Université de Lille

UMR 8198 Evolution Ecologie Paléontologie (EEP)

Sylvain Legrand

Evo-Eco-Paléo

Structuration du laboratoire

- Laboratoire fondé en **2015** (GEPV+PAL du laboratoire Geosystèmes)
- **2 tutelles** : Université de Lille (FST) et CNRS (Écologie & Environnement)
- Direction : **Pr Xavier Vekemans**
- **2 équipes** : Evolution-Ecologie et Paléontologie
- Environ **50 permanents** pour un total de **80 personnes**
- Localisation **SN2** et **SN5**

Structuration du laboratoire

Direction : DU Xavier Vekemans **DU-adjoints** Taniel Danelian et Pascal Touzet

Equipe **Evolution et Ecologie**
Resp. Vincent Castric
Resp. adj. J.-F. Arnaud

- Groupe **Ecologie & génétique évolutives à diverses échelles EGEDE** (S. Billiard)
- Groupe **Spéciation** (C. Roux)
- Groupe **Ecologie des interactions et changement global** (N. Hautekèete)
- Groupe **Génomique évolutive et systèmes de reproduction chez les plantes** (C. Poux)

Equipe **Paléontologie**
Resp. Catherine Crônier

- Thème **Dynamique de la paléobiodiversité en temps profond** (B. Cascales-Minana)
- Thème **Tempo et mode de l'évolution morphologique** (B. Van Bocxlaer)
- Thème **Reconstructions des paléoécosystèmes** (S. Clausen)

Services communs et plateaux techniques

- Service administratif et logistique
- Service informatique (cellule bio-analyse / bio-informatique)
- Collections de recherche et patrimoniales
- Plateau de biologie moléculaire et biochimie
- Plateau de préparations paléontologiques et palynologiques
- Plateau d'imagerie

Plateforme Serres et Cultures en Environnements Contrôlés

Plateau de bioinformatique

- **2 ingénieurs** : **Mathieu Genète** (CNRS) et **Clément Mazoyer** (Ulille)
- **1 cluster de calcul** : 480 cœurs, 1.8 To RAM

Frontal du cluster SLURM

frontal	coeurs physiques	coeurs logiques	proc	cache	RAM
bee	2 * 20 = 40	40	Intel(R) Xeon(R) Silver 4210R CPU @ 2.40GHz	13.75 Mo	62 Go

Liste des noeuds

noeuds	coeurs physiques	coeurs logiques	proc	cache	RAM
agrion	2 * 64 = 128	128	AMD EPYC 7452 32-Core Processor @ 2.35 Ghz	128 Mo	515 Go
seed	2 * 48 = 96	96	AMD EPYC 7401 24-Core Processor @ 2.00 Ghz	64 Mo	386 Go
nereis	2 * 32 = 64	64	Intel(R) Xeon(R) CPU E5-4610 v2 @ 2.30GHz	16 Mo	257 Go
bryophyta	2 * 96 = 192	192	Intel(R) Xeon(R) Gold 6252 CPU @ 2.10GHz	35.75 Mo	644 Go

Participation à la promotion de la bioinformatique

- **Formations de l'Université de Lille en bioinformatique et biostatistiques**
  Master Bioinformatique (MISO, OSB), Master EVOBIO, Master Biotechnologies, Master Biologie Santé, Master IBVEM

  —>Introduction bioinfo, Bash, Python, cloud computing, statistiques, stages, projets

- **Formations Billile au Collège Doctoral**
  Cycle d'introduction à la bioinformatique et phylogénie

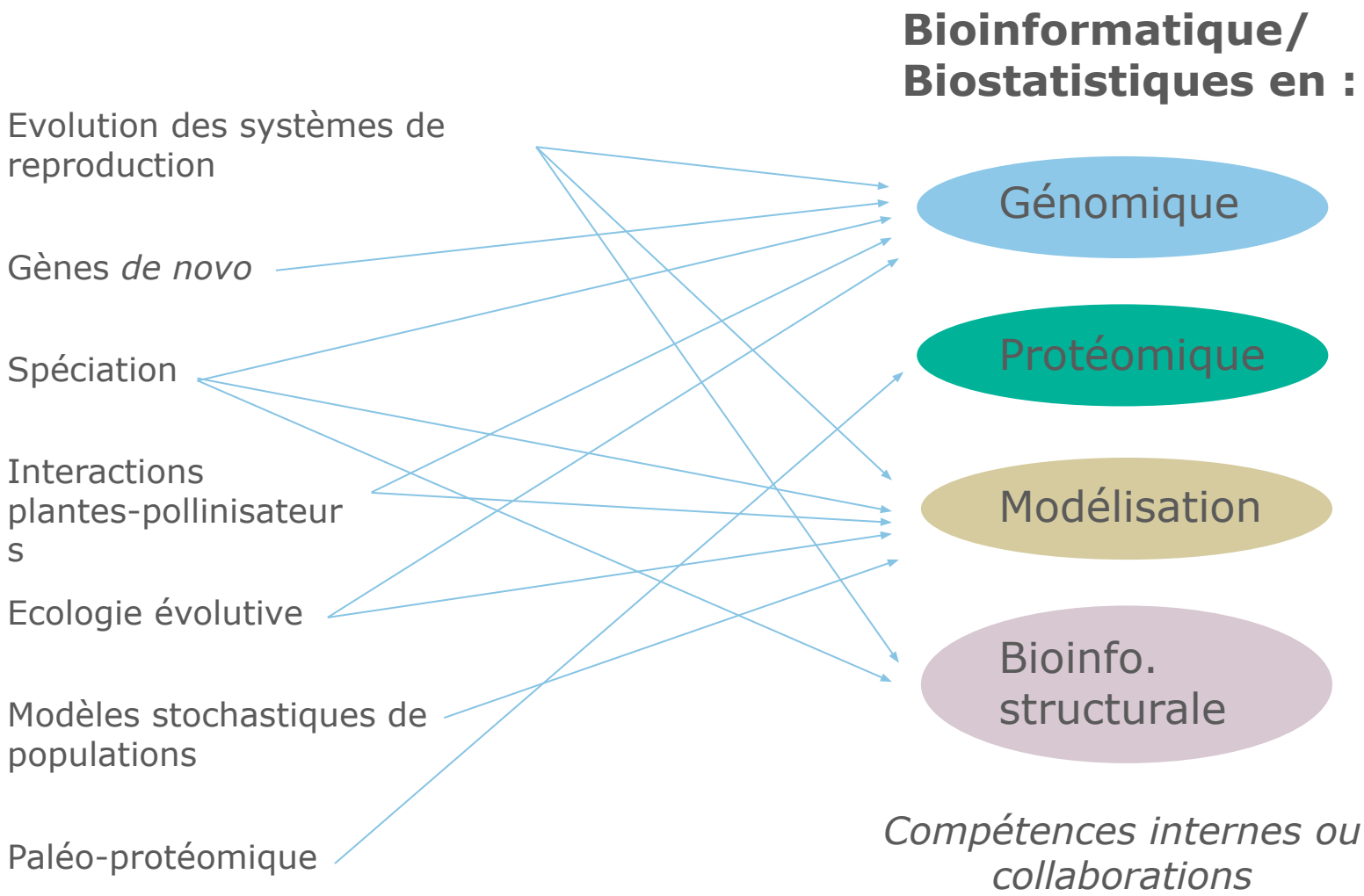
- **Ecole Bioinformatique de Roscoff**

- Comité Scientifique Interne de **Bilille**

- **Fête de la science**

- **Projet CDP** Protein Interaction Evolution (PIE) EEP, UGSF, CRIStAL, MSAP

La bioinfo dans les projets de recherche d'EEP

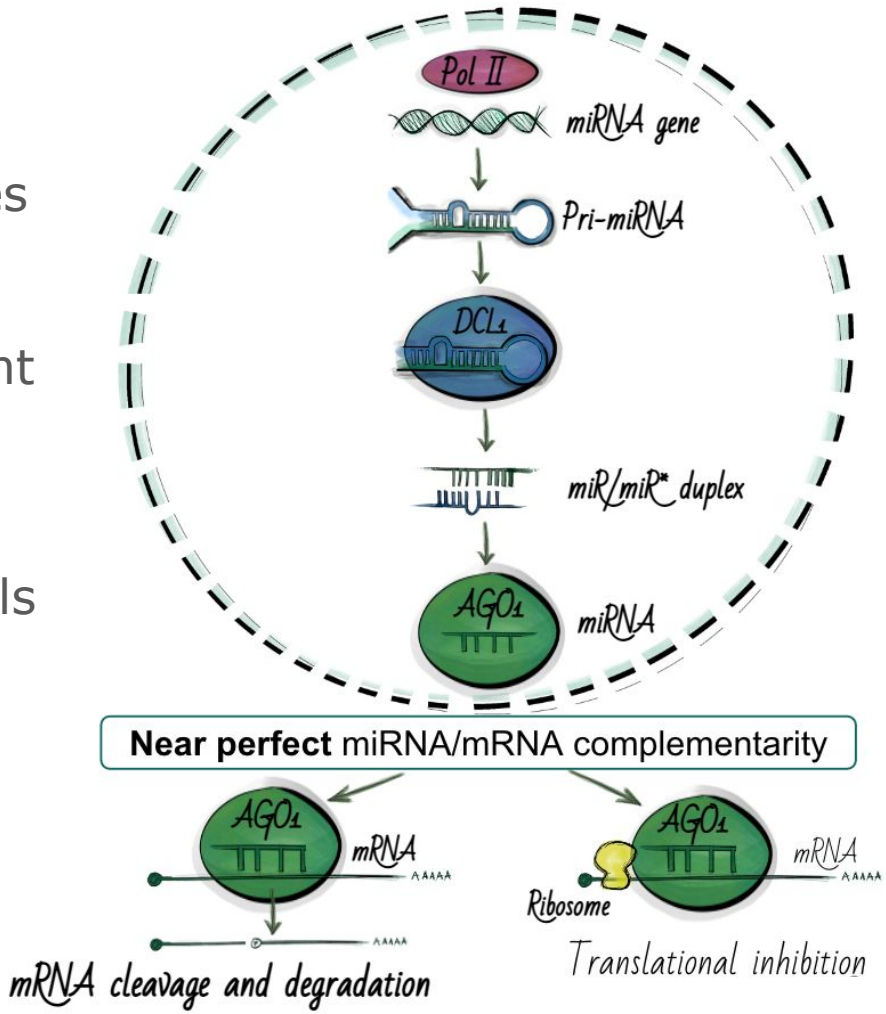


Evolution des micro-ARNs chez *Arabidopsis halleri* et *lyrata*
Thèse de Flavia Pavan

- Les **microARN** (miARNs) sont une classe de **petits ARNs non codants** qui régulent l'**expression des gènes** au niveau post-transcriptionnel chez les eucaryotes

- Alors que certains gènes de miARN sont profondément **conservés**, la majorité semble être **spécifique** à une espèce

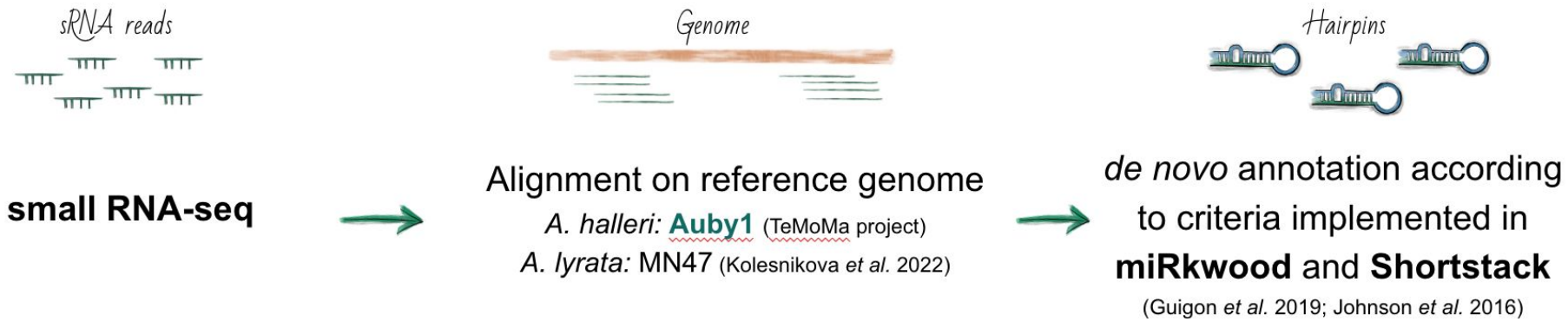
- Quels sont les processus par lesquels ils **émergent** et s'**intègrent** dans les **réseaux de régulation** cellulaire ?



Evolution des micro-ARNs chez *Arabidopsis halleri* et *lyrata*
Thèse de Flavia Pavan

Pour étudier l'évolution récente des gènes miRNA, il est nécessaire d'avoir des espèces qui ont connu une **divergence récente** —> divergence *A. halleri* et *A. lyrata* = ~1 million d'années

<u>Annotation des miRNA</u> chez différents individus *A. halleri* (19) et *A. lyrata* (13)

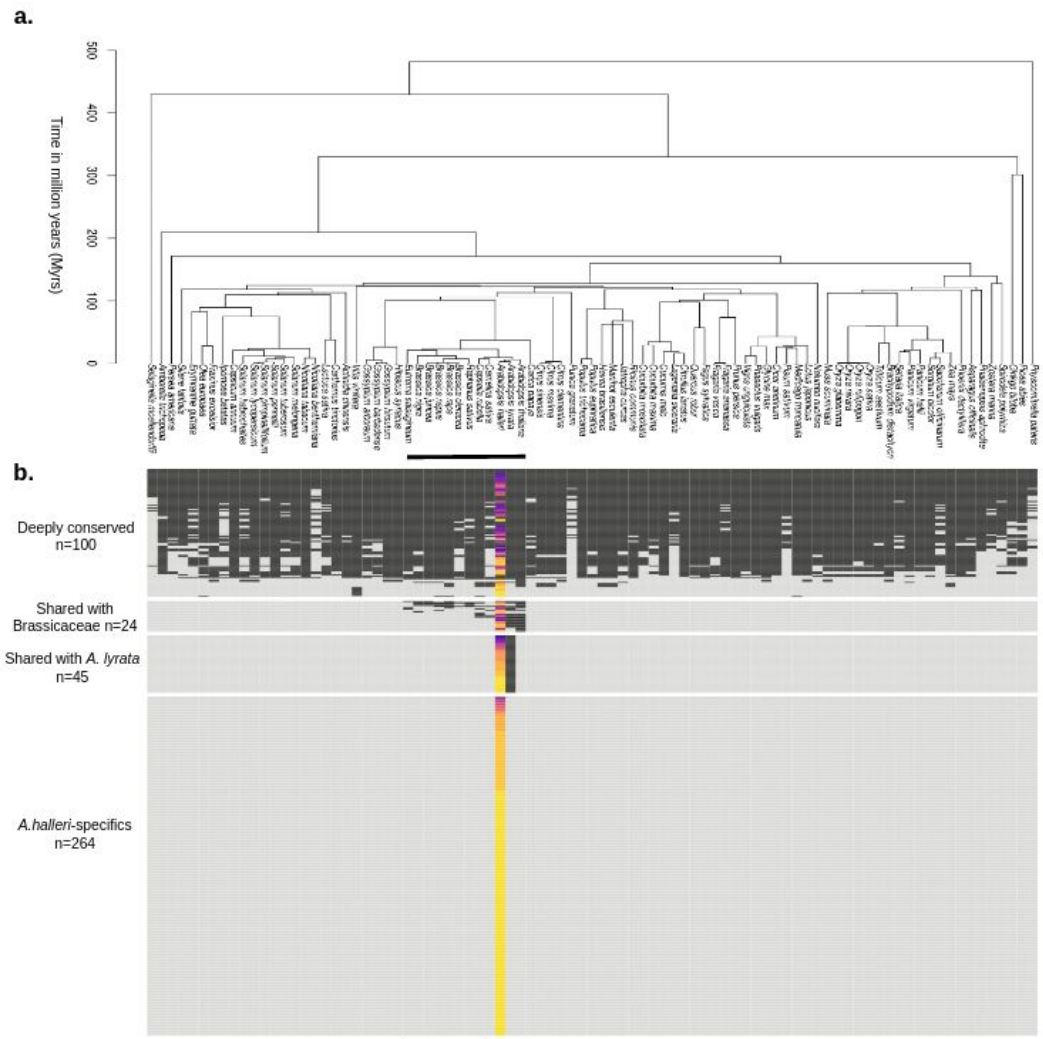


- **433** et **232** miRNAs respectivement identifiés chez *A. halleri* et *A. lyrata*

Evolution des micro-ARNs chez *Arabidopsis halleri* et *lyrata*
Thèse de Flavia Pavan

<u>**Conservation** des miRNAs</u>

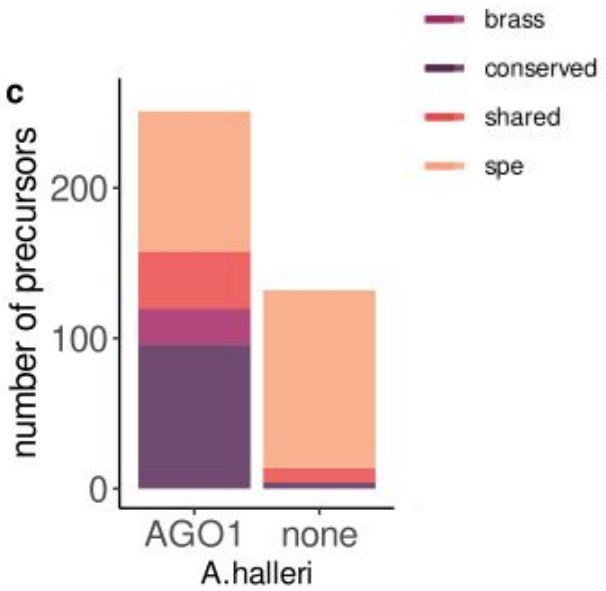
- Un grand nombre de **miRNAs spécifiques**
- **Apparition rapide** de nouveaux miRNAs mais **peu** sont **conservés à long terme**
- Spécifiques : uniquement des faux positifs ?



Evolution des micro-ARNs chez *Arabidopsis halleri* et *lyrata*
Thèse de Flavia Pavan

Validation des prédictions

- Par une approche d'**immunoprécipitation** des protéines AGO (AGO1 et AGO4) et séquençage des petits ARNs associés
- La quasi-totalité des miRNAs conservés sont **chargés dans AGO1**
- >1/3 des miRNAs spécifiques sont chargés dans AGO1

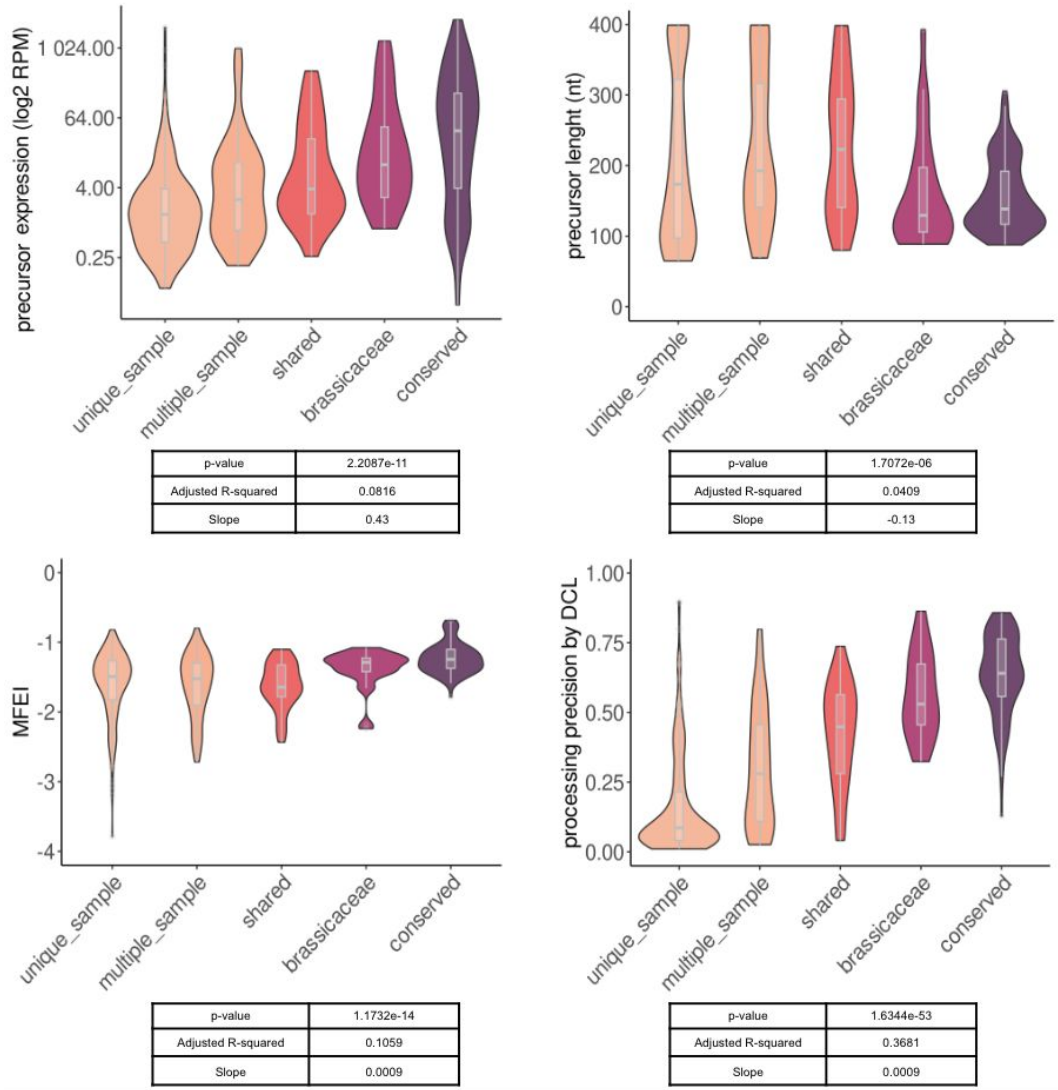


Evolution des micro-ARNs chez *Arabidopsis halleri* et *lyrata*
Thèse de Flavia Pavan

Caractéristiques des miRNAs

Au cours de l'évolution :

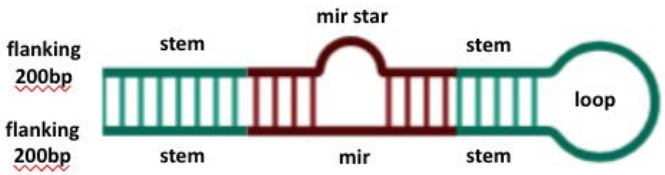
- **L'expression** du précurseur du miARN **augmente**
- La **longueur** du précurseur **diminue**
- La **stabilité du précurseur** diminue
- La **précision de clivage** du précurseur **augmente**



Evolution des micro-ARNs chez *Arabidopsis halleri* et *lyrata*
Thèse de Flavia Pavan

Contraintes fonctionnelles au sein des pre-miRNAs

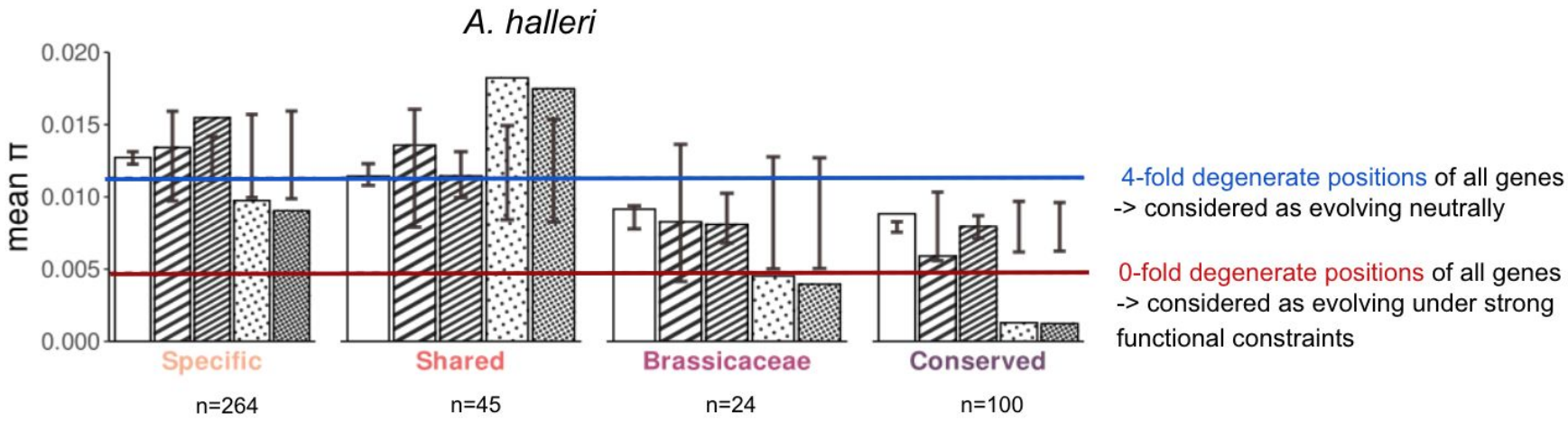
- Mesure de la **diversité nucléotidique** (polymorphisme ∏) à partir de données de séquençage de génome (Illumina) : GATK4/VCF tools/Python (150 individus pour *A. halleri*, 72 pour *A. lyrata*)



- Plus grande **contrainte fonctionnelle** pour le **duplex miRNA/miRNA***

Evolution des micro-ARNs chez *Arabidopsis halleri* et *lyrata*
Thèse de Flavia Pavan

Contraintes fonctionnelles au sein des pre-miRNAs

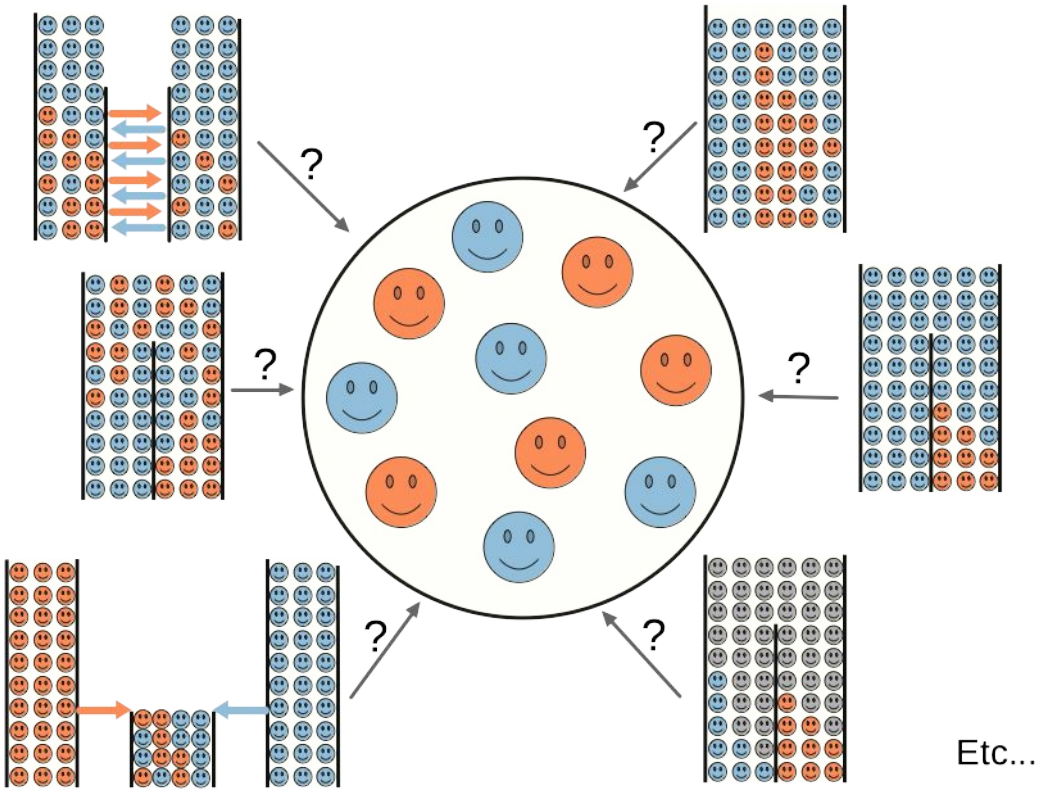


- Le **polymorphisme global** est **plus élevé pour les pre-miARN récents** que pour les anciens
- Les **contraintes sélectives augmentent au cours de l'évolution**

DILS par Camille Roux

- Outils de génétique des populations qui permet **à partir de données génomiques d'inférer une histoire démographique**

- Quelle **histoire démographique** explique-t-elle le mieux la diversité génétique actuelle ?



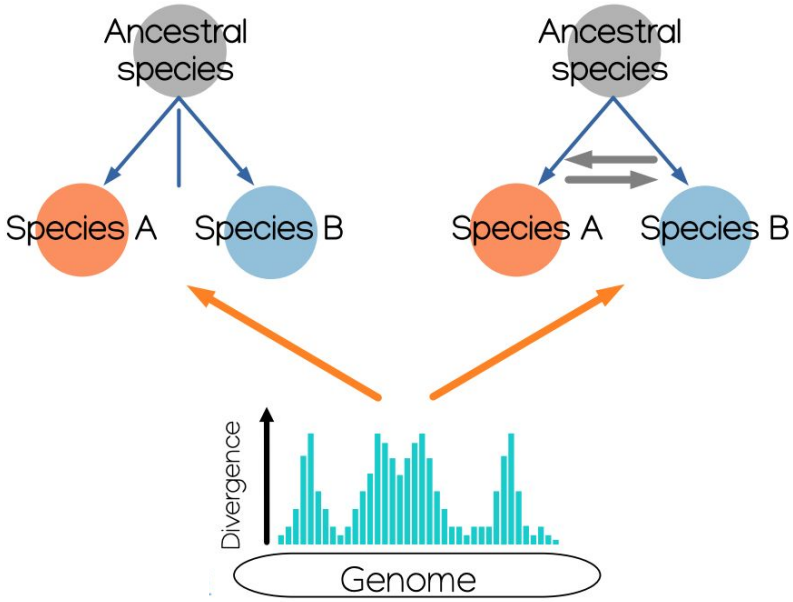
Fraïsse C, *et al*. DILS: Demographic inferences with linked selection by using ABC. Mol Ecol Resour. 2021 Nov;21(8):2629-2644

DILS par Camille Roux

- **(1) Analyse des données génomiques** d'individus d'une population —> mesures de différentes statistiques qui décrivent la diversité génétique le long du génome. ex: divergence

- **(2) Simulations de données** pour ces mêmes statistiques pour différentes histoires démographiques

- **(3)** Par **machine learning : quel scénario explique-t-il** le mieux **les données réelles** ?



Sylvain Legrand
Maître de Conférences
UMR CNRS 8198 EVO-ECO-PALEO
Evolution, Ecologie et Paléontologie
Université de Lille - Faculté des Sciences et Technologies
Bât SN2, bureau 208 - 59655 Villeneuve d'Ascq

sylvain.legrand@univ-lille.fr | http://eep.univ-lille.fr/
Tél. +33 (0)3 20 43 40 16