

# Stage Master/Ingénieur en Bioinformatique et Génomique.

## THEME : Analyse Comparée des Génomes de Plantes et d'Animaux.

### ✓ Contexte scientifique :

Afin d'étudier l'évolution des génomes de plantes et d'animaux, la synténie (conservation des gènes entre espèces) et les duplications totales de génomes sont étudiées. En comparant les génomes séquencés d'espèces modernes animales et végétales l'équipe PPAV (Paléogénomique des Plantes pour l'amélioration Variétale) de l'INRA de Clermont-ferrand (<http://www4.clermont.inra.fr/umr1095/Equipes/Recherches/Paleogenomique-des-plantes-pour-l-amelioration-varietale>) a défini un ancêtre à 5 chromosomes chez les Monocots, 7 chez les Dicots et 10 chez les vertébrés permettant de retracer l'histoire évolutive génomique des plantes et des animaux au niveau structural (Murat et al. 2012). L'équipe s'intéresse également à des aspects fonctionnels des gènes et de régulation des génomes liés à l'évolution, notamment concernant l'impact des duplications totales de génomes sur l'expression/méthylation des gènes dupliqués au sein des génomes modernes. Dans ce sens, l'équipe s'investit de façon importante dans le développement d'outils et de méthodes permettant de mettre en évidence des relations d'homologies entre et au sein des génomes de plantes et d'animaux afin de reconstruire leurs ancêtres et de comprendre leurs histoires évolutives.

A ce jour, le travail de mise en évidence de ces relations de synténie et de paralogie chez les plantes et les animaux séquencés a été effectué et publié (*cf* liste publication jointe). L'objectif du stage est d'étudier différents méthylomes chez les plantes et les animaux afin de comprendre et de comparer les modes de régulation de ces génomes ainsi que de comprendre certaines différences génomiques entre les plantes et les animaux tel que la présence ou l'absence de chromosomes sexuels. Sur la base des travaux de l'équipe, le stagiaire devra assimiler les concepts ainsi que les différents outils et méthodes utilisés par l'équipe afin de centraliser toutes les connaissances nécessaires à ce travail, il devra mettre en place une méthode afin d'étudier ces méthylomes (data mining, mapping, interprétation, ...) et l'automatiser. Le développement de programmes sera donc également indispensable pour répondre aux objectifs.

### ✓ Compétences requises

Maîtrise de langages de programmation (C, perl, JAVA, SQL, PHP...) ainsi que des outils d'interfaçage (HTML, CGI MySQL, PHP, perl CGI) appliqués aux systèmes d'exploitation classiques (IRIX, Linux, SOLARIS, OSF1, MacOS X). Exploitation des séquences (comparaison de séquences, utilisation de banque de données biologiques, alignements multiples, détection/annotation de gènes/repeat, phylogénie...). Maîtrise des techniques de sélection, exploration, modification et modélisation de grandes bases de données (Access...). Méthodes statistiques (régression, AFCM, ACP, analyse discriminante, classification, test de permutation). Visualisation des données.

### ✓ Laboratoire d'accueil :

INRA/UBP UMR 1095 Génétique Diversité et Ecophysiologie des Céréales  
5 chemin de Beaulieu 63100 Clermont-ferrand

✓ Durée et période souhaitées : 6 mois.

### ✓ Encadrement :

Jérôme Salse : Directeur de recherche ([jsalse@clermont.inra.fr](mailto:jsalse@clermont.inra.fr))

Florent Murat : Ingénieur de recherche Bio-informatique ([Florent.MURAT@clermont.inra.fr](mailto:Florent.MURAT@clermont.inra.fr))

Tél : 04 73 62 43 80.

### ✓ Moyens accordés :

Toutes les ressources, l'infrastructure (bio-informatique) et l'encadrement scientifique associés à la pleine réalisation du projet ainsi que la finalisation du mémoire de fin d'étude.

### ✓ Profil stagiaire souhaité :

Maîtrise de la programmation, bases en statistiques, intérêt pour la biologie.

### ✓ Quelques publications de l'équipe dans le domaine

1. Salse J, Abrouk M, Bolot S, Guilhot N, Courcelle E, Faraut T, Waugh R, Close TJ, Messing J, Feuillet C (2009) Reconstruction of monocotelydoneous proto-chromosomes reveals faster evolution in plants than in animals. **PNAS USA**. 106:14908-14913.
2. Salse J, Abrouk M, Murat F, Masood Quraishi U, Feuillet C (2009) Improved standards and new comparative genomics tools provide new insights into grasses paleogenomics. **Briefings in Bioinf.** 10(6):619-30.
3. Murat F, Xu JH, Tannier E, Abrouk M, Guilhot N, Pont C, Messing J, Salse J (2010) Ancestral Grass Karyotype Reconstruction Unravels New Mechanisms of Genome Shuffling as a Source of Plant Evolution. **Genome research**. 20(11):1545-1557.
4. Salse J (2012) In silico archeogenomics unveils modern plant genome organisation, regulation and evolution. **Current Opinion in Plant Biology**. 15(2):122-30.
5. Murat F, Peer YV, Salse J (2012) Decoding plant and animal genome plasticity from differential paleo-evolutionary patterns and processes. **Genome Biology and Evolution**. 4(9):805-16.