

# Session de formation 2019





**Bioinformatics platform dedicated to the genetics and genomics of tropical and Mediterranean plants and their pathogens**

genome assembly SNP detection  
phylogeny structural variation  
comparative genomics transcriptome assembly differential expression  
GWAS pangenomics  
population genetics metagenomics  
polyploidy



Rice



Banana



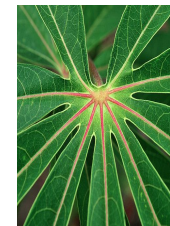
Palm



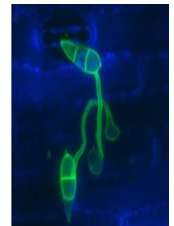
Sorghum



Coffee



Cassava



Magnaporthe



Larmande Pierre  
Sabot François  
Tando Ndomassi  
**Tranchant-Dubreuil  
Christine**



Comte Aurore  
Dereeper Alexis



**Orjuela-Bouniol Julie**



Bocs Stephanie  
De Lamotte Frédéric  
**Droc Gaetan**  
Dufayard Jean-François  
Hamelin Chantal  
Martin Guillaume  
Pitollat Bertrand  
**Ruiz Manuel**  
**Sarah Gautier**  
Summo Marilyne



**Bioversity  
International**

**Rouard Mathieu**  
Guignon Valentin  
Catherine Breton



**Mahé Frédéric**  
**Ravel Sébastien**



Sempere Guilhem



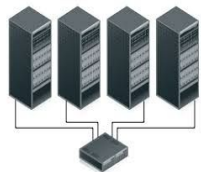
## Workflow manager

TOOLBOX  
Toolbox for generic NGS analyses

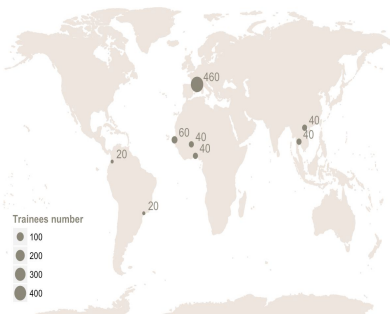
●●●●●  
SNAKEMAKE

Galaxy

## HPC and trainings....



37 courses organized last 7 years



IRD  
Institut de Recherche  
pour le Développement

cirad

## Genome Hubs & Information System



Gigwa

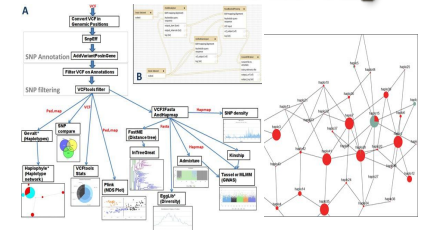
SNPs and Indels

GreenPhyl

Family Id	Family Name	Number of sequences	Status
GP000010	Cytochrome P450 superfamily	6542	●●●
GP000017	AP2/EREBP transcription factor family: ERF/ERF3 group (partial)	5142	●●●
GP000020	NAC transcription factor family	4574	●●●
GP000028	MADS transcription factor family		
GP000018	Hamam peroxidase superfamily		
GP000095	General substrate transporter superfamily		
GP000022	Subtilisin-like Serine Proteases family		
GP000019	NPF, NRT1/PTR FAMILY		

Gene families

SNIPlay



<https://github.com/SouthGreenPlatform>



@green\_bioinfo

# South Green

bioinformatics platform



Erwan Corre



Marie Simonin  
Sébastien Cunnac



Etienne Loire  
Julie Reveillaud



Florentin Constancias



Valentin Klein



Valérie Noël



Emmanuelle Beyne



[@ltropBioinfo](https://twitter.com/ltropBioinfo)



**And more collaborators !**

- 18-19/03 - Guide de survie à Linux - IRD
- 21/03 - Initiation à l'utilisation du cluster CIRAD - CIRAD
- 22/03 - Initiation à l'utilisation du cluster itrop - IRD
- 15-16/04 - Initiation au gestionnaires de workflow SG & Gigwa - IRD
- 18-19/04 - Guide du Jedi en Linux & bash - CIRAD
- 13-16/05 - Python - IRD
- 17/05 - Initiation aux analyses de données transcriptomiques - IRD
- 21/05 - Utilisation avancée du cluster IRD - IRD
- 23-24/05 - Initiation aux analyses de données métagénomiques - IRD
- 6/06 - Manipulation de données et figures sous R - CIRAD
- 25-27/09 - Assemblage et annotation de transcriptomes - IRD



# Modules de formation 2019

- Toutes nos formations :  
<https://southgreenplatform.github.io/trainings/>
- Topo & TP : [Linux For Dummies](#)
- Environnement de travail : [Logiciels à installer](#)

# Assembling and functionally annotating a de-novo transcriptome from RNA-seq data



[www.southgreen.fr](http://www.southgreen.fr)

<https://southgreenplatform.github.io/trainings>





## Introduction

[survival] Basic commands in SLURM : [see](#)

## Reads quality checking

- Practice 1. Checking Reads Quality [see](#)

## Assembly with Trinity

- Practice 2. Assembling transcriptome from RNA-seq [see](#)
- Practice 3. Assessing transcriptome assembly quality [see](#)

## Annotation with Trinotate

- Practice 5 : Functional annotation [see](#)

## Differential Expression

- Practice 4 : Differential Expression Analysis (DE) [see](#)

## Conclusions



Erwan Corre **ABIMS**



Julie Orjuela-Bouniol



Christine Tranchant-Dubreuil



Stéphanie Bocs-Sidibe



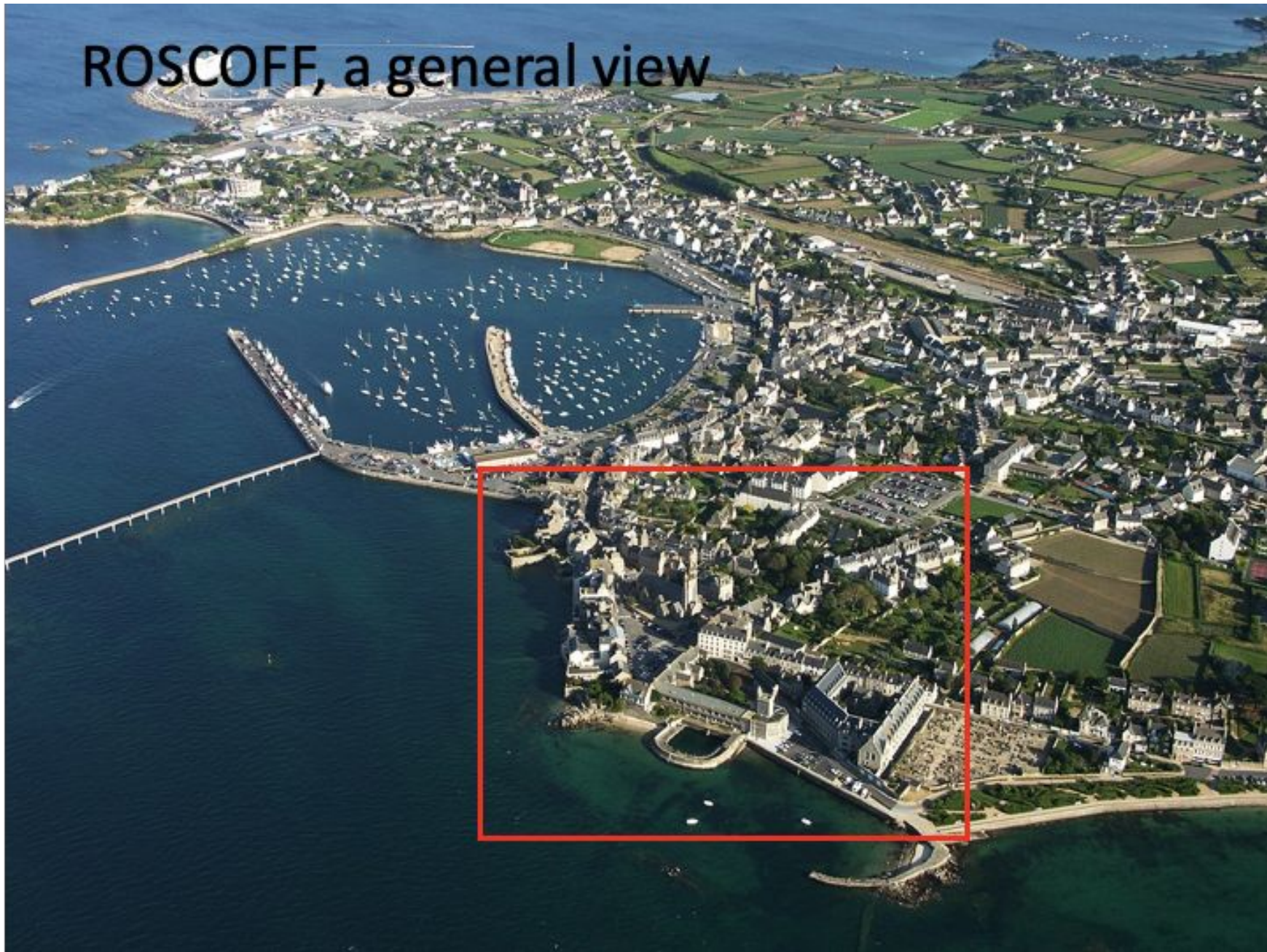
David Lopez



Aurore Comte



# ROSCOFF, a general view





# La Station Biologique de Roscoff



Recherche : 2 UMR, 1 FR, 1 UMI

Observation : INSU

Enseignement

Accueil

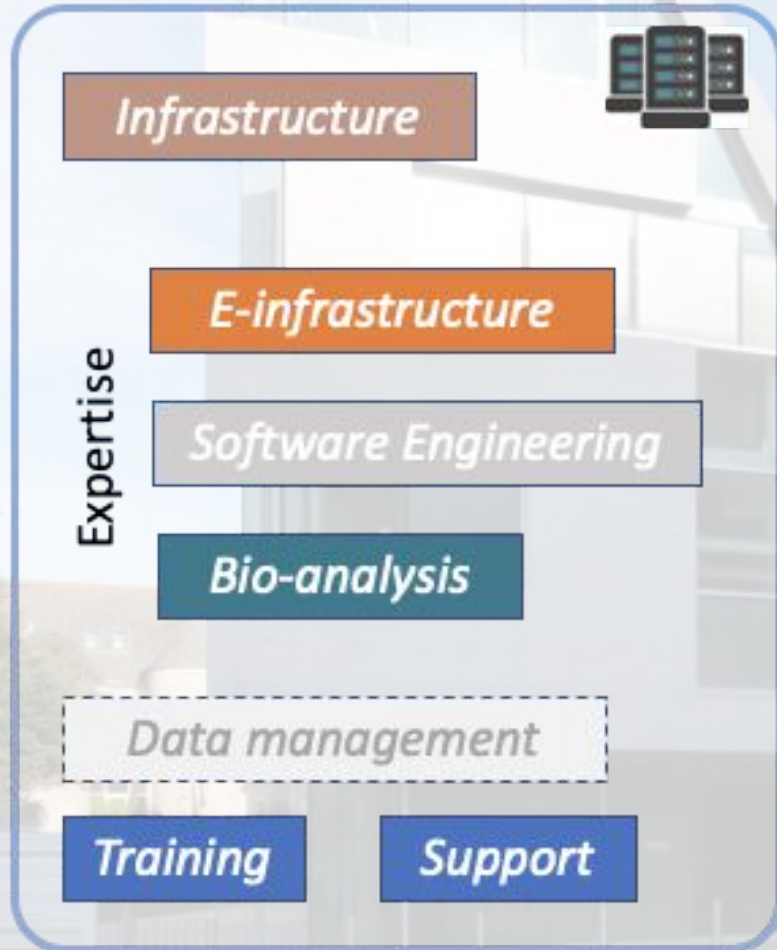
CNRS Sorbonne Université  
300 personnes

## Recherche:

- la biochimie et le développement des algues brunes et rouges,
- l'écophysiologie de la faune hydrothermale et l'adaptation aux milieux extrêmes,
- la diversité du phytoplancton et du zooplancton,
- l'évolution et la génétique des populations,
- l'écologie benthique.
- chimie marine, phénomènes de circulation océanique des masses d'eau.

# ABIMS : Analysis and Bioinformatics for Marine Science

## Activities



## User Communities

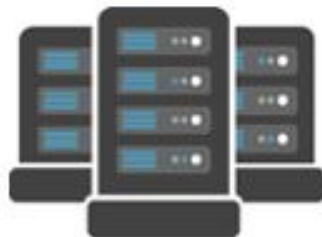


# ABIMS : Analysis and Bioinformatics for Marine Science

2019



Storage : 3 PB



Cluster : 1,900 CPU  
(3<sup>rd</sup> largest french open infrastructure)

900 support tickets/year



Int. Bandwidth : 1 Gb/s  
Ext. Bandwidth : 10 Gb/s



100 internal accounts  
500 external accounts

x 15

## Transcriptome analysis:

Arctic Krill  
Red abd Brown algae  
Marine Pico Eukaryotes  
Radiolaires  
Artemia  
Sepia, Pterois, Heterotis , Saculine  
Sabellaria alveolata  
Habalone

Mucor  
Vertebrates Jaws  
Quercus.  
Canard Foie gras.  
Mouse







# Practice

1

*Aller sur la practice 1 [Checking Reads](#)  
[Quality](#) du github.*



# Practice

2

*Aller sur la practice 2 [Assembling transcriptome from RNA-seq](#) du [github](#).*



# Practice

3

*Aller sur la practice 3 [Assessing transcriptome assembly quality](#) du github.*



# Practice

5

*Aller sur la practice 5 [Functional annotation](#) du github.*



# Practice

4

*Aller sur la practice 4 [Differential Expression Analysis \(DE\)](#) du [github](#)*



# Merci pour votre attention !

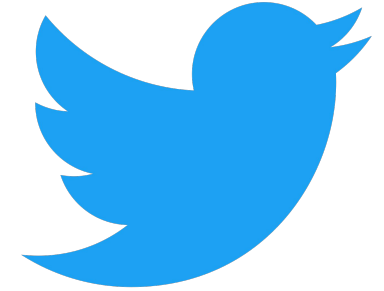


Le matériel pédagogique utilisé pour ces enseignements est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions (BY-NC-SA) 4.0 International:

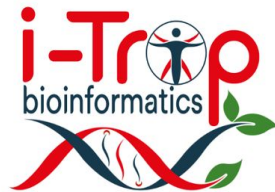
<http://creativecommons.org/licenses/by-nc-sa/4.0/>



**SUIVEZ NOUS SUR TWITTER !**



South Green : [@green\\_bioinfo](https://twitter.com/green_bioinfo)



I-Trop : [@ltropBioinfo](https://twitter.com/ltropBioinfo)



**N'oubliez pas de nous citer !**

### **Comment citer les clusters?**

"The authors acknowledge the IRD i-Trop HPC at IRD Montpellier for providing HPC resources that have contributed to the research results reported within this paper. URL: <http://bioinfo.ird.fr/> "

"The authors acknowledge the CIRAD UMR-AGAP HPC (South Green Platform) at CIRAD montpellier for providing HPC resources that have contributed to the research results reported within this paper. URL: <http://www.southgreen.fr>"