



# Bilan 2022

## Effectifs du CATI

- Sortie 2022
  - **Mathnum** : Clément Birbes – Arnaud Di Franco - Andrea Dreau - Jean Mainguy - Annick Moisan - Paul Terzian - Gabryelle Agoutin
  - **GA** : Sandrine Contenot Cagnon
- Entrée 2022 : Jean-Marc Frigerio (UMR BIOGECO – Bordeaux – Dept ECODIV) anciennement membre CATI PROSODIe - Maïna Vienne (MathNum)
- 45 membres actifs

## Réunions mensuelles

- 25 Janvier : Luc Jouneau : “Présentation iSEE pour l’exploration de données scRNA-Seq”
- 14 Mars : Valentin Costes : “Méthylome spermatique et prédiction fertilité des taureaux”
- 14 Avril : Projet d’un site web
- 20 Mai : Points d’organisation ; Organisation AG ; Projets soumis à la DIPSO
- 6 Septembre : Luc Jouneau : “Marmite Norvégienne : un exemple low-tech d’économie d’énergie”
- 17 Octobre : Gérald Salin et Elise Maigne : “Bilan émission GES MIAT et plate-forme Génomique”
- 17 Novembre : Christophe Klopp/Ludovic Duvaux : Etat de l’art Assemblage de génomes
- 9 Décembre : Michelle Halstead : Séquençage et analyse Cut&Tag (réporté)

## Actions inter-CATI

- Séminaire CATI/PEPI/Pépinière numérique 2 membres participants (Claire Hoede et Philippe Bordron)
- Soumission projets à la DIPSO – Demande de soutien à projets innovants et structurants 2022
- Co-construction du 3<sup>ème</sup> Hackathon inter-CATIs Omiques 2022 (Claire Hoede) (financement reçu de 14346€ - complément de budget par les 5 CATIs + soutien prise en charge du transport par le CATI BIOS4Biol pour les membres du CATI BIOS4Biol)

- 3<sup>ème</sup> Hackathon inter-CATI Omiques 2022 qui s'est tenu à Aussois du 5-8 décembre. 8 participants (Bordron Philippe; Duvaux Ludovic; Morin emmanuelle; Guyomar Cervin; Mainguy Jean; Darnige Eden; Sabban Jules; Tournayre Jérémy)
- Projet SK8 : Mise en place d'un service d'hébergement d'applications R-Shiny (Hackathon des acteurs du projet ; rencontres inter-CATIs IUMAN – IMOTEP - Formation Kubernetes)

## Avancement projets CATI (feuille de route 2021)

- **omics du Lombric :**

Des scientifiques ont été contactés pour connaître leurs besoins mais ils n'ont pas donné suite pour l'instant. Sans réponse de leur part -> abandon de cette perspective

- **MultiQC:**

Module de e-training sur l'interprétation du rapport multiQC

4 participants (Claire Hoede, Sarah Maman, Yannick Lippi, Cervin Guyomar)

Mise en place d'un GitBook: <https://bios4biol.pages.mia.inra.fr/Help4MultiQC/> et d'un suivi des réunions et des actions: [https://forgemia.inra.fr/bios4biol/Help4MultiQC/-wikis/Suivi\\_projet](https://forgemia.inra.fr/bios4biol/Help4MultiQC/-wikis/Suivi_projet)

Utilisation du support e-learning "FastQC : outil de contrôle et d'interprétation de la qualité des séquences" disponible sur le site de la FPN

Autres supports utilisés:

<https://elearning.formation-permanente.inrae.fr/course/view.php?id=196&section=1>

<https://github.com/nf-core/rnaseq/blob/master/docs/output.md#quality-control>

[https://github.com/hbctraining/Intro-to-rnaseq-hpc-salmon/blob/master/lessons/qc\\_fastqc\\_assessment.md](https://github.com/hbctraining/Intro-to-rnaseq-hpc-salmon/blob/master/lessons/qc_fastqc_assessment.md)

Journée de travail collaboratif le 8 Juin 2022

Réflexion pour déposer un sujet de stage (financé par le CATI).

- **Données simulées**

Projet en panne pour l'instant par manque de temps des personnes directement concernées (Equipe FROGS – Géraldine et Maria). Soumission d'un sujet pour un stage tutoré 2023 au master bioinfo de Toulouse (pas de nécessité de financement).

- **Formations GIT**

Toulouse : co-construit avec Estelle Ancelet de MIAT (28 membres du CATI inscrits)

Jouy : En collaboration avec la FP locale, une formation sur 2 jours a eu lieu les 28 et 29 Juin 2022 (8 participants). Il n'est pas prévu de refaire une nouvelle session de cette formation en 2023.

- Participation du CATI au congrès **ECCB** (12-21 septembre 2022 à Sitges, Espagne)

7 membres du CATI y ont présenté leur poster (Jean Mainguy, Claire Hoede, Maina Vienne, Vincent Darbot, Paul Terzian, Elisa Maigné et Cervin Guyomar ) :

- Joanna Fourquet, Maina Vienne, Jean Mainguy, Vincent Darbot, Pierre Martin, et al.. metagWGS: a workflow to analyse short and long HiFi metagenomic reads. ECCB2022, Sep 2022, Sitges, Spain. <hal-03788263>
- Jean Mainguy, Adrien Castinel, Olivier Bouchez, Sylvie Combes, Carole Iampietro, et al.. Strengths and limits of long read metabarcoding. ECCB 2022, Sep 2022, Meliá Sitges, Spain. <hal-03848228>
- Paul Terzian, Céline Vandecasteele, Christine Gaspin, C et al. Training DeepSignal models to call CpG methylation in pig and quail ONT reads. ECCB 2022, Sep 2022, Meliá Sitges, Spain.
- Cyril Kurylo, Cervin Guyomar, Sarah Djebali and Sylvain Foissac. Improving genome annotations with RNA-seq data: a scalable and reproducible workflow for Transcripts And Genes Assembly, Deconvolution, Analysis (TAGADA). ECCB 2022, Sep 2022, Meliá Sitges, Spain. <https://doi.org/10.7490/f1000research.1119197.1>
- Élise Maigné, Céline Noirot, Jérôme Mariette, et al. ASTERICS: A Tool for the ExploRation and Integration of omiCS data. ECCB 2022, Sep 2022, Meliá Sitges, Spain.

## Avancement projet CATI : construction d'un site web pour le CATI

Principaux acteurs: Jeremy Tournayre et Anne Frambourg.

Objectif: disposer d'un site web pour le CATI pour :

- que les membres du CATI trouvent
  - l'ensemble de la documentation interne à un seul endroit (comptes-rendus, enquêtes, présentations, etc),
  - qui a quelles compétences (trombinoscope)
  - un agenda
  - les actualités

- Autres selon besoins et envies
- que tous (membres du cati, nouveaux arrivants dans le CATI, responsables fonctionnels des agents, départements, dispo etc.) puissent être informés de ce qu'est le CATI et de ce qu'on y fait.

Techno:

- Basé sur le framework wordpress (achat licence achat licence du thème Divi ~250 €).
- Un espace Nextcloud (achat espace de travail, pour l'instant la tarification est suspendue. Avant la suspension le Terra était à 50€/an (sans réplication). A la reprise de la tarification, ce sera au moins 50€/an pour le Terra.),
- une boîte aux lettres partagée
- une machine virtuelle (Marie-Stéphane Trottard, Didier Laborie)

Point d'avancement du projet le 3 février 2023 (1 par trimestre)

## Organisation des journées annuelles 26-28 septembre 2022 (en présentiel à Toulouse)

- Lundi:
  - Accueil + nouveaux arrivants
  - ouverture à de nouveaux pans scientifiques et techniques :
    - Axes de recherche sur les organoïdes - Martin Beaumont (CR - GenPhySE)
    - Analyses de données métabolomiques - Fabien Jourdan (DR - Toxalim)
  - démo d'une techno de construction de site web (Jeremy Tournayre et Valérie Vidal)
- Mardi: journée cohésion d'équipe sur la thématique de la découverte du patrimoine Toulousain
  - rallye découverte; restaurant; croisière en bateau électrique sur la Garonne; visite à la découverte de la violette sur une péniche dédiée; Soirée à Ma biche sur le Toit
- Mercredi matin:
  - bilan des productions de l'année avec les responsables scientifiques du CATI
  - Démarrage de la co-construction des pages du site web du CATI
  - restaurant

36 agents ont participé aux journées.

## Bilan des productions CATI 2022

Les productions du CATI se définissent par des productions qui:

- émanent de discussions/idées/rerelations d'agents au sein du CATI
- sont des projets menés par plusieurs membres du CATI (une production CATI peut être une production issue d'une plateforme)

N.B : Un document en annexe ("Une image - Un projet - AG CATI BIOS4Biol 2022.pdf") détaille les différentes productions.

### **Workflows**

PAQMir : analyse et annotation des petits ARNs (A.Frambourg, L.Jouneau)

Analyse de variants RNA-Seq - NextFlow (M.Bernard, M.Charles, C.Guyomar, S.Maman)

WoodySV - PangenOak : Analyse de variants structuraux par construction de graphes pangénomiques (L.Duvaux)

metaWGS : Ajout d'une fonctionnalité de binning (M.Vienne, V.Darbot, J.Mainguy, C.Noiro, G.Pascal, C.Hoede)

VarPipeline : Découverte incrémental de variants (P.Deiais, P.Legoueix)

FROGSFUNC : Intégration de PICRUST2 pour la prédiction et la quantification fonctionnelle des écosystèmes microbiens (V.Darbot, M.Bernard, G.Pascal)

### **Benchmark**

Test d'outils pour l'affiliation taxonomique des données WMS ou WGS (E.Jausseme, M.Manno, C.Hoede, C.Noiro, J.Mainguy, M.Vienne, C.Kuchly, E.Darnige)

### **Formations**

Formation R avancé : package data.table (P.Terzian, E.Maigné)

Help4MultiQC: rédaction d'un gitbook pour l'aide à l'interprétation des résultats compilés par MultiQCreport (C. Guyomar, C. Hoede, S. Maman, Y. Lippi)

Tutoriels de formation pour FROGS: développement du site web et production de pages de tutoriels des outils. (M. Bernard, L. Auer, V. Darbot, G. Pascal)

Alignement de lectures courtes et recherche de variants de petite taille (C.Noiroto, P.Bardou, C.Cabau)

## **Infrastructure, Système et Base de données**

Acquisition cluster 2022 (MS.Trotard, P.Deiais, D.Laborie)

Remplacement du gestionnaire de ticket OTRS -> Znuzy (G.Salin, D.Laborie, MS.Trotard)

Lancement du service SK8 (serveur Shiny) (E.Maigné)

Evolution LIMS NGL (E.Darnige, C.Kuchly, G.Salin, J.Sabban, C.Noiroto, R.Therville)

Base de données de fichier de données hétérogènes (P.Bardou, C.Cabau, C.Klopp)

## Présentations du CATI à d'autres instances

- Participation aux séminaires DIPSO de présentation des différents CATIs - hiver 2022
- Présentation du CATI au département SA : Luc Jouneau devait assister le comité de direction du département SA dans l'identification de leurs besoins dans le domaine du numérique mais nous n'avons pas eu de nouvelle de la part de SA sur ce point depuis avril 2022
- Participation de Géraldine à la journée du numérique organisée par la DIPSO le 6 septembre 2022 sur le centre de Toulouse.
- Rédaction de deux articles pour le journal interne Reg@rd du département GA.

## Perspectives

- Demande de financement pour Hackathon inter-CATI avec une partie Work4Graph - du 18 au 21 septembre 2023 à Roscoff auprès de l'appel à financement de la dipso.
- Finalisation du site web B4B
- Préparations des journées annuelles 2023, et du hackathon 2023.
- Intervenant prévu pour 2023:
  - janvier: Jérôme Montfort (Plateforme Régionale Transcriptome)
  - février: Point projet site web Jeremy et Anne - avancements + besoins
  - mars: Sylvain Foissac (GenPhySE)
- Aide au financement d'un stage M2 (annotation fonctionnelle de transcrits de champignons) - Lucas Auer (Nancy)
- Stage tutoré M1 sur la simulation de données métabarcoding - Vincent Darbot

