avantages se gère au moment de l'accès à la ressource, car l'accès à la ressource génétique est essentiel, même en biologie de syn. Il est nécessaire de faire le lien entre les ressources génétiques, les données de séquençage et les données de phénotypage, car c'est à partir de l'accumulation de ces deux types de données que l'on peut caractériser une ressource, montrer son intérêt et ainsi lui donner la visibilité et de la valeur. Cette étape constitue déjà une forme de partage des avantages.

² Un gène natif est un gène naturel qui n'a fait l'objet d'aucune modification biotechnologique. Depuis 2016, un amendement interdisant le brevetage des « produits issus de procédés essentiellement biologiques » a été adopté en France.

³ Créée en 1967, l'OMPI est l'Organisation mondiale de la propriété intellectuelle chargée d'élaborer un système international de propriété intellectuelle.

- De la même façon, le partage des avantages sous forme de développement de capacités est trop souvent fait a posteriori alors qu'il devrait être prévu en amont. Cela pourrait se traduire par le développement de partenariats équitables et donc la co-construction des programmes de recherche entre détenteurs et utilisateurs de ressources génétiques. C'est un prérequis pour aborder l'accès aux bases de données de séquençage. L'open access ne peut suffire comme partage des avantages dès lors qu'il n'y a pas d'équité dans les capacités d'utilisation des données.
- Une réglementation basée sur le droit du producteur de bases de données a été imaginée, où, à partir d'une extraction substantielle, un partage pourrait se mettre en place avec le fournisseur de la ressource.
- Une solution mutualisée au niveau international sur les données avec un réel transfert de technologie et de renforcement des capacités a été proposée, un peu à l'image du modèle du TIRPAA, et permettrait de faciliter le travail juridique. Il a été rappelé que les détenteurs de races animales ne sont pas prêts à intégrer un accord type TIRPAA pour leurs ressources biologiques, mais que tous voient un intérêt à la mutualisation de leurs données tant que cet accès libre partagé permet un retour significatif d'informations utiles pour chacun.

Des réflexions ont été exprimées sur les formes possibles de renforcement de capacités et de transfert de technologie :

- Des chercheurs ont rappelé qu'il était *de facto* obligatoire dans les contrats établis par l'IRD, de faire participer les partenaires, de les former et de les inclure dans les publications, dans une logique de diffusion de savoir.
- Le ministère de l'Agriculture et de l'alimentation a rappelé l'importance de la question du partage des avantages non-monétaires.
- Cette forme de partage des avantages, notamment le renforcement des capacités individuelles et institutionnelles en analyse de données (bioinformatiques) pourrait faire l'objet d'une initiative forte des ministères. Des instituts de recherche comme l'IRD et le Cirad ont des expériences à partager en matière de renforcement de capacités au Sud.
- Dans le cadre du développement de la génomique et des nouvelles techniques de séquençage, des professionnels ont souligné l'importance de partager les outils d'analyse des données.
- Le renforcement des capacités doit viser les étudiants et les communautés de chercheurs, mais aussi les communautés de producteurs pour veiller à l'équilibre en termes de capacités de négociation.

Anca Leroy, point focal Nagoya au ministère de la transition écologique et solidaire, a souhaité apporter quelques précisions :

- Les ressources phytogénétiques sont traitées de manière différente dans le cadre de la FAO, du protocole de Nagoya, du règlement communautaire européen et de la loi pour la biodiversité de 2016.
- Au niveau international, la France s'exprime à travers la présidence tournante du Conseil de l'Union européenne, actuellement l'Autriche, et la Commission européenne. La France sera bien présente à la COP en Egypte.
- La position de la France pour la prochaine COP en Egypte est très claire : le champ d'application du protocole de Nagoya ne couvre pas la DSI.
- Le règlement communautaire européen ne couvre pas la due diligence pour la DSI, car il n'y a pas de définition consensuelle au niveau international donc on ne peut pas discuter d'un élargissement du protocole de Nagoya.
- Au niveau international, l'enjeu primordial sur la DSI concerne moins la CDB et le protocole de Nagoya, que les négociations sur le futur accord concernant le droit de la mer au-delà des juridictions nationales où il y a des enjeux majeurs.
-

Clôture : Jean Lanotte au ministère de l'Agriculture et de l'alimentation a remercié l'ensemble des participants et a précisé que l'étude va permettre d'alimenter les réflexions du ministère, mais qu'il y a encore des sujets à approfondir. L'étude devrait être largement diffusée.