

Partie I : Caractéristiques structurales du génome d'*Arabidopsis thaliana*

1. Introduction

Arabidopsis thaliana (arabette de Thalius) est la première plante dont le génome fut séquencé en totalité en 2000. Elle a été choisie dans les années 1980 comme espèce de référence pour les études en génétique moléculaire végétale en tant que modèle dicotylédone.

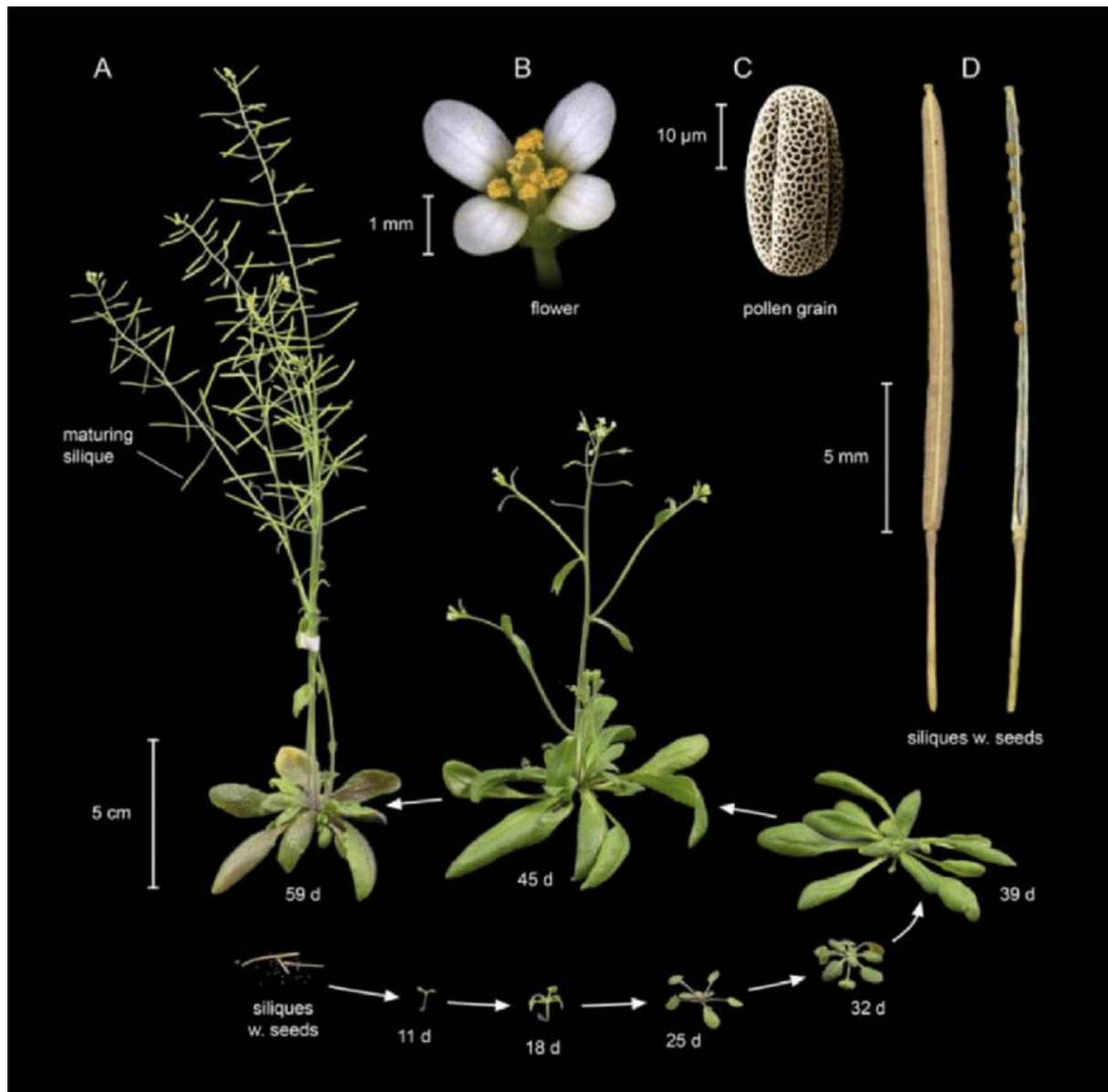


Fig.1. *Arabidopsis thaliana*. (A) *A. thaliana* à différents stades de son cycle de vie. (B) Fleur d'*A. thaliana*, (C) grain de pollen (micrographie électronique à balayage) et (D) siliques matures (gousses à graines ; à gauche : fermées ; à droite : ouvertes avec quelques graines non éclatées restantes) (adapté de Krämer, 2015).

2. Caractéristique morphologique d'*Arabidopsis thaliana*

C'est une petite plante dicotylédone très commune de la famille des Brassicaceae (crucifères). Elle présente des propriétés qui ont facilité son adoption comme modèle :

- un cycle de reproduction court, de l'ordre de 6 semaines ;
- une petite taille permettant de la cultiver facilement et en grand nombre, tant in vitro en boîtes de Pétri qu'en serre ;
- une grande prolificité (plusieurs milliers de graines de petite taille [0,5 mm de long] par plante, logées dans des fruits appelés siliques) ;
- une autogamie facilitant d'obtention de lignées pures (cf. Annexe 1) ;
- une petite taille du génome nucléaire, de l'ordre de 125 méga paires de bases (Mpb) ;
- une faible proportion de séquences répétées dans son génome (10 à 15 %).

3. Génome d'*Arabidopsis thaliana*

Le génome nucléaire d'*Arabidopsis thaliana* est de petite taille, environ 125 méga paires de bases (Mpb). Il possède 5 paires de chromosomes. Le séquençage, achevé en 2000, a permis de répertorier quelque 28 000 gènes. Il présente une faible proportion de séquences répétées (seulement 10 à 15 %), ce qui contraste avec d'autres plantes.

La structure du génome est le résultat de deux cycles de duplication de tout le génome, suivis d'un brassage et de pertes de gènes. Le génome a également connu au moins 30 duplications de segments. Caractéristiques Spécifiques aux Plantes ("Plant-specific") :

- **Transfert d'ADN** : Plus de 800 gènes nucléaires sont d'origine probable du génome des chloroplastes, indiquant un transfert d'ADN toujours en cours.
- **Éléments transposables** : Ils représentent au moins 10 % du génome.
- **Gènes uniques** : Présence de gènes spécifiques, comme ceux impliqués dans la biosynthèse de la paroi cellulaire, le transport des nutriments, la photosynthèse (comme la Rubisco), et ceux liés à la turgescence et aux réponses adaptatives à la vie sessile (phototropiques, gravitropiques).
- **Gènes de Résistance (Gènes R)** : Nombreux, hautement polymorphes, mais dispersés dans le génome.

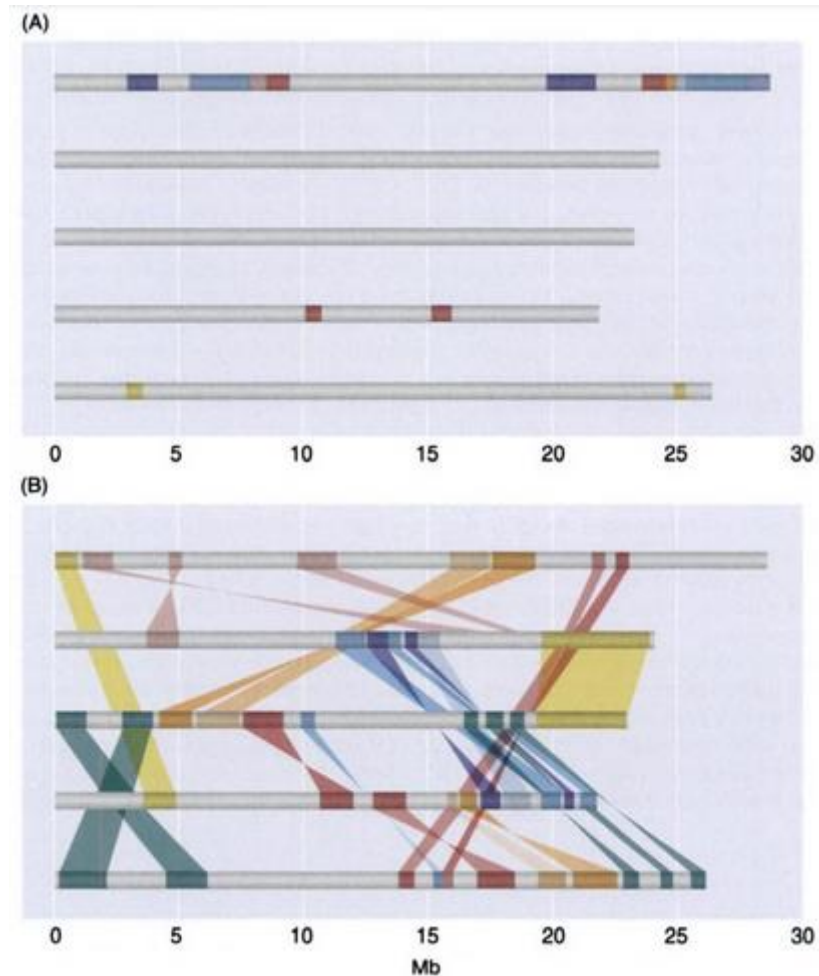


Fig.2. Les duplications chromosomiques dans le génome d'*Arabidopsis thaliana*.

Des alignements de tout le génome de la plante modèle Arabidopsis thaliana indiquent qu'il y a eu au moins 30 duplications de segments dans son génome.

(A) On peut voir sept duplications intra-chromosomiques représentées sous forme de duplications de blocs de couleur présents sur trois des cinq chromosomes ; cinq duplications se sont produites sur le premier chromosome et les chromosomes quatre et cinq présentent chacun une duplication.

(B) Représentation schématique de deux douzaines de duplications inter-chromosomiques de segments. Lorsque la bande est torsadée, cela indique qu'une inversion a accompagné l'événement de duplication. (Adapté de AGS 2000.). Frank et al. 2009

4. Bases de données relatives à *Arabidopsis thaliana*

Les données issues du séquençage et des études fonctionnelles sont coordonnées et mises à disposition de la communauté scientifique par des ressources en ligne.

4.1. TAIR (The Arabidopsis Information Resource) C'est le site Web principal du projet de séquençage. Il permet aux chercheurs de combiner la séquence génomique avec des banques d'EST, des cartes génétiques et physiques. Il fournit des liens vers des données de génétique fonctionnelle et moléculaire, la bibliographie propre à chaque gène, et une liste de souches mutantes.

3.2 Flagdb++

FLAGdb++ regroupe des connaissances génomiques et post-génomiques qui favorisent les études structurales et fonctionnelles sur une sélection de génomes végétaux, notamment celui d'*Arabidopsis thaliana*. à travers une interface graphique conviviale.

3.3 AraCyc (Arabidopsis thaliana Biochemical Pathways)

AraCyc est une **base de données biochimique** développée par **TAIR** (The Arabidopsis Information Resource) spécifiquement pour la plante modèle *Arabidopsis thaliana*. Son objectif est de fournir une représentation complète et conviviale du **métabolisme** d'Arabidopsis.

Points Clés :

- **Contenu** : Elle répertorie **plus de 170 voies métaboliques**, incluant des informations détaillées sur les **gènes**, les **protéines/enzymes**, les **réactions**, les **composés** et leur **localisation subcellulaire**.
- **Outils** : Elle utilise le logiciel **Pathway Tools** pour permettre la **visualisation graphique** des voies, allant de la vue d'ensemble du métabolisme jusqu'à la structure chimique des composés.
- **Construction et Qualité** : Initialement construite automatiquement à partir de la base de référence inter-espèces **MetaCyc**, la base AraCyc est ensuite **affinée et annotée manuellement** par des curateurs.
- **Spécificité Végétale** : Ce processus de curation permet d'ajouter des voies spécifiques aux plantes non présentes dans MetaCyc (ex : biosynthèse des **brassinostéroïdes** ou des **caroténoïdes**), augmentant ainsi la qualité et l'exhaustivité de la base.

3.4 Le centre de ressources *Arabidopsis thaliana* pour la génomique (INRA Versailles)

Le Nottingham Arabidopsis Stock Center (NASC) est une ressource cruciale en ligne. Il met à disposition des populations destinées à la cartographie, comme les lignées pures recombinantes (RI), ainsi que des souches mutantes. Il centralise l'information structurale, les données d'expression génique, et les amorces pour le génotypage.

Partie II : Caractéristiques structurelles du génome du riz

1. Introduction

Oryza sativa (le riz) est le modèle choisi pour représenter les monocotylédones dans les études de génomique, complétant la dichotomie des angiospermes avec *Arabidopsis thaliana* (dicotylédone). Le séquençage de son génome a débuté en 1997.

1. Intérêt du génome du riz par rapport à celui des autres céréales

Le riz possède une taille de génome relativement petite (environ 430 Mpb sur 12 chromosomes haploïdes) comparée aux autres grandes céréales. Son génome est environ 5 fois plus petit que celui du maïs et environ 40 fois plus petit que celui du blé. Les techniques de transformation du riz sont disponibles, de même que de nombreuses ressources génétiques.

2. Intérêt du génome du riz en lui-même

- Importance Agronomique Majeure : Le riz est une céréale essentielle pour l'alimentation mondiale, étant le principal aliment pour plus de la moitié de la planète et représentant 30 % de la production mondiale de céréales.
- Modèle pour les Monocotylédones : Sa petite taille génomique, par rapport aux autres céréales, en a fait le modèle de choix pour les études sur les monocotylédones.
- Synténie : La comparaison de sa séquence génomique avec celle d'*Arabidopsis* a révélé un grand degré de synténie (conservation de l'arrangement des gènes homologues).
[Image illustrating the synteny and gene arrangement differences between a segment of the rice genome and multiple segments of the *Arabidopsis thaliana* genome]

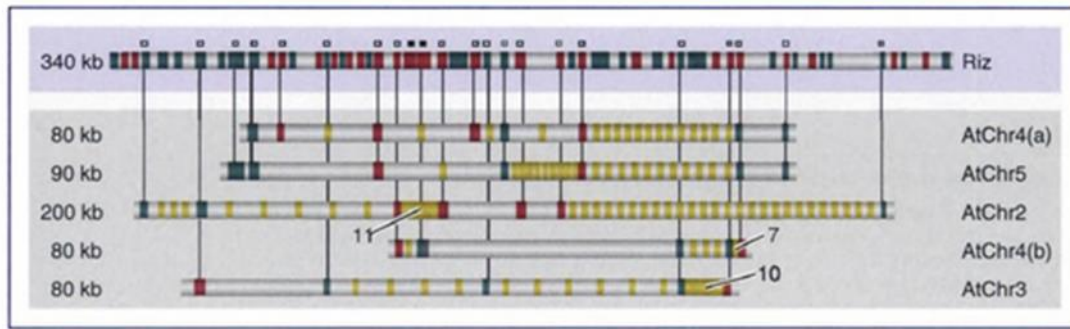


Fig.3. La synténie entre le riz et Arabidopsis. Vingt des 54 gènes compris dans un segment de 340 kb du génome du riz (en haut), conservent le même ordre dans cinq régions différentes de 80 à 200 kb du génome d'Arabidopsis (dessins du bas, non représentées à l'échelle). On trouve les gènes conservés sur les deux brins (cadres rouges et verts), mais ils sont séparés par un nombre variable d'autres gènes (boîtes jaunes) chez Arabidopsis. Les boîtes ombrées au-dessus du chromosome de riz indiquent que ce gène est conservé mais dans l'orientation opposée sur les chromosomes d'Arabidopsis. Frank et al. 2009

3. Bénéfices du séquençage complet du génome du riz

- Amélioration des Cultures : L'un des objectifs majeurs du séquençage est l'identification de gènes contrôlant les caractères agronomiques.
- Génomique Comparative pour les Céréales : L'analyse du génome du riz permet de rechercher et d'étudier les gènes homologues chez d'autres monocotylédones de grande culture (blé, maïs, etc.) grâce au phénomène de synténie. Les études sur le riz fournissent ainsi un cadre de référence pour l'amélioration d'autres céréales.