



Présentation de la plate-forme

INRA - MIAT - Plate-forme Bioinformatique GenoToul
Chemin de Borde Rouge,
BP52 627,
31326 Castanet Tolosan cedex

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

L'équipe au 01/06/19

COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)

COEUR DE METIER
LA SEQUENCE

20 ANS DE VIE

11,5 ETP dont 5 CDD



Christine Gaspin (30% FTE)
DR INRA / Scientific animation
[+33 \(0\)5 61 28 52 82](tel:+330561285282)
christine.gaspin@inra.fr



Christophe Klopp (30% FTE)
IR INRA / Technical animation
[+33 \(0\)5 61 28 50 36](tel:+330561285036)
christophe.klopp@inra.fr



Claire Hoede (100% FTE)
IR INRA / Development and data analysis
[+33 \(0\)5 61 28 53 05](tel:+330561285305)
claire.hoede@inra.fr



Didier Laborie (80% FTE)
IE INRA / System administrator
[+33 \(0\)5 61 28 54 27](tel:+330561285427)
didier.laborie@inra.fr



Jérôme Mariette (100% FTE)
IE INRA / Development and data analysis
[+33 \(0\)5 61 28 57 25](tel:+330561285725)
jerome.mariette@inra.fr



Marie-Stéphane Trotard (100% FTE)
IE INRA / System administrator
[+33 \(0\)5 61 28 52 76](tel:+330561285276)
marie-stephane.trotard@inra.fr



Céline Noirot (100% FTE)
IE INRA / Development and data analysis
[+33 \(0\)5 61 28 57 24](tel:+330561285724)
celine.noirot@inra.fr



Patrice Dehais (50% FTE)
IE INRA / System administrator
[+33 \(0\)5 61 28 57 08](tel:+330561285708)
patrice.dehais@inra.fr



Sabrina Legoueix (20% FTE)
IE INRA / Development and data analysis
[+33 \(0\)5 61 28 57 08](tel:+330561285708)
sabrina.legoueix@inra.fr



Annick Moisan (40% FTE)
IE INRA / Quality management
[+33 \(0\)5 61 28 53 35](tel:+330561285335)
annick.moisan@inra.fr



Emilien Lasguignes (100% FTE)
CDD France génomique / Development and data analysis
[+33 \(0\)5 61 28 54 93](tel:+330561285493)
emilien.lasguignes@inra.fr



Joanna Fourquet (100% FTE)
CDD LabEx ECOFACT / Development and data analysis
[+33 \(0\)5 61 28 50 68](tel:+330561285068)
joanna.fourquet@inra.fr



Jean Mainguy (100% FTE)
CDD SeqOccIn (Region Occitanie / FEDER) co-encadré par
Géraldine Pascal GenPhySE / Development and data analysis
[+33 \(0\)5 61 70 52 87](tel:+330561705287)
jean.mainguy@inra.fr



Paul Terzian (100% FTE)
SeqOccIn (Region Occitanie / FEDER) co-encadré par Céline
Vandecasteele Get-Plage et Christophe Klopp Sigeneae /
Development and data analysis
[+33 \(0\)5 61 70 52 87](tel:+330561705287)
paul.terzian@inra.fr



Andreea Dreau (100% FTE)
CDD SeqOccIn (Region Occitanie / FEDER) encadrée par Matthias
Zytnicki MIAT / Development and data analysis
[+33 \(0\)5 61 28 55 73](tel:+330561285573)
andreea.dreau@inra.fr

L'infrastructure

COMPETENCES

- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS

(1DR, 2IR, 7IE)



COEUR DE METIER
LA SEQUENCE



20 ANS DE VIE



LOGICIELS (>600)

& BANQUES (237)

EQUIPEMENTS

- Clusters de calcul :
~3.000 cœurs (6000 threads)
- Espace stockage
~1 Po x 2
- Espace serveurs
- Espace partagé calcul
~1 Po

>1000 comptes

>2000 tickets / an



L'infrastructure

COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)



COEUR DE METIER
LA SEQUENCE



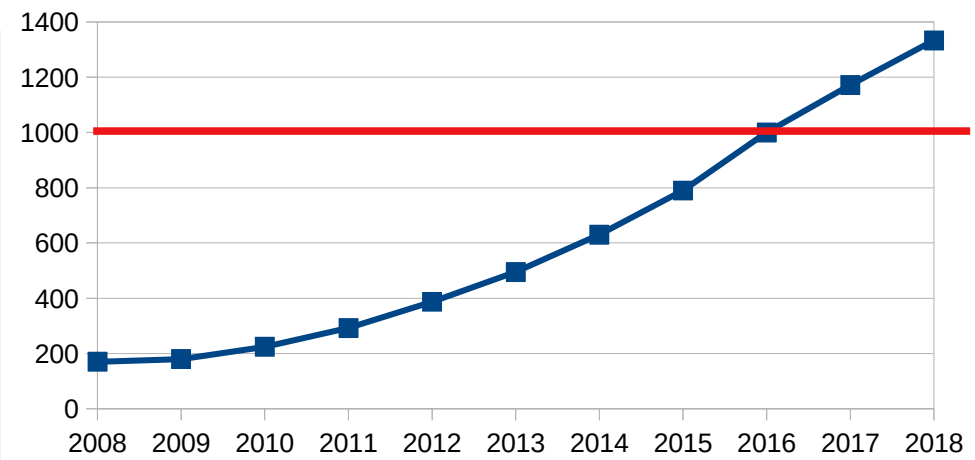
20 ANS DE VIE



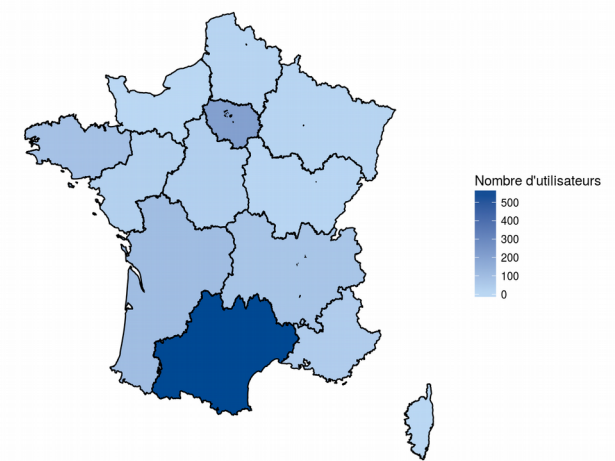
LOGICIELS (>600)
& BANQUES (237)

EQUIPEMENTS
- Clusters de calcul :
~3.000 cœurs (6000 threads)
- Espace stockage
~1 Po x 2
- Espace serveurs
- Espace partagé calcul
~1 Po

Nb utilisateurs GenoToul Bioinfo



Nombre d'utilisateurs par région



COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)

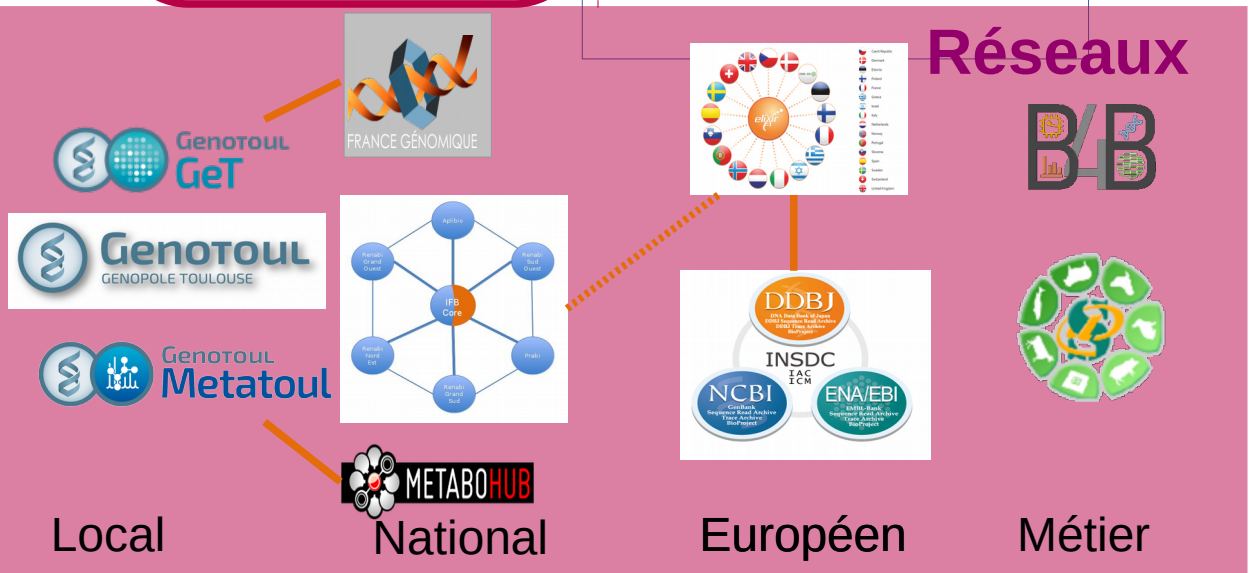
COEUR DE METIER
LA SEQUENCE

20 ANS DE VIE

LOGICIELS (471) & BANQUES (237)

EQUIPEMENTS
- Clusters de calcul :
~3.000 cœurs (6000 threads)
- Espace stockage
~1 Po x 2
- Espace serveurs
- Espace partagé calcul
~1 Po

Réseaux



COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)



COEUR DE METIER
LA SEQUENCE



20 ANS DE VIE

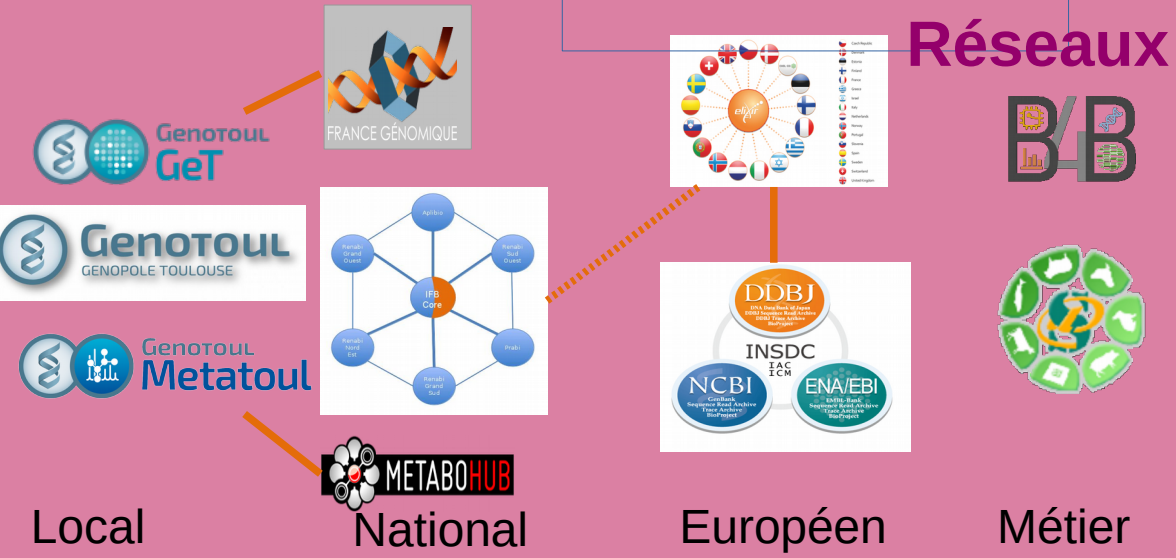




LOGICIELS (>600)
& BANQUES (237)

EQUIPEMENTS
- Clusters de calcul :
~3.000 cœurs (6000 threads)
- Espace stockage
~1 Po x 2
- Espace serveurs
- Espace partagé calcul
~1 Po

Réseaux





COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)

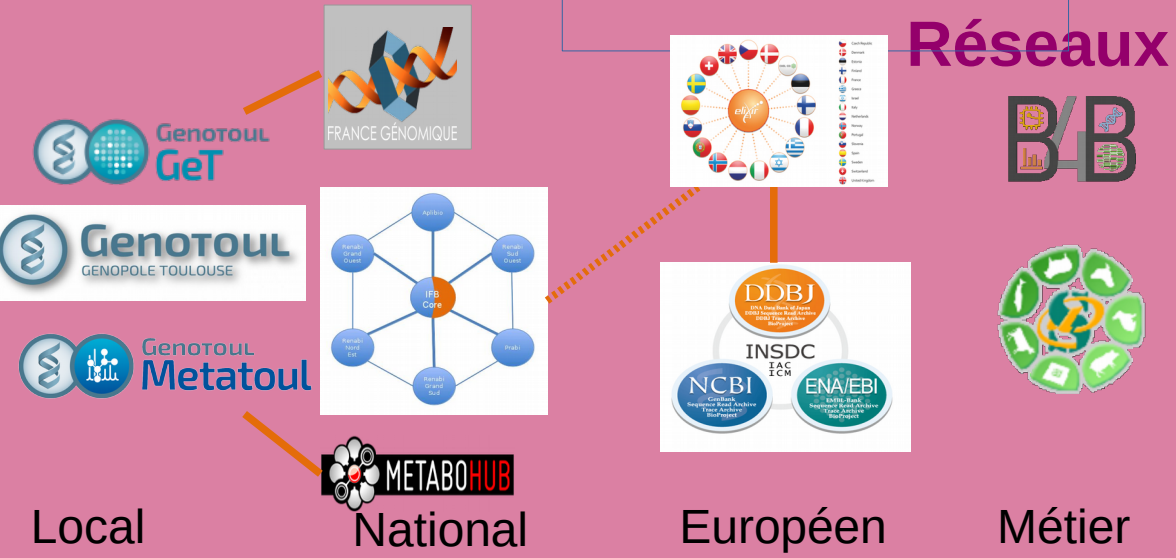
COEUR DE METIER LA SEQUENCE

20 ANS DE VIE

LOGICIELS (>600) & BANQUES (237)

EQUIPEMENTS
- Clusters de calcul : ~3.000 cœurs (6000 threads)
- Espace stockage ~1 Po x 2
- Espace serveurs
- Espace partagé calcul ~1 Po

Réseaux



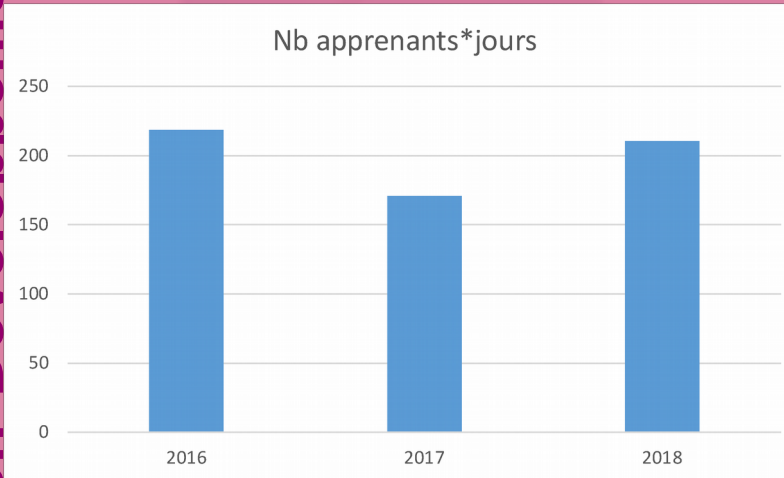
FORMATIONS

~20 Jours/an en partenariat

**CNRS
CIRAD
INSERM
INRA
Université
Privé**

Organizers	Vacancies	Registration
S. Manan, C. Cabou	-	-
G. Pascal, S. Manan, S. Rodriguez, M. Bernat, L. Cauquil	-	-
S. Manan, S. Rodriguez	-	-
P. Bardou, S. Rodriguez	-	-
G. Pascal, S. Manan, S. Rodriguez, M. Bernat, L. Cauquil	0	Register now

**40 % INRA
5 % privé**



~8 Jours/an
- Master Bioinfo UPS
- IFB/ELIXIR



COMPETENCES

- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)



COEUR DE METIER
LA SEQUENCE



20 ANS DE VIE



LOGICIELS (>600)
& BANQUES (237)

EQUIPEMENTS

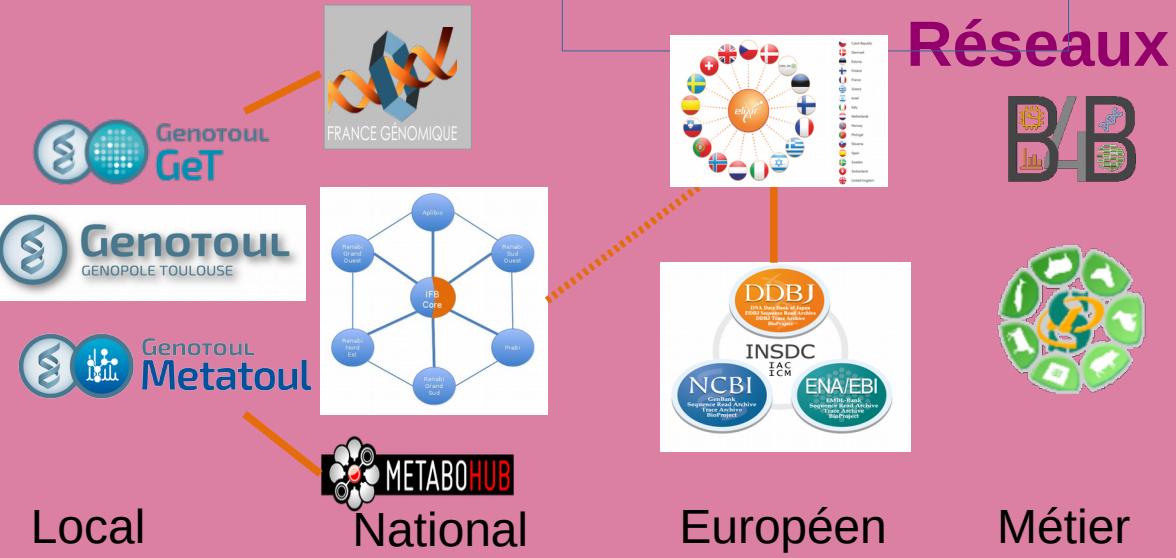
- Clusters de calcul :
~3.000 cœurs (6000 threads)
- Espace stockage
~1 Po x 2
- Espace serveurs
- Espace partagé calcul
~1 Po

FORMATIONS

- Des liens tissés avec les chercheurs/ingénieurs de MIAT et de GenPhyse (NED, Sigenae)
 - RNAseq
 - Métagénomique
 - Cycle Galaxy
- Des liens renforcés autour de Galaxy avec l'université via Sigenae

Accompagnement / Développement / Animation

Réseaux



COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)

COEUR DE METIER
LA SEQUENCE

20 ANS DE VIE

LOGICIELS (>600)
& BANQUES (237)

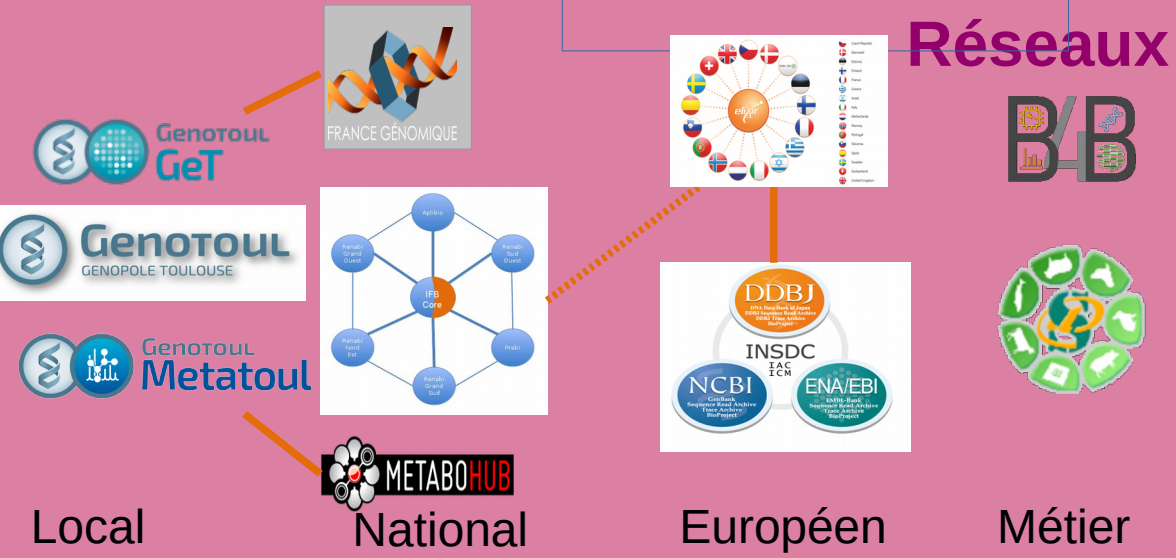
EQUIPEMENTS
- Clusters de calcul :
~3.000 cœurs (6000 threads)
- Espace stockage
~1 Po x 2
- Espace serveurs
- Espace partagé calcul
~1 Po

FORMATIONS

HEBERGEMENT
- ~ 80 Machines virtuelles (dont 35 externes)
- 1800 To de stockage (dont ~1Po occupé par les contributeurs (Sigenae, get-plage) dont ~800 To loués (56 contrats en cours)

Accompagnement
Animation Développement

Réseaux



COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)



COEUR DE METIER LA SEQUENCE



20 ANS DE VIE

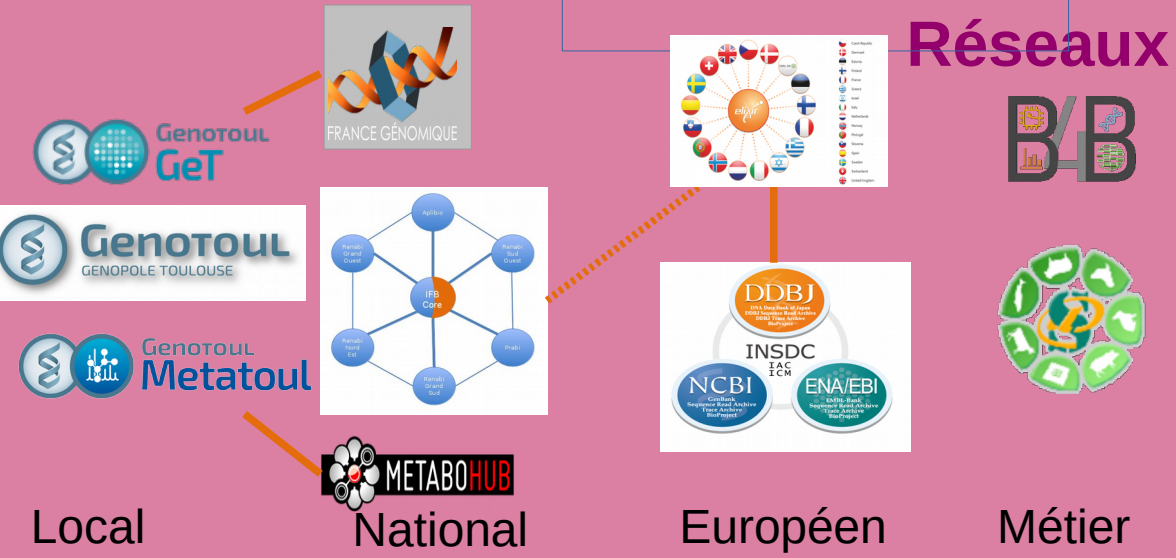




LOGICIELS (>600) & BANQUES (237)

EQUIPEMENTS
- Clusters de calcul : ~3.000 cœurs (6000 threads)
- Espace stockage ~1 Po x 2
- Espace serveurs
- Espace partagé calcul ~1 Po

Réseaux



FORMATIONS

HEBERGEMENT
- ~ 80 Machines virtuelles (dont 35 externes)
- 1800 To de stockage (dont ~1Po occupé par les contributeurs (Sigena, get-plage) dont ~800 To loués (78 contrats en cours)

ONT promethion, 09/2018, 7 To



Illumina NovaSeq
02/2018, 6 To

Accompagnement
Animation Développement



COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)



COEUR DE METIER
LA SEQUENCE



20 ANS DE VIE



LOGICIELS (>600)
& BANQUES (237)

EQUIPEMENTS
- Clusters de calcul :
~3.000 cœurs (6000 threads)
- Espace stockage
~1 Po x 2
- Espace serveurs
- Espace partagé calcul
~1 Po

Accompagnement
Développement
Animation

FORMATIONS

HEBERGEMENT

ACCOMPAGNEMENT PROJETS
Environ 25 projets en parallèle chaque année

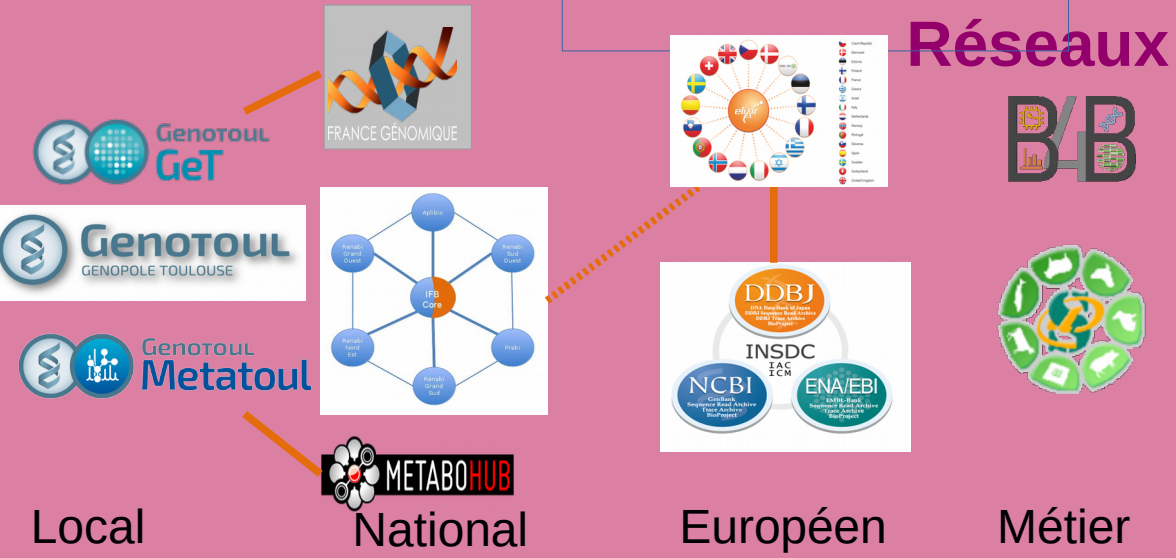
2 PIA, 1 ANR, 1 CPER
1 projet européen (coll. MIAT)
1 Labex (D. Pontier & X. Charpentier)
1 projet Région/FEDER SeqOccin
Projet région co-porté 6 M€
Avec le privé
3 axes (+ transversal) :
Assemblage + variations
Méthylation
Métagénomique

En cours d'évaluation :
- ANR coord. C Gaspin en coll. avec l'équipe de MF Sagot

>15 projets sous contrat propre

13 publications avec un membre de la PF en tant que co-auteurs en 2018

Réseaux



COMPETENCES
 - Bioinformatique
 - Informatique

ÉQUIPE DE 10 PERMANENTS
 (1DR, 2IR, 7IE)

COEUR DE METIER
LA SEQUENCE

20 ANS DE VIE

LOGICIELS (>600)
& BANQUES (237)

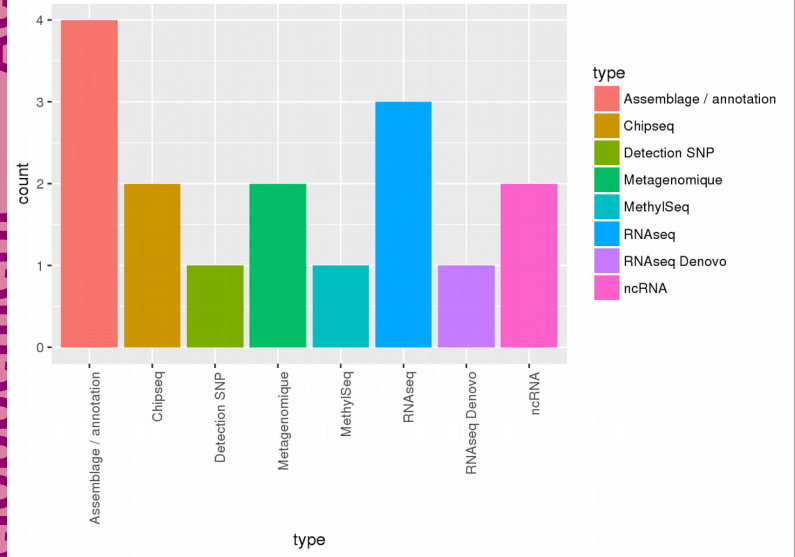
EQUIPEMENTS
 - Clusters de calcul :
 ~3.000 cœurs (6000 threads)
 - Espace stockage
 ~1 Po x 2
 - Espace serveurs
 - Espace partagé calcul
 ~1 Po

Animation Développement Accompagnement

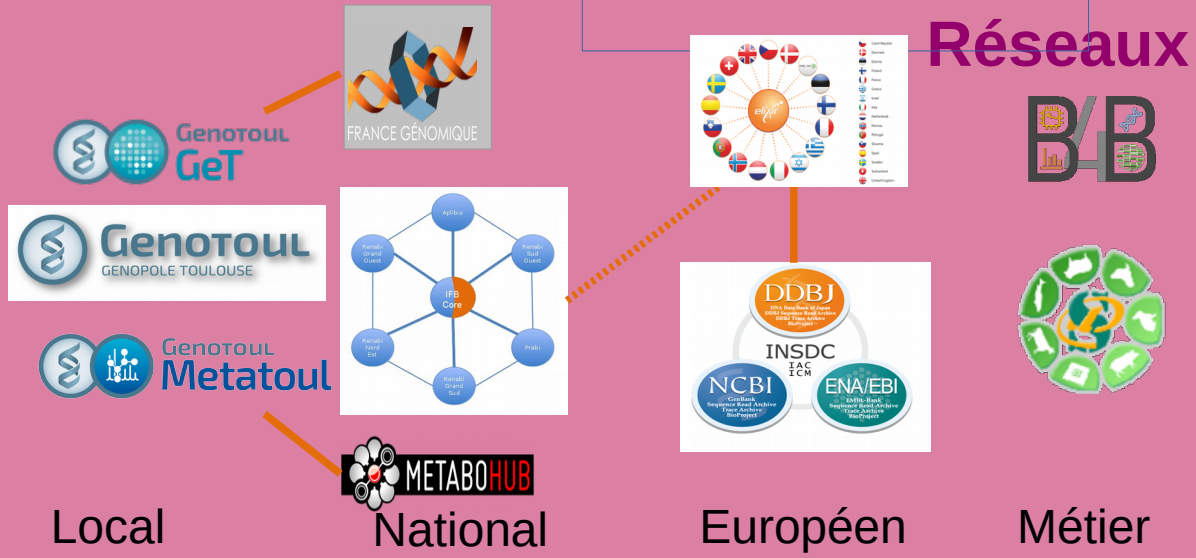
FORMATIONS

HEBERGEMENT

ACCOMPAGNEMENT PROJETS
 Environ 25 projets en parallèle chaque année



Réseaux



En cours d'évaluation :
 - ANR coord. C Gaspin en coll. avec l'équipe de MF Sagot

>15 projets sous contrat propre

13 publications avec un membre de la PF en tant que co-auteurs en 2018

COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)

COEUR DE METIER LA SEQUENCE

20 ANS DE VIE

LOGICIELS (>600) & BANQUES (237)

EQUIPEMENTS
- Clusters de calcul : ~3.000 cœurs (6000 threads)
- Espace stockage ~1 Po x 2
- Espace serveurs
- Espace partagé calcul ~1 Po

Accompagnement
Développement
Animation

FORMATIONS

HEBERGEMENT

ACCOMPAGNEMENT PROJETS
Environ 25 projets / an

DEVELOPPEMENT D'APPLICATIFS

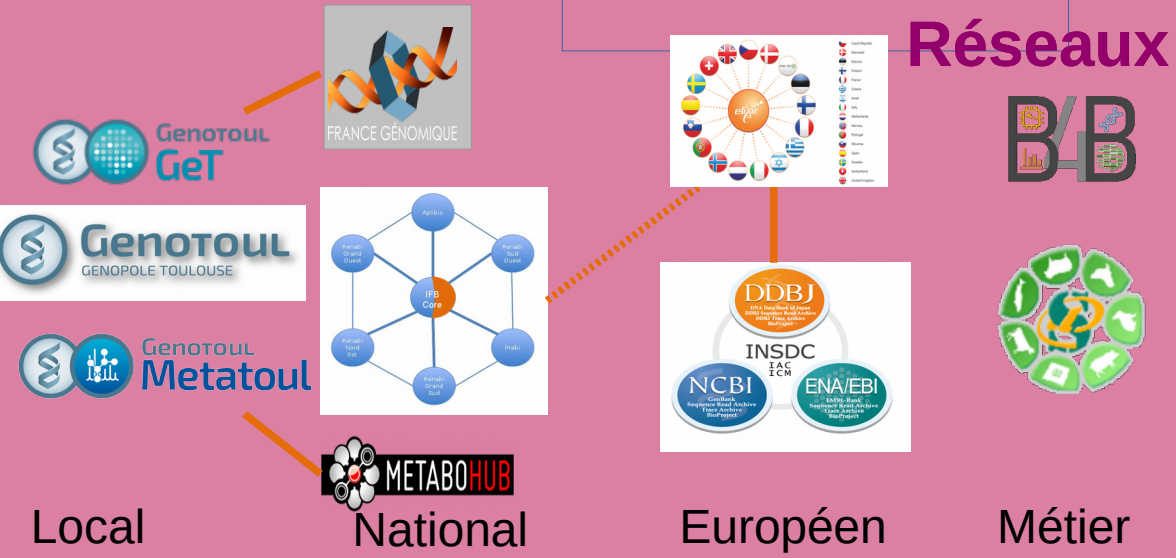
- Gestion et traitement des séquences issues du haut débit / ARNnc & réseaux de régulation
- PIA Infrastructure France Génomique
- PIA Infrastructure IFB

4 projets de développement en maintenance, 2 projets actifs

11 logiciels open source dont 9 publiés

3 publications d'outils développés en collaboration en 2018

Réseaux



COMPETENCES
 - Bioinformatique
 - Informatique

ÉQUIPE DE 10 PERMANENTS
 (1DR, 2IR, 7IE)

COEUR DE METIER LA SEQUENCE

20 ANS DE VIE

LOGICIELS (>600) & BANQUES (237)

EQUIPEMENTS
 - Clusters de calcul : ~3.000 cœurs (6000 threads)
 - Espace stockage ~1 Po x 2
 - Espace serveurs
 - Espace partagé calcul ~1 Po

Animation Développement Accompagnement

FORMATIONS

HEBERGEMENT

ACCOMPAGNEMENT PROJETS
 Environ 25 projets / an

DEVELOPPEMENT D'APPLICATIFS

D-Genies
 D-GENIES – for Dot plot large Genomes in an Interactive, Efficient and Simple way – is an online tool designed to compare two genomes. It supports large genome and you can interact with the dot plot to improve the visualisation.
 We use minimap version 2 to align the two genomes. Then, the PAF file is parsed and plotted into an interactive plot written with d3.js library.
 D-Genies also allows to display dot plots from other aligners by uploading their PAF or MAF alignment file.
 Cabannes E, Klapp C. (2018) D-GENIES: dot plot large genomes in an interactive, efficient and simple way. PeerJ 6:e4958 <https://doi.org/10.7717/peerj.4958>

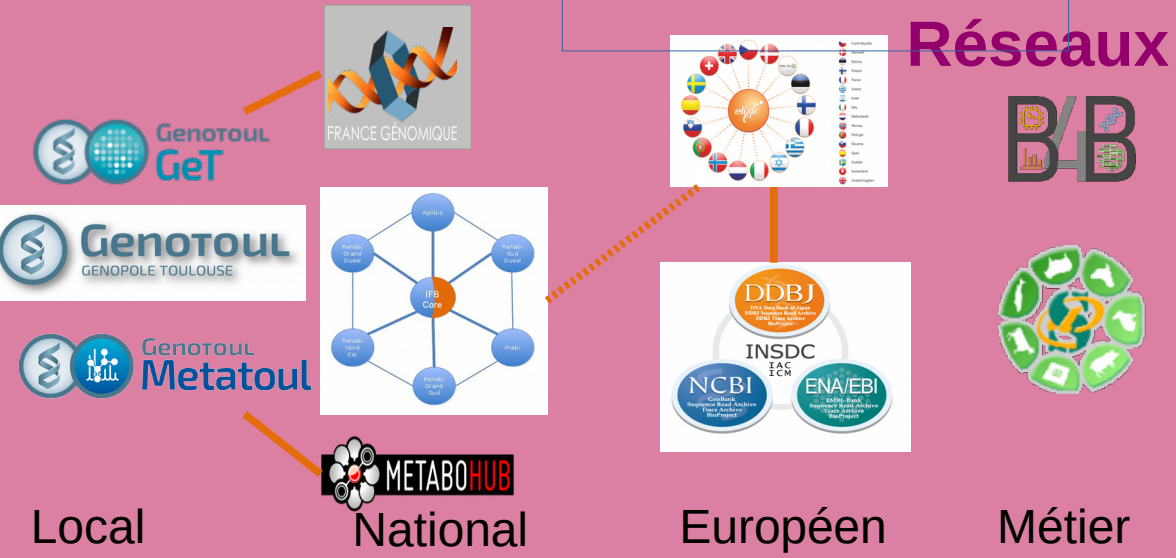
D-Genies Summary

Dot plots are widely used to quickly compare sequence sets. They provide a synthetic similarity overview, highlighting repetitions, breaks and inversions. Different tools have been developed to easily generate genomic alignment dot plots, but they are often limited in the input sequence size. D-GENIES is a standalone and WEB application performing large genome alignments using minimap2 software package and generating interactive dot plots. It enables users to sort query sequences along the reference, zoom in the plot and download several image, alignment or sequence files. D-GENIES is an easy to install open source software package (GPL) developed in Python and JavaScript.

D-Genies Availability

D-Genies is available under the GNU General Public License (GPL). The package is coming with full documentation and test data. Sources are available here.

Réseaux



COMPETENCES
- Bioinformatique
- Informatique

ÉQUIPE DE 10 PERMANENTS
(1DR, 2IR, 7IE)

COEUR DE METIER
LA SEQUENCE

20 ANS DE VIE

LOGICIELS (>600)
& BANQUES (237)

EQUIPEMENTS
- Clusters de calcul :
~3.000 cœurs (6000 threads)
- Espace stockage
~1 Po x 2
- Espace serveurs
- Espace partagé calcul
~1 Po

Accompagnement
Développement
Animation

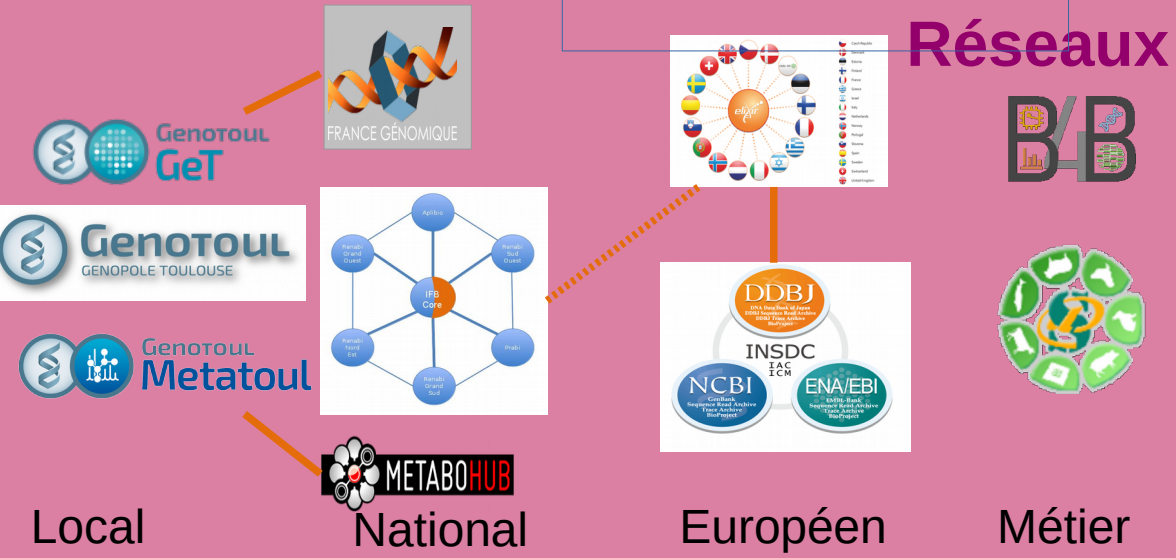
FORMATIONS

HEBERGEMENT

ACCOMPAGNEMENT PROJETS
Environ 25 projets / an

DEVELOPPEMENT D'APPLICATIFS

Réseaux



ANIMATION
Journées bioinfo/biostats Occitanie

Concours interne 2019

- IE développeur d'applications
- Fiche de poste complète :
http://jobs.inra.fr/offers/emploi_perm/concours/concours-internes-ingenieurs-cadres-et-techniciens/?concours=117735&campagne=23129&domaine=®ion=&corps=&entite_rattachement=&discipline=&annonce=26825&annonce_id=14#annonce
- Attention le profil s'intitule : "Ingénieur-e en études d'environnements géo-naturels et anthropisés" alors que c'est un profil de développeur d'application que nous recherchons.
- Les conditions requises pour concourir sont détaillées dans le guide du candidat disponible sur la page d'accueil du concours (section s'inscrire).
- Les conditions d'ancienneté requises dans la fonction publique (agents de l'État et des collectivités territoriales) pour un IE : 5 ans au premier janvier de l'année du concours.
- N'hésitez pas à faire circuler l'information et à nous contacter pour toutes questions. Contacts : christine.gaspin@inra.fr, christophe.klopp@inra.fr
- Clôture des inscriptions : 27 juin 2019.