

Session de formations South Green - 1^{er} semestre 2019

La plateforme South Green, avec ces partenaires, organise une nouvelle série de modules de formations allant de l'initiation à linux aux analyses de données de séquençage en passant par l'utilisation des clusters de calcul itrop et CIRAD ou l'initiation à python.



18-19 Mars	Guide de survie à Linux - salle 151 -IRD
21 Mars	Initiation à l'utilisation du cluster CIRAD – salle 40 (bat. 6) – CIRAD Lavalette
22 Mars	Initiation à l'utilisation du cluster itrop - salle 151 -IRD
15-16 Avril	Initiation au gestionnaires de workflow SG & Gigwa - salle 151 -IRD
18-19 Avril	Guide du Jedi en Linux & bash -salle Bibliothèque (bat. K) – CIRAD Baillarguet
13-16 Mai	Python - salle 151 -IRD
17Mai	Initiation aux analyses de données transcriptomiques - salle 151 -IRD
21 Mai	Utilisation avancée du cluster IRD - salle 151 -IRD
23-24 Mai	Initiation aux analyses de données métagénomiques - salle 151 -IRD
6 Juin	Manipulation de données et figures sous R – salle B07 (bat. B) - CIRAD Baillarguet
26-28 Juin	Assemblage et annotation de transcriptomes - salle 151 -IRD

Formateurs



- Aurore Comte - UMR IPME, i-Trop
- Sébastien Cunnac - UMR IPME
- Valentin Klein - UMR DIADE, i-Trop
- Valérie Noël - UMR MIVEGEC, i-Trop
- Julie Orjuela-Bouniol - UMR BOREA, i-Trop
- François Sabot - UMR DIADE
- Marie Simonin - UMR IPME
- Ndomassi Tando - UMR DIADE, i-Trop
- Christine Tranchant-Dubreuil - UMR DIADE, i-Trop



- Bertrand Pitollat
- Gaëtan Droc
- Gautier Sarah
- Stephanie Sidibe-Bocs
- Marilyne Summo



- Sébastien Ravel
- Frédéric Mahé



- Guilhem Sempere



- Florentin Constancias



- Etienne Loire
- Julie Reveillaud



- Erwan Core



- Catherine Breton
- Alberto Cenci

Guide de survie à Linux	4
Initiation à l'utilisation du cluster bioinformatique CIRAD	5
Initiation à l'utilisation du cluster bioinformatique i-Trop/IRD	5
Initiation aux gestionnaires de workflow South Green (Galaxy et TOGGLE) et à l'application Gigwa	6
Linux avancé et initiation bash	7
Initiation au langage Python	8
Initiation aux analyses de données transcriptomiques	9
Utilisation Avancée d'un Cluster de Calcul	9
Initiation aux analyses de données metabarcoding	10
Manipulation de données et figures sous R	10
Assemblage et annotation des transcrits avec la suite Trinity	11

Guide de survie à Linux

Des commandes de base pour débiter sous un serveur linux aux commandes pour filtrer des fichiers sans connaissances de programmation

Date	18-19/03/19, 9h-17h	Salle 151 - IRD
Formateurs		
Christine Tranchant (DIADE)		Florentin Constancias (Qualisud)
Ndomassi Tando (DIADE)		Valérie Noël (MIVEGEC)
Guilhem Sempéré (Intertryp)		
Prérequis	Aucun	
Programme		
<ul style="list-style-type: none">• <i>ssh, pwd, ls</i>• arborescence sous linux (chemin absolu / relatif) &<ul style="list-style-type: none">○ <i>pwd, ls, cd, mkdir, rm, cp, ln -s, wget</i>○ Metacaracteres : *, [ensemble], [!ensemble]• Fichiers : <i>less, head, tail, wc, sort, cut, unique</i>• Connexion cluster : <i>ssh, scp</i>• Lancement d'une analyse bioinfo (blast, alignement...)• Manipulation de fichiers de sortie<ul style="list-style-type: none">○ <i>less, head, tail, wc</i>○ Recherche : <i>grep</i>○ <i>sort, cut, uniq</i>• Notion E/S : Redirection dans un fichier, vers une autre commande		

Initiation à l'utilisation du cluster bioinformatique CIRAD

Date	21/03/19 - 9h00 - 17h00	Bâtiment 6, salle 40 - CIRAD Lavalette
Formateurs		
Bertrand Pitollat (AGAP)		Gaëtan Droc (AGAP)
Ndomassi Tando (DIADE)		
Prérequis	Connaître les commandes de base en linux	
Programme		
<ul style="list-style-type: none">• Architecture infrastructure• Arborescence• SGE (qsub, qrsh, qhost, qacct, qstat,qdel)• Modules environnement• bonnes pratiques• Scripting de base : config sge, transfert de fichiers, utilisation des modules environment		

Initiation à l'utilisation du cluster bioinformatique i-Trop/IRD

Date	22/03/19 -9h00 - 17h00	Salle 151 - IRD
Formateurs		
Ndomassi Tando (DIADE)		Valérie Noël (MIVEGEC)
Bertrand Pitollat (AGAP)		Valentin Klein (DIADE)
Prérequis	Connaître les commandes de base en linux	
Programme		
<ul style="list-style-type: none">• Architecture infrastructure• Arborescence• SGE (qsub, qrsh, qhost, qacct, qstat,qdel)• Modules environnement• bonnes pratiques• Scripting de base : config sge, transfert de fichiers, utilisation des modules environment		

Initiation aux gestionnaires de workflow South Green (Galaxy et TOGGLE) et à l'application Gigwa

Date	15-16/04/19 - 9h - 17h	Salle 151 - IRD
Formateurs		
Aurore Comte (IPME)		Jean François Dufayard (AGAP)
Marilyne Summo (AGAP)		Julie Orjuela-Bouniol (BOREA)
Sébastien Ravel (BGPI)		François Sabot (DIADE)
Guilhem Sempéré (Intertryp)		
Prérequis	Connaître les commandes de base en linux (+ conseillé utilisation cluster)	
Programme		
<ul style="list-style-type: none">• “Détection de SNP” via les gestionnaires de workflows Galaxy & TOGGLE• Analyse de données SNP sous Gigwa. <p>Gigwa est un outil graphique permettant de gérer, interroger et exploiter d’importants volumes de données de génotypage d’une manière conviviale et rapide. Il se présente, sous la forme d’une application web, qui peut être aussi bien utilisée sur un poste local que déployée sur des serveurs pour maintenir à disposition ce type de données, les partager, les explorer en temps réel.</p>		

Linux avancé et initiation bash

Date	18-19/04/19 - 9h - 17h	Bibliothèque, Bat K - CIRAD Baillarguet
Formateurs		
Christine Tranchant (DIADE)		Valérie Noël (MIVEGEC)
Gautier Sarah (AGAP)		François Sabot (DIADE)
Frédéric Mahé (BGPI)		
Ndomassi Tando (DIADE)		
Prérequis	Connaître les commandes de base en linux	
Programme		
<ul style="list-style-type: none">• Manipuler des fichiers avec commande sed, awk• Apprendre à écrire des scripts basiques Shell (Bash) et d'utiliser la ligne de commande de façon optimale pour automatiser une série d'opérations• Ecrire des scripts simples pour lancer la même analyse sur plusieurs fichiers• Effectuer le debugging d'un script shell		

Initiation au langage Python

Date	13 au 16 mai 2019 - 9h - 17h (4 jrs)	salle 151 IRD
Formateurs		
Etienne Loire (ASTRE)		Valérie Noël (MIVEGEC)
Sébastien Ravel (BGPI)		Julie Orjuela-Bouniol (BOREA)
Valentin Klein (DIADE)		
Prérequis	Connaître les commandes de base en linux	
Programme		
Introduction à la programmation en utilisant le langage Python. A la fin de cette formation les participants pourront écrire des programmes en python et manipuler des données biologiques et comprendre les programmes plus complexes écrits par d'autres.		
Les points suivants seront abordés:		
<ul style="list-style-type: none">• Structures de données (listes, dictionaries)• Typages des données• Structures de controle (conditions et boucles)• Fonctions• Règles d'écriture• Mise en pratique d'écriture et exécution des programmes• Gestion d'options et création/edition des fichiers• Expressions regulieres• Lecture/ecriture des fichiers bioinformatiques (csv, fasta, sam...)		

Initiation aux analyses de données transcriptomiques

Date	17/05/2019 - 9h-17h	salle 151 IRD
Formateurs		
Catherine Breton (Bioversity)		Julie Orjuela-Bouniol (BOREA)
Gautier Sarah (AGAP)		Sébastien Cunnac (IPME)
Prérequis	Connaître les commandes de base en linux + TOGGLE et Galaxy	
Programme		
<ul style="list-style-type: none">• Méthodes avec alignement avec TOGGLE: hisat2, stringtie• Méthodes de quasi-alignement avec Galaxy : kallisto, salmon• Normalisation et expression différentielle de gènes sous Galaxy et R: edgeR, Deseq2• Clustering, réseaux de co-expression: pivot, R		

Utilisation Avancée d'un Cluster de Calcul

Date	21/05/19 -9h00 - 17h00	salle 151 IRD
Formateurs		
Ndomassi Tando (DIADE)		François Sabot (DIADE)
Bertrand Pitollat (AGAP)		Sébastien Ravel (BGPI)
		Valérie Noël (MIVEGEC)
Prérequis	Connaître les commandes de base en linux, savoir lancer des jobs sur un cluster	
Programme		
<ul style="list-style-type: none">• Savoir lancer des jobs array• Créer son propre module environment• Utiliser un conteneur singularity• Installer singularity sur sa machine (linux)• Créer ses propres conteneurs singularity• Utiliser ses conteneurs sur un cluster ayant singularity		

Initiation aux analyses de données metabarcoding

Date	23/05-24/05/2019 - 9h - 17h	salle 151 IRD
Formateurs		
Frédéric Mahé (BGPI)		Julie Orjuela-Bouniol (BOREA)
Marie Simonin (IPME)		Julie Reveillaud (ASTRE)
Florentin Constancias (Qualisud)		Guilhem Sempéré (Intertryp)
Prérequis	Connaître les commandes de base en linux + Galaxy et R	
Programme		
<ul style="list-style-type: none">• Introduction au metabarcoding• Pre-processing, clustering, affiliation taxonomique (FROGS) Manipulation et visualisation des tables d'OTUs avec le package PhyloSeq (sur FROGS et sur R)• Génération des ASV (Amplicon Sequences Variants) tables avec DADA2• Manipulation et visualisation des tables d'ASV sur R		

Manipulation de données et figures sous R

Date	6 juin	Salle B07, Batiment B - CIRAD Baillarguet
Formateurs Frédéric Mahé (BGPI) Etienne Loire (ASTRE) Florentin Constancias (Qualisud)		
Prérequis	Avoir installé R et R studio (pour les personnes venant avec leur propre ordinateur)	
Programme <ul style="list-style-type: none">• Introduction à R,• Introduction à R studio et au "tidyverse",• Chargement de données avec readr,• Manipulation de données avec dplyr et tidyr,• Aperçu des packages lubridate (données temporelles), stringr (chaînes de caractère) et forcats (manipuler des données qualitatives),• Créer de belles figures avec ggplot et ses extensions		

Assemblage et annotation des transcrits avec la suite Trinity

Date	26/06-28/06/2019 - 9h-17h	Salle 151 - IRD
Formateurs		
Erwan Corre (Plateforme AbiMs Roscoff), Gautier Sarah (AGAP) Julie Orjuela-Bouniol (BOREA) Stephanie Sidibe-Bocs (AGAP)		
Prérequis	Connaître les commandes de base en linux	
Programme		
Contrôle de qualité et nettoyage des reads		
Assemblage de novo des transcrits avec Trinity		
<ul style="list-style-type: none">• Évaluation de la qualité de l'assemblage• Identification des transcrits différentiellement exprimés (DE)		
Annotation fonctionnelle des transcrits avec Trinotate et prédictions des régions codantes avec TransDecoder		