

Cette lettre d'information est destinée aux membres des équipes de recherche utilisant la plate-forme bio-informatique GenoToul. Elle a pour but de vous informer sur les évolutions de l'équipe, les nouveaux outils, services, conditions d'utilisation, projets et formations mis en place.

1/ Les prochains cycles d'apprentissage

A) 4 jours de formation à la métagénomique 16S et 18S sous Galaxy seront organisés du 10 au 13 octobre 2016 (reste 2 places)

Vous avez ou allez obtenir des données de métagénomique 16S ou 18S Miseq ou 454 ? Vous souhaitez les analyser efficacement sans utiliser la ligne de commande ? Les équipes Sigenae, NED (GenPhySE), TWB et bioinfo genotoul vous proposent pour cela une formation de 4 jours indivisibles sous environnement Galaxy. La première demi-journée sera consacrée, après une initiation à l'environnement Galaxy à la prise en main de l'instance Galaxy Toulousaine. Pendant deux jours vous apprendrez à utiliser le pipeline FROGS dont la description est disponible à l'url :

http://bioinfo.genotoul.fr/fileadmin/BIO_INFO_STAT_2015/oral/FROGS_GenotoulBioinfo.pdf. Il est constitué des étapes principales suivantes : nettoyage, clustering, affiliation taxonomique, filtre et statistiques descriptives que nous vous détaillerons. Enfin la dernière journée et demie vous permettra d'interpréter les résultats grâce aux outils statistiques proposés.

Pré-requis nécessaire : connaissance de R ou d'un autre langage de programmation pour la partie statistique.

B) 2,5 jours de formation à l'alignement et à la détection de variants à partir de NGS sous Galaxy du 5 au 7 octobre 2016

La première journée sera consacrée à la présentation de l'environnement Galaxy, puis à la prise en main de l'instance toulousaine. Suivront ensuite 1,5 jours de formations vous permettant de comprendre les principaux formats de fichiers manipulés, d'apprendre à aligner les reads sur un génome de référence et de détecter SNP et petits indels via la suite GATK

Aucun pré-requis n'est nécessaire.

Ces formations sont organisées sur le site INRA de Toulouse Auzeville.

Les tarifs sont disponibles à l'adresse suivante : <http://bioinfo.genotoul.fr/index.php?id=115>.

Les inscriptions s'effectuent sur cette page : <http://bioinfo.genotoul.fr/index.php?id=10>.

2/ Mise à jour de l'instance Galaxy toulousaine

L'équipe Sigenae met à votre disposition la toute nouvelle version de l'instance Galaxy Sigenae / BioInfo Genotoul. Cette version est accessible depuis la page d'accueil (<http://sigenae-workbench.toulouse.inra.fr>) via le lien "Access to new Galaxy version". L'accès aux historiques ainsi que plusieurs outils ont été améliorés. Les outils les plus utilisés ont été migrés. Si certains outils vous manquent, n'hésitez pas à nous en faire part afin que nous puissions les migrer au fur et à mesure de vos besoins. D'ici un trimestre, la nouvelle version remplacera l'ancienne. Nous vous conseillons donc de récupérer depuis la version actuelle :

- vos workflows ("download workflow"),
- vos métadonnées (versions des outils Galaxy et bioinfo utilisés)
- vos jeux de données utiles (outil "Download my Galaxy dataset on Genotoul (work)"),

Galaxy is a workbench available for biologists from Sigenae Platform. Galaxy objectives are:

- Make bioinfo Linux tools accessible to biologists.
- Hide the complexity of the infrastructure.
- Allow creation, execution and sharing of workflows.

[Access to new Galaxy version](#)



WELCOME TO GALAXY WORKBENCH



pour pouvoir si besoin les importer dans la nouvelle version, les espaces de stockage des données étant différents entre les deux versions de Galaxy. L'équipe du support galaxy est disponible via l'e-mail : sigenae-support@listes.inra.fr pour répondre à toutes vos éventuelles questions.

3/ Vous êtes 1000 utilisateurs de la plateforme bioinfo genotoul

La plate-forme GenoToul vient d'atteindre le millier de comptes utilisateurs actifs sur la plate-forme. Par conséquent, nous sommes contraints de maintenir une politique stricte de gestion des comptes. Les comptes Linux inutilisés depuis +1 an sont donc systématiquement désactivés et les données purgées. De plus, nous ne pouvons plus conserver les données des comptes expirés sans demande explicite (via un mail au support : support.genopole@toulouse.inra.fr).

4/ Prochains départs dans l'équipe

Céline Noirot (Titulaire IE Bioinfo) et Alexandre Heurteaux (CDD IE Bioinfo/Biostat, Financement GenEndurance) quittent l'équipe à la fin du mois d'août. Céline prend de nouvelles responsabilités au CHU de Toulouse à l'Institut Fédératif de Biologie et Alexandre commence une thèse dans l'équipe « Dynamique chromatinienne et prolifération cellulaire » dirigée par Olivier Cuvier. Les membres de l'équipe sont présentés sur cette page : <http://bioinfo.genotoul.fr/index.php?id=4> régulièrement mise à jour.

5/ Extension du work

L'infrastructure de calcul Genotoul a été interrompue lundi 25 et mardi 26 juillet pour étendre significativement l'espace disque /work. Les quotas disques ne seront **PAS** augmentés; nous avons besoin de cette extension pouvoir y stocker les fichiers temporaires de calcul de l'ensemble de nos utilisateurs (dont le nombre est en forte croissance).

A cette occasion, nous avons mis à jour des firmwares matériels ainsi que le système de fichiers GPFS en Spectrum Scale 4.2 (corrections de bogues et nouvelles fonctionnalités).

6/ Les traitements de données sur le serveur frontal genotoul sont interdits

Pour que l'ensemble des utilisateurs de genotoul puissent travailler dans de bonnes conditions nous avons dû interdire le traitement des données sur le serveur frontal. Les serveurs genotoul sont réservés exclusivement à la connexion, au transfert de données, à la compilation, au test de la ligne de commande et à la soumission de jobs sur le cluster de calcul. Tout traitement de données lancé directement (sans «qsub», «qarray», «qrsh» ou «qlogin») sera systématiquement interrompu sans préavis par les administrateurs système. Ces quatre commandes sont explicitées dans notre FAQ (<http://bioinfo.genotoul.fr/index.php?id=11>) à la question : « Which commands can I use to submit my job » ?

7/ Un nouveau site web en préparation

Un nouveau site web est en cours de construction. Nous espérons qu'il vous apportera satisfaction. Lorsque qu'il sera mis en production (certainement en septembre / octobre), n'hésitez pas à nous faire remonter les problèmes que vous pourriez rencontrer.

8/ Journée Bioinfo-Biostat le 2 décembre prochain à Auzeville-Tolosane

La journée bioinfo-biostat se déroulera le 2 décembre 2016 sur le site INRA d'Auzeville. Comme d'habitude vous serez prochainement sollicités par e-mail pour l'appel à soumissions. Les inscriptions seront ouvertes en septembre et le programme sera disponible en novembre. Le comité d'organisation est constitué des personnes suivantes : Sébastien Dejean, Monique Falières, Christine Gaspin, Sandrine Laguerre, Raphaël Mourad, Céline Noirot, Matthias Zytnicki.

9/ Remerciements et publications

Vous êtes de plus en plus nombreux à utiliser l'infrastructure de Genotoul Bioinfo et à la citer dans vos publications. Nous vous en remercions. Le nombre de publications remerciant la plateforme Bioinfo GenoToul est un indicateur de notre utilité pour nos recherches de financements par exemple. C'est pourquoi nous vous proposons le modèle de phrase suivant lorsque vos publications traitent d'un sujet ayant bénéficié des ressources de la plateforme : "We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrenees for providing help and/or computing and/or storage resources".

Dans le cadre d'une collaboration, vous pouvez directement citer la / les personnes qui ont participé au projet de la manière suivante : Name, bioinformatics platform Toulouse Midi-Pyrenees, MIAT UR 875, INRA Auzeville CS 52627 31326 Castanet Tolosan cedex.

Vous trouverez sur notre site web (<http://bioinfo.genotoul.fr/index.php?id=54>), la liste des publications ayant utilisé l'infrastructure bioinfo Genotoul. Si vous avez publié en utilisant nos ressources et que votre article n'y apparaît pas, merci de nous en communiquer les références par mail à anim.bioinfo@toulouse.inra.fr.

Pour toute demande d'information ou de travaux, veuillez envoyer un mail à support.genopole@toulouse.inra.fr en précisant vos noms et coordonnées.