



Séminaire de restitution

Etude sur l'utilisation de données de séquençage de ressources génétiques pour l'agriculture et l'alimentation

Fondation pour la recherche sur la biodiversité

Charlotte Navarro
Chargée d'étude

Membres
Fondateurs
de la FRB:





Restitution de l'Etude sur l'utilisation de données de séquençage de ressources génétiques pour l'agriculture et l'alimentation

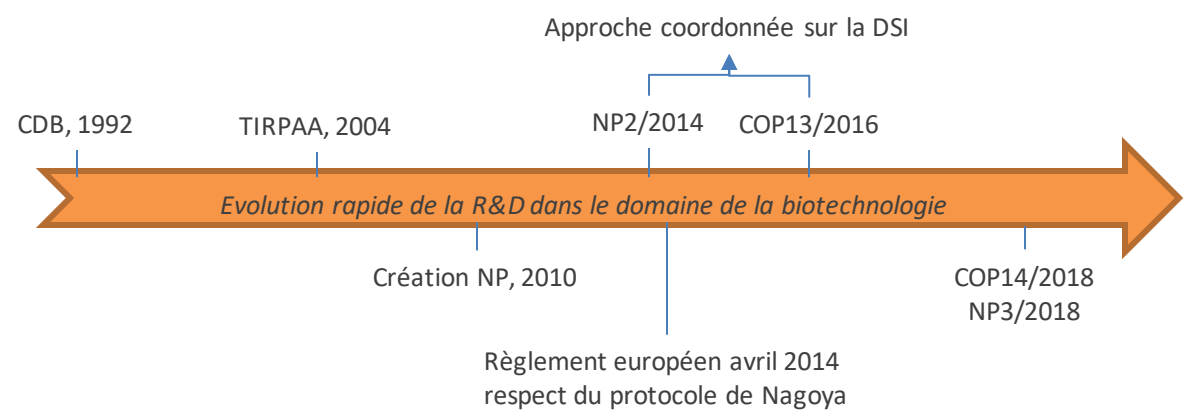
Plan de restitution:

1. Contexte de l'étude
2. L'objet de cette étude
3. Les utilisations



1. Contexte de l'étude

- Enjeux et actualités
- La commande de l'étude
- La méthodologie de l'étude





1. Contexte de l'étude

Définition dans l'Article 2 de la Convention sur la diversité biologique.

- **Les ressources génétiques** font référence au matériel génétique ayant une valeur effective ou potentielle. Elles contiennent le matériel génétique, d'origine végétale, animale, microbienne ou autre, contenant des unités fonctionnelles de l'hérédité. Les ressources génétiques peuvent être in situ (en collections) ou ex situ (sur le terrain). **Les ressources « in silico » ne sont pas comprises dans cette définition.**



Notions essentielles

De la biologie moléculaire à la génomique

Le génome

Les gènes

Codage chimique : A, T, C, G

Séquence : succession des éléments de base.

Séquençage: permet de déterminer la séquence d'une molécule d'ADN

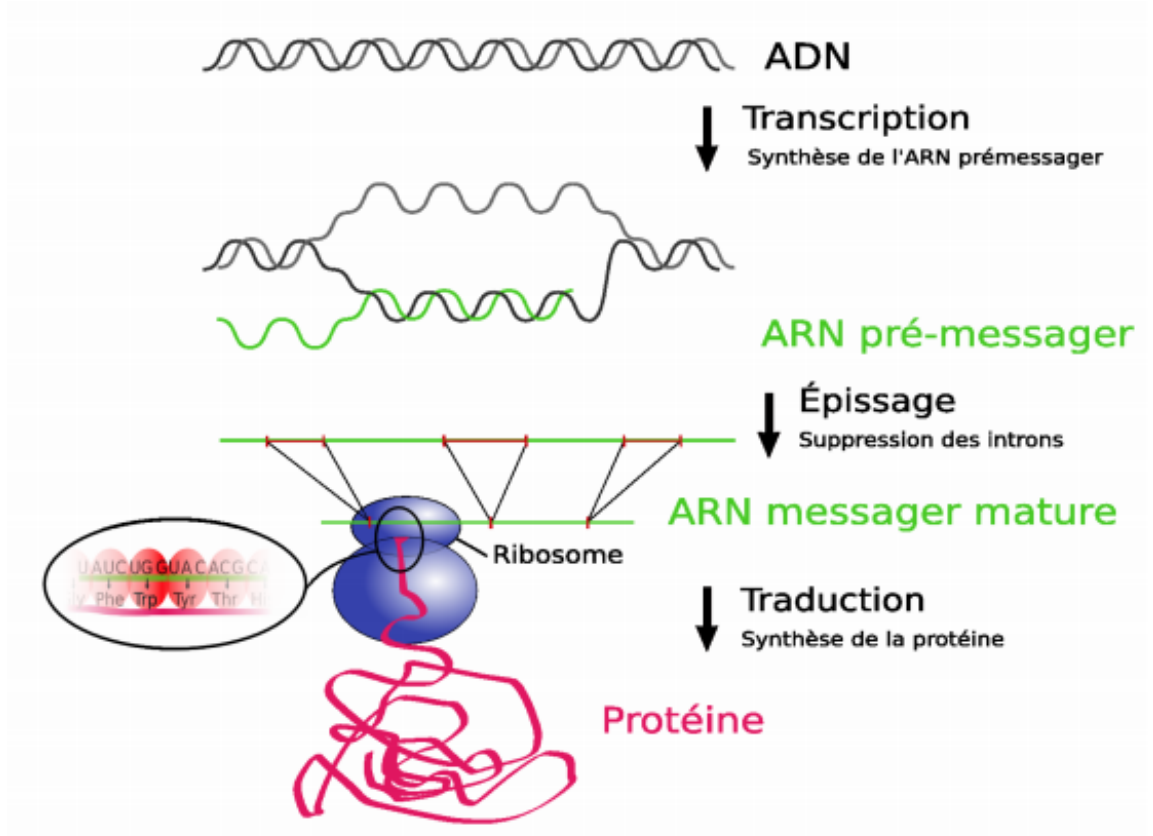


FIGURE 1.1 – Le *dogme central* de la biologie moléculaire
 Source: Peterlongo, 2016



Evolution de la génomique

(Source: Christine Gaspin, directrice de recherche à l'UMR MIAT Inra)

1955 : séquençage de la première protéine.

1965 : premier atlas de séquences protéiques (accès aux fonctionnalités).

1970 : premières méthodes reposant sur des algorithmes pour comparer les séquences.

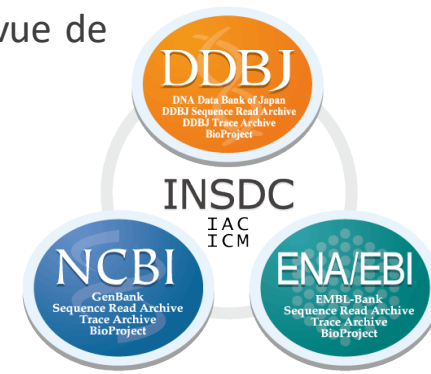
1974 : Sanger met au point le séquençage de l'ADN (accès à l'ensemble du génome).

1988 : Création du réseau collaboratif international EMBnet et début de l'utilisation d'internet par la recherche au niveau mondial pour l'échange des données.

1990 : Lancement du projet « Génome humain » et naissance du logiciel BLAST (très largement utilisé). Mise en place de bases de données qui s'échangent leurs données : collaboration internationale et mise à jour quotidienne. Attribution d'un identifiant (« DOI » aujourd'hui) pour identifier les séquences lors du dépôt de séquences dans une base de données.

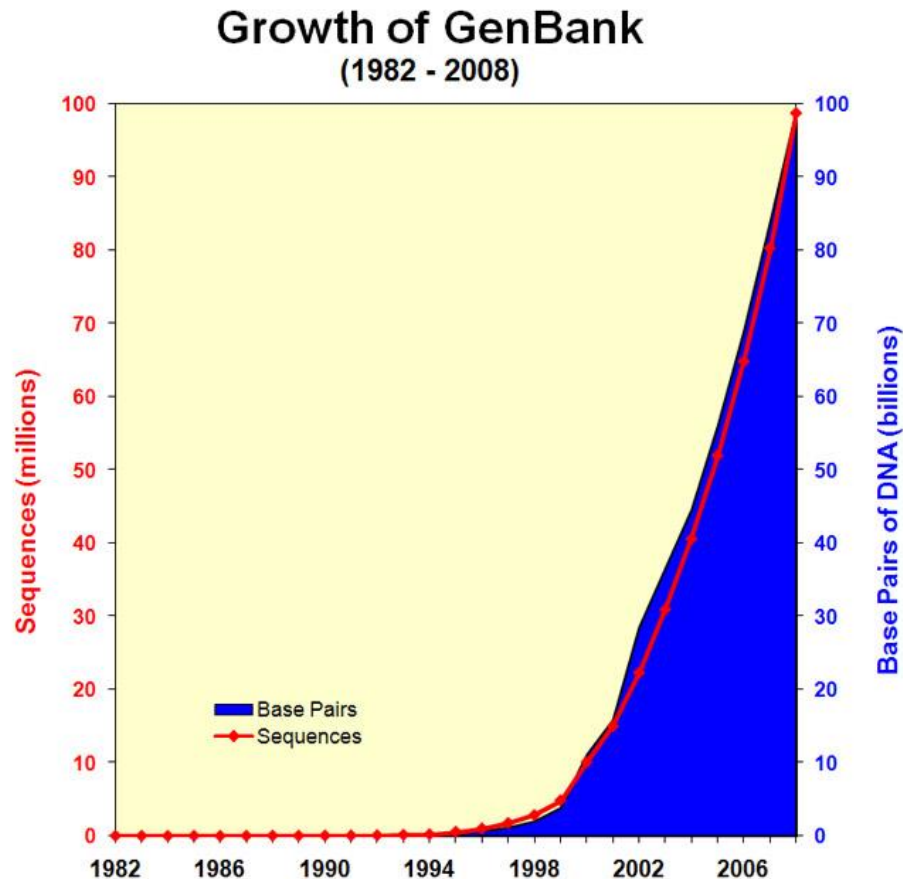
Dès les années 1990 : production collective et partage des données biologiques en vue de l'exploitation dans le cadre de consortiums internationaux.

Aujourd'hui: Le séquençage des génomes se démocratise.





Croissance du nombre de séquence stockée dans GenBank de sa naissance à 2008, (NCBI)



En décembre 2017, *GeneBank* stockait 206 millions de séquences correspondant à 250 milliards de bases, <https://www.ncbi.nlm.nih.gov/genbank/statistics/>).



Le réseau national de pôles de génomique

- France génomique
- Programme Investissement d'avenir
- CEA, Inra, CNRS, INSERM, Institut Pasteur, Fondation Paris Sciences & Lettres, CERBM-GIE
- Genoscope d'Evry, GenoToul, MGX-Montpellier GenomiX, GenOuest
- GIS IBISA, 2007: labellisation et soutien aux plateformes et infrastructures



Les plateformes de France en génomique, www.france-genomique.org



2. L'objet de cette étude

Les données de séquençage,
« Digital Sequence Information »

- Pas de définition scientifique, « DSI » terme de référence;
- Désaccords sur le caractère matériel ou non, indépendant de la ressource génétique ou non
- La DSI, un produit de la recherche, une représentation intellectuelle de la ressource génétique
- Terme polysémique
- Terme évolutif
- **De quoi parle-t-on?**

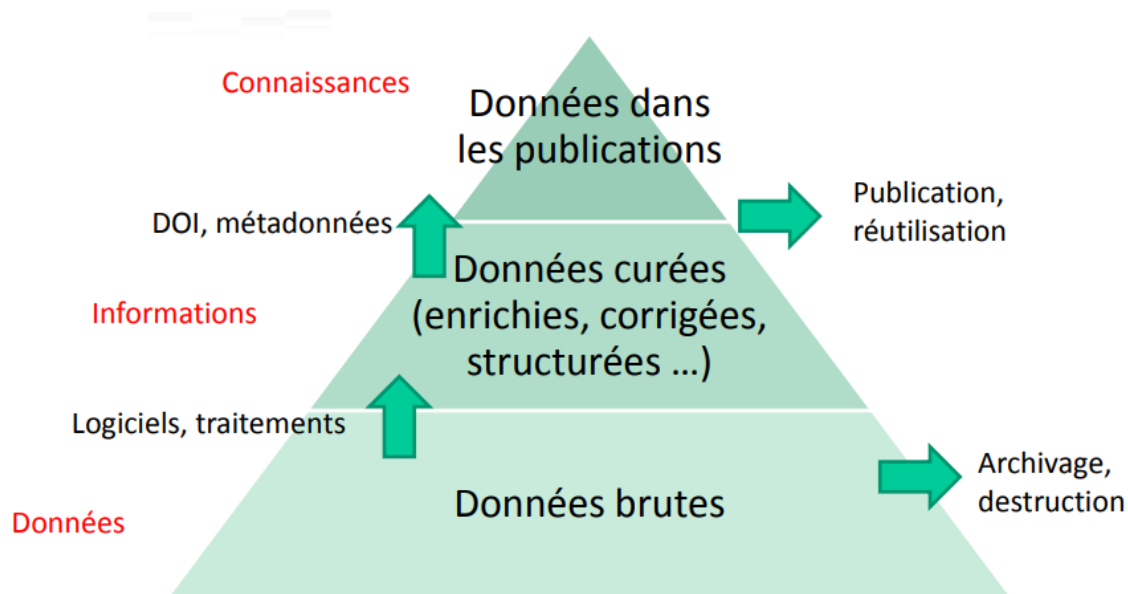


2. L'objet de cette étude: typologies proposées

- Chainage des programmes ou Protocole bioinformatique (« pipeline »)

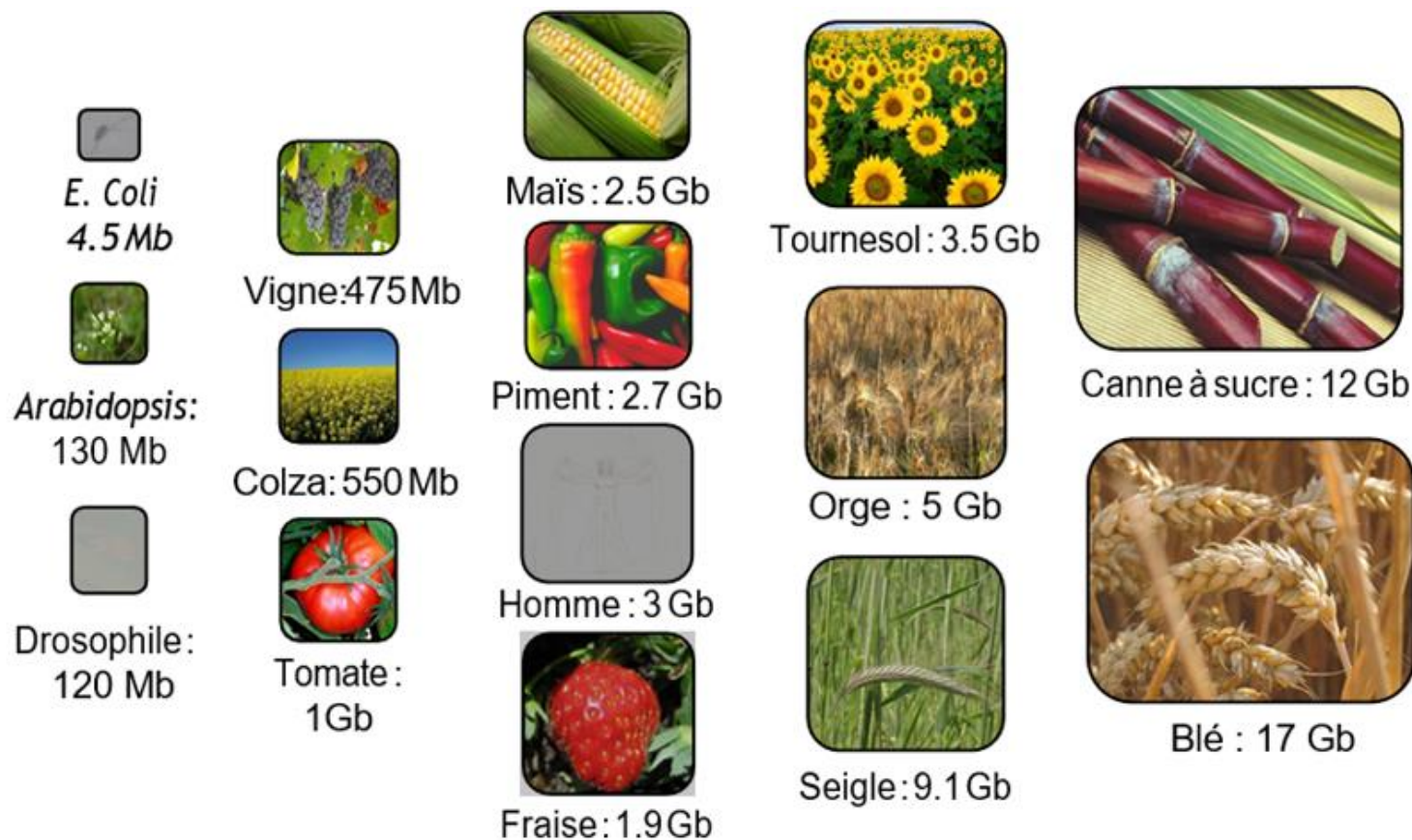


- Données brutes/données nettoyées/données assemblées



Les données de séquençage de ressources génétiques pour l'agriculture et l'alimentation

- Les types de ressources génétiques considérées : végétales, animales, aquatiques, forestières, de microorganismes et d'invertébrés.
- La taille du génome influe sur le séquençage
- Les espèces d'intérêt économique bénéficient davantage des programmes de génomique





Utilisations du séquençage du génome des ressources génétiques

- Caractérisation des RG
- Analyse de la diversité allélique
- Recherche des allèles les plus efficaces/intéressants
- Détermination et Identification des gènes présents dans les zones du génome (cartographie génétique)
- Sélection assistée par marqueurs : choix des géniteurs, tri des descendants obtenus
- Mettre au point des outils (puces à ADN) capables de suivre l'activation de gènes selon certaines conditions
- Clonage plus facile des gènes pour des possibilités de transformations génétiques mieux ciblées, plus efficaces utilisant des gènes spécifiques ou cherchant à éteindre ou désactiver l'expression de certains gènes indésirables
- ...



L'utilisation de données de séquençage de ressources zoogénétiques :

- L'exemple de BiodivA: pour la conservation de races traditionnelles, locales et/ou anciennes françaises. Présentation par Michèle Tixier-Boichard, Inra
- Vache laitière (2009) : outils de génotypage pour identifier les gènes responsables de la baisse de fertilité des troupeaux
- Séquençage du génome du porc en 2012 : modèle pour les maladies humaines
- Séquençage du mouton Texel : amélioration génétique du développement musculaire





L'utilisation de données de séquençage de ressources phytogénétiques

- Espèce modèle « arabette » (*Arabidopsis Thaliana*) un génome de petite taille qui facilite les manipulations en laboratoires: caractérisation des gènes (structure, fonction, régulation)





Le projet Genius, « Ingénierie cellulaire : amélioration et innovation technologiques pour les plantes d'une agriculture durable » : outils pour une modification ciblée des caractères agronomiques

- 12 espèces : 9 cultivées, 3 modèles
- Réduction des intrants, adaptation au changement climatique, utilisation de la biomasse végétale
- Modifications possibles: perte de fonction d'un gène, modification d'un gène, insertion ciblée d'un gène
- Techniques: modifications ciblées des gènes (méganucléases, TALENs, CRISPR-Cas9)



Vers des plantes cultivées présentant des caractères de culture et de qualité améliorés pour l'alimentation humaine et animale et d'autres utilisations, Peter Rogowsky

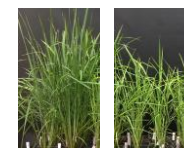


Exemples de recherche dans le cadre du projet Genius

- Outils de sélection: **Maïs**, obtention de gamètes diploïdes → pour la propagation asexuée des cultures
- Qualité des produits: **Pomme de terre**, production d'un amidon composé uniquement d'amylopectine → additif pour l'industrie alimentaire et de la colle
- Temps de floraison: **Pomme**, floraison très précoce → pour un cycle de vie raccourci et l'adaptation au changement climatique
- Adaptation au stress abiotique: **Riz**, tolérance à la salinité → pour la culture sur des terres marginales et l'adaptation au changement climatique
- Résistance aux maladies: **Tomate**, résistance au potyvirus → pour la protection des plantes et la réduction de pesticides



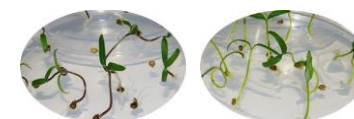
Apple



Rice, wheat



Poplar



Tomato

Vers des plantes cultivées présentant des caractères de culture et de qualité améliorés pour l'alimentation humaine et animale et d'autres utilisations,

Peter Rogowsky



Le peuplier noir, une espèce d'intérêt économique et écologique

- Séquençage du génome du premier arbre : *Populus trichocarpa* en 2006 et croissance des ressources génomiques
- Caractériser la diversité génétique notamment pour la conservation
- Intégration de l'information génomique dans le programme d'amélioration





L'utilisation de données de séquençage de ressources génétiques d'invertébrés

- Dès 1994, cartographie génétique de l'Abeille
- Situation catastrophique: perte d'adaptation des souches locales et augmentation de la consommation de sucre en France; syndrome d'effondrement des colonies d'abeilles
- Projet BEEHOPE: gestion durable de l'apiculture, écologie de l'Abeille noire, étude de l'adaptation au changement climatique
- Apport du séquençage: évaluation de la diversité génétique (marqueurs moléculaires); séquençage de génomes entiers
- Conservation: mise en place de conservatoires génétiques d'abeilles d'origine naturelle





L'utilisation de données de séquençage de ressources génétiques aquatiques

- Importance de la conchyliculture européenne (moules, huîtres, palourdes)
- Mortalité due à des virus, bactéries, parasites
- Projet Vivaldi: prévenir et contrôler les maladies des bivalves marins
- Apport du séquençage: Sélection génétique pour la résistance/tolérance aux maladies et compréhension des interactions entre coquillages – environnement – pathogènes
- Projet transfrontalier, constitution d'un réseau international





Conclusion

- Intérêts du séquençage du génome de ressources génétiques pour l'agriculture et l'alimentation
- L'évolution des techniques de séquençage concerne d'autres domaines: santé, cosmétique, ...
- La donnée de séquençage : différents états et utilisations
- Des problèmes soulevés: le traitement, stockage des données, les statut des données
- Débat autour de l'utilisation des nouvelles techniques de selection variétale



Merci pour votre attention

Prochains événements autour de la DSI:

- Novembre 2018: COP 14 en Egypte
- Janvier 2019 : FAO/CRGAA



À suivre, les présentations de projets de génomique de ressources génétiques pour l'agriculture et l'alimentation et les initiatives pour la gestion des données de séquençage





Les techniques de séquençage

- Le séquençage consiste à lire l'ADN ou l'ARN contenu dans les cellules d'un échantillon biologique pour produire un texte numérique composé d'une succession de lettres (ACGT ou ACGU) (Peterlongo, 2016).
- Le séquençage de novo qui consiste à obtenir la séquence d'un organisme entier (assemblage de données de séquence d'un génome inconnu);
- le reséquençage du génome complet pour répertorier les variations nucléotidiques et structurales;
- le reséquençage ciblé d'ADN (ou « capture ») par amplicon ou hybridation ;
- le génotypage qui consiste à l'extraction de l'ADN et à son analyse sur puce.



L'utilisation des données de séquençage de ressources génétiques forestières

Pin maritime: marqueurs moléculaires au service de l'amélioration génétique

- Dès les années 1960 : constitution d'arbres "élites"
- Notions de cycle de sélection, croisement, greffage
- > Apport des techniques de génotypage et séquençage: stratégies de sélection et production de variétés plus courtes.
- Identification de l'origine géographique des RGF
- Etudier le régime de reproduction dans les vergers
- Certifier les RGF du programme d'amélioration
- Simplifier les cycles de sélection
- Prédire la valeur d'un arbre par la génomique pour accélérer les cycles de sélection

