

## Répondre aux besoins de traitement à grande échelle en sciences de la vie

La plateforme met à disposition de ses utilisateurs une infrastructure matérielle et logicielle performante dédiée aux sciences de la vie comprenant :

- Un cluster de calcul de plus de 3000 cœurs
- Un espace de stockage de plus de 2 Po
- Des serveurs hébergeant plus de 100 machines virtuelles
- Près de 1000 logiciels du domaine mutualisés
- Plus de 200 banques de données mises à jour, indexées et partagées

Fin 2017, l'équipe rassemble sept personnels permanents (5,4 ETP) et un non permanent (1 ETP).



F. Cabanette



Ch. Gaspin



C. Hoede



Ch. Klopp



D. Laborie



J. Mariette



C. Noiroit



M.-S. Trotard

Elle s'appuie sur son savoir-faire interne et son réseau de compétences pour accompagner les besoins des programmes scientifiques sur :

- Le traitement à grande échelle des données « omiques »
- Le développement d'outils innovants en concertation avec la communauté scientifique et les autres plateformes nationales
- La formation continue à l'utilisation de l'infrastructure et au traitement de données
- Le support aux plateformes de production de données



### Responsable scientifique :

Christine Gaspin

### Responsable technique :

Christophe Klopp

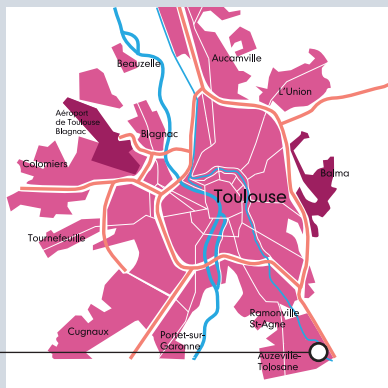
### Contact :

anim.bioinfo@toulouse.inra.fr

Site web : <http://bioinfo.genotoul.fr>

 @BioinfoGenotoul

### Localisation des équipements



Genotoul  
Bioinfo  
INRA

Les faits marquants scientifiques :

## L'analyse comparative de génomes au rendez-vous...

La disponibilité d'un nombre croissant de génomes séquencés, d'espèces proches ou plus éloignées, a donné un formidable élan à leur analyse comparative. Dans ce domaine, la plateforme bioinformatique met en avant cette année deux de ses contributions portant respectivement sur la reconstruction de génomes ancestraux de plantes et l'analyse de l'évolution des ARN non codants (ARNnc) régulateurs et de leur co-évolution avec les gènes codants dans le genre *Listeria*.



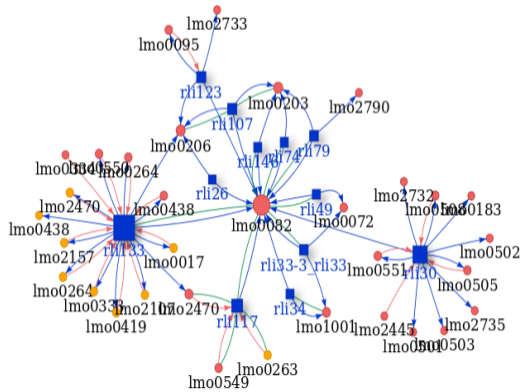
### Reconstruction du génome ancestral de plantes

En comparant la séquence d'ADN d'une trentaine de génomes de plantes, le groupe qui a mené ce travail a reconstruit le génome ancestral le plus récent de la première plante à fleurs. Ce génome ancestral serait constitué de 15 chromosomes et de plus de 20.000 gènes. L'apparition de copies de gènes ainsi que la réorganisation des chromosomes ancestraux au cours de l'évolution seraient à l'origine de nouvelles espèces végétales. Pour en savoir plus : Murat et al., Nature Genetics, 2017

### Analyse de l'évolution des ARNnc régulateurs et de leur co-évolution avec les gènes codants dans le genre *Listeria*

Cette analyse *in silico* a permis de produire l'histoire évolutive des ARN non codant (ARNnc) régulateurs et de leur co-évolution avec les gènes codants pour des protéines. Elle a été mise en œuvre sur près de 80 génomes du genre *Listeria* et a permis de mettre en évidence un réseau de 12 ARNnc qui co-évoluent avec des gènes codants pour des protéines de la paroi ainsi que des facteurs de virulence.

Pour en savoir plus : Cerutti et al., BMC Genomics, 2017



## PUBLICATIONS

- Murat F, Armero A, Pont C, Klopp C, Salse J. Reconstructing the genome of the most recent common ancestor of flowering plants. Nat Genet. 2017 Apr; 49(4):490-496.
- Cerutti F, Mallet L, Painset A, Hoede C, Moisan A, Bécavin C, Duval M, Dussurget O, Cossart P, Gaspin C, Chiappello H. Unraveling the evolution and coevolution of small regulatory RNAs and coding genes in *Listeria*. BMC Genomics. 2017 Nov 16; 18(1):882.

Le fait marquant technologique :

## L'infrastructure de calcul et de stockage innove...

La période 2017-2018 marque l'acquisition de nouveaux équipements, qui seront hébergés dans le data center du site Inra d'Auzeville-Tolosane. Cette nouvelle infrastructure, qui s'inscrit à terme dans le projet d'infrastructure nationale IFB vise à déployer des solutions répondant aux besoins des utilisateurs des sciences de la vie sur les différents nœuds du réseau de la fédération nationale académique.



Data center du site INRA



Le nouveau cluster

Le nouveau service apportera de l'agilité au cluster de calcul traditionnel. Il permettra d'augmenter les performances et la capacité du système de stockage parallèle. Il deviendra aussi plus facile de déployer des environnements logiciels complets correspondant à la demande.

### 2017 : les nouveaux nœuds de calcul sont en test

- 48 nœuds avec deux processeurs Intel Broadwell E5-2683v4 (16 cœurs, 2,1 GHz), 256 Go de RAM
- 1 nœud Bigmem SMP avec 48 cœurs, 2,2 GHz, 1536 Go de RAM, 22 To de stockage dédié
- 1 nœud de visualisation avec deux processeurs Intel Broadwell E5-2683v4 (16 cœurs, 2,1 GHz), 128 Go de RAM, carte GPU Nvidia Tesla K40
- Une interconnexion haut débit et faible latence (56Gbits/s)

### 2018 : les pratiques changent...

- Pour une répartition optimale des jobs, SLURM remplacera SGE
- La visualisation 3D à distance devient possible
- Les technologies du cloud seront évaluées.

## VALORISATION

### Nos utilisateurs

Plus de 1100 utilisateurs utilisent les ressources de la plateforme bioinformatique Génotoul Bioinfo en 2017 dont plus de 50% sont localisés en région Occitanie.

### Ils nous remercient ou publient avec nous.

Plus de 50 articles remercient la plateforme bioinformatique Génotoul Bioinfo en 2017 (<http://bioinfo.genotoul.fr/index.php/about-us/publications>)