



# Workflow4Metabolomics: Towards an international computing infrastructure and a tools showcase for Metabolomics

Gildas Le Corguillé, Franck Giacomoni, Christophe Caron

## ► To cite this version:

Gildas Le Corguillé, Franck Giacomoni, Christophe Caron. Workflow4Metabolomics: Towards an international computing infrastructure and a tools showcase for Metabolomics. GCC2017 Galaxy Community Conferences, Jun 2017, Montpellier, France. 2017. hal-01604843

**HAL Id: hal-01604843**

**<https://hal.science/hal-01604843>**

Submitted on 2 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License



# 2017 Galaxy Community Conference

Le Corum  
Montpellier, France  
26-30 June 2017

Hack the Galaxy: 26-27 June  
Training: 27-28 June  
Conference: 29-30 June

<https://gcc2017.scienceconf.org/>  
<https://gcc2017.sched.com/>



#usegalaxy @galaxyproject

## Sponsors



INDIANA UNIVERSITY



<https://gcc2017.sciencesconf.org/page/sponsors>

# Bienvenue !

Welcome to the 2017 Galaxy Community Conference (GCC2017), the eighth annual gathering of the Galaxy Community. GCC2017 features two days of training, two days of Hack the Galaxy, and a two day meeting with accepted and keynote talks, posters, vendor exhibits, birds-of-a-feather (BoF) gatherings, demos, and lightning talks - all about high-throughput biological research and its supporting compute infrastructure. This event also features plenty of time for networking and impromptu gatherings.



We would like to give an enormous thank you to our sponsors and hosts, our instructors, the Scientific Committee, BoF organisers, speakers, demo and poster presenters, and to everyone who contributed to making this event a success.

Merci beaucoup,

The GCC2017 Organising Committee

## Hosts





# Institut Français de Bioinformatique



**IFB consolidates French National centres, services and bioinformatic resources into one Federation**

**IFB** is accelerating life sciences research by delivering tools, data collection, training and **Cloud** Infrastructure for collaborative and reproducible science

**IFB** is the French node of the **European research Infrastructure ELIXIR**

**IFB** contributes to the vibrancy of **Galaxy** community by coordinating the French Galaxy Working Group



# Network

Wifi is available throughout all conference facilities.

Wifi-name: GCC2017

Password: galaxy2017

# Social Media

Unless requested by the speaker, Tweeting and other social media activity are actively encouraged. Post early, post often.



#usegalaxy

@galaxypoint

# Meals

The official conference dinner is Thursday evening. It will be held at Le Château du Pouget. Buses will make the journey from Le Corum.

Coffee, other beverages, and light snacks will be available during breaks. Lunch is provided on the first day of Hack the Galaxy, and during the conference. On Tuesday and Wednesday participants can grab lunch in the city center.



# Getting Around

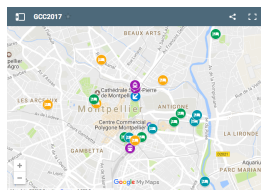


GCC2017 takes place at Le Corum Conference Centre. Le Corum is located in the center of Montpellier, and is easily reached on foot or via public transportation.



*See the Conference Centre floor plan on the last page for a map of where events are.*

Suggested lodging and several social events are elsewhere around the city centre. A custom Google Map with all conference venues and suggested lodging is also available.



If you are staying in the CROUS student lodging La Colombière, you will want to take the TaM Blue line (#1) from the Hôpital Lapeyronie station to the Corum station. See <http://www.tam-voyages.com/> for information about all TaM lines.

*See the map on the second to last page, and <http://bit.ly/gcc2017map>.*



# INRA

## SCIENCE & IMPACT

Galaxy: advancing the digital  
agricultural revolution.

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

# GCC2017 Events Overview

Mon 26 June	Tue 27 June	Wed 28 June	Thu 29 June	Fri 30 June
<b>Hack the Galaxy Dev &amp; User Hacks</b> 2 intense days of Galaxy development & data analysis		<b>Training</b> 5 tracks, 15 sessions: covering the full spectrum of Galaxy usage, admin & development	<b>GCC2017 Meeting</b> 2 days of accepted talks, poster sessions, birds-of-a- feather, lightning talks, exhibits, keynotes, and networking.	
	<b>Training</b> Full day, single track, using Galaxy			<b>Gala Dinner</b>

## ELIXIR: A DISTRIBUTED INFRASTRUCTURE FOR LIFE-SCIENCE DATA

ELIXIR unites Europe's leading life science organisations in managing and safeguarding the increasing volume of data being generated by publicly funded research. It coordinates and sustains bioinformatics resources across 20 European countries.


[@ELIXIREurope](https://twitter.com/ELIXIREurope)
[www.elixir-europe.org](http://www.elixir-europe.org)

## ELIXIR support to Galaxy users and developers

- **ELIXIR Galaxy Working Group:** a network of Galaxy administrators, developers and users
  - **Developing training materials** as part of the Galaxy Training network
  - Collecting training materials through the ELIXIR Training Portal (TeSS)
  - Building a Galaxy ecosystem with **integrated and interconnected resources** for Tools, Workflows and Services
  - Establishing system-agnostic, reproducible and executable environments for bioinformatics software with **BioConda** and **BioContainers**
- Organising training courses and community events for Galaxy users:
- **21 July 2017, Prague, Czech Republic**  
Making Galaxy Work for You (ISMB/ECCB tutorial)
  - **8-12 January 2018, Oslo, Norway**  
Galaxy Admin Workshop (Registration opens in Autumn 2017)

Visit ELIXIR Galaxy Working Group page to find out more:  
[www.elixir-europe.org/galaxy-wg](http://www.elixir-europe.org/galaxy-wg)



# Hack the Galaxy

There are two parallel and co-located *Hack the Galaxy* events at GCC2017. The *Dev Hack* focuses on extending Galaxy code and capabilities, while the *User/Data Hack* focuses on addressing challenging data integration and analysis problems.

*Hack the Galaxy* runs for two days. **Preregistration is required to attend.**

## Monday 26 June: Hack the Galaxy

Time	Sully 1 & Sully 2 (Level 1)
08:00	Conference Desk Open (until 22:00)
09:00	<b>Introductions, Project Ideas, Brownian Motion, Coalescing</b>
10:00	<b>Hack the Galaxy (Dev or Data)</b>
12:30	Catered lunch
13:30	<b>Hack the Galaxy</b>
17:30	Dinner on your own
<b>Continued hacking at Le Corum or the venue of your choice</b>	
22:00	Le Corum Closes

## Tuesday 27 June: Hack the Galaxy

Time	Sully 1 & Sully 2 (Level 1)
08:00	Conference Desk Open (until 18:00)
09:00	<b>Group Reports</b>
09:40	<b>Hack the Galaxy (Dev or Data)</b>
12:30	Lunch on your own
13:30	<b>Hack the Galaxy</b>
18:00	Le Corum Closes
18:30	<b>Dinner on your own</b> <b>Continued hacking at venue of your choice</b>

# Training

Two full days of training are offered. The first day offers a single track focusing on how to use Galaxy to address biological questions. The second day offers five tracks and 15 topics covering Galaxy usage, development, administration and tool packaging.

Training is hands-on. **Please bring an Ethernet-capable laptop. Ethernet cables and power are provided at each desk. (Wi-fi will be available, but discouraged.)**

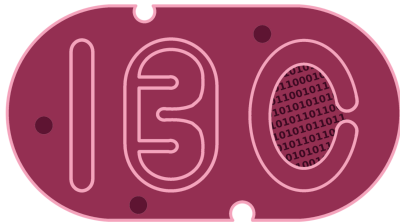
## Tuesday 27 June: Using Galaxy Training

Time	Barthez (Level 2)
08:00	Conference Desk Open (until 18:00)
09:00	<b>Galaxy 101: A Gentle Introduction to Galaxy</b> Mallory Freeberg, Mo Heydarian
11:30	Lunch on your own
12:30	<b>RNA-Seq Analysis in Galaxy</b> Mo Heydarian, Mallory Freeberg
15:00	Break
15:30	<b>Visualisation of BIG DATA in Galaxy</b> Aysam Guerler, Saskia Hiltermann, Jeremy Goecks
18:00	Done
18:30	Dinner on your own



# Wednesday 28 June: Galaxy Training

Time	Barthez (Level 2)	Rondolet (Level 2)	Sully 1 (Level 1)	Sully 2 (Level 1)	Sully 3 (Level 1)
08:00	Conference Desk Open (until 18:00)				
09:00	Galaxy Architecture John Chilton, Nate Coraor	Writing & Publishing Galaxy Tools M van den Beek, D Blankenberg, D Bouvier, M Čech, G Le Corguillé, N Soranzo, N Stoler	RNA-Seq Analysis in Galaxy Mo Heydarian, Mallory Freeberg	Small Genome de novo Assembly Using Galaxy Simon Gladman	Introduction to Galaxy Admin: Setting up Galaxy as a Service Jochen Bich, Trond Thorbjørnsen, Nikolay Aleksandrov Vzov
11:30	Lunch on your own				
12:30	Advanced Customisation of a Galaxy Instance Björn Grünig, Erick Rasche, Martin Čech	Conda & Containers for Tool Dependencies – a Developer's Perspective Abdulrahman Azab, Marius van den Beek, John Chilton, Nicola Soranzo	Visualisation of BIG DATA in Galaxy Aysam Guerler, Jeremy Goecks	How to Analyze Metagenomics & Amplicon in Galaxy Bérénice Batut, Saskia Hiltmann	Galaxy for Proteomics Clemens Blank, Timothy Griffin, Pratik Jagtap, James Johnson
15:00	Break				
15:30	Advanced Accelerated Galaxy Admin Enis Afgan, Dan Blankenberg, Martin Čech, John Chilton, Nate Coraor, Simon Gladman	Galaxy for Training & Education Bérénice Batut, Björn Grünig	ChIPseq Analysis Using DeepTools & Macs Devon Ryan	Visualization Development in Galaxy Jeremy Goecks, Aysam Guerler, Saskia Hiltmann	SNP Analysis & GWAS Using SNIPlay Alexis Dereeper, Yann Hueber



# Institut de biologie computationnelle

*IBC is a multidisciplinary research institute focused on Computational Biology, gathers circa 100 researchers from several disciplines: life sciences, computer science and math. It was launched in 2012 for promoting interdisciplinary thinking to address critical biological questions.*

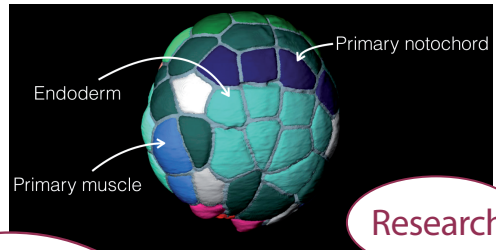
One industrial member and 8 national trustees

**Scientific domains:** bioinformatics, molecular biology, medicine, agronomy, evolutionary biology, plant & animal biology, math & computer science.

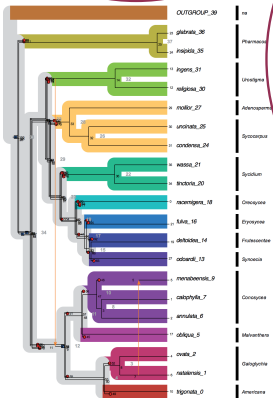
**Overarching aim:** Designing computational methods for analysing and modelling massive data in biology, health, agronomy and environment.

## Biological applications

- genomics (eg. vertebrate, rice, cancer)
- transcriptomics
- virus and pathogens (HIV, malaria, etc.)
- embryonic development and stem cells
- population genetics
- Next Generation Sequencing analysis tools
- ontologies for agronomy & medicine
  - phylogenetic inference
  - protein structure



## Education

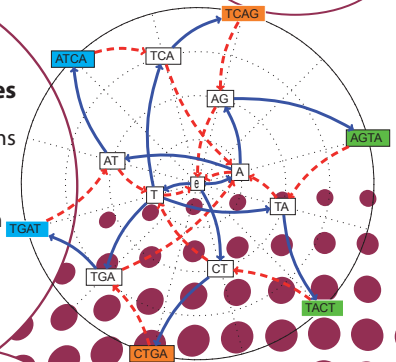


## Computational techniques

- algorithms on sequences & graphs
- statistical inference
- machine learning
- modelling for biology & evolution
- image processing
- heterogeneous databases
  - information retrieval
  - big data

## Animation

## Research



**Contacts:** Eric Rivals  
 ibc-contact@lirmm.fr  
 04 67 41 86 64



# Conference: Thursday 29 June

Time	Einstein Auditorium (Level 0)
08:00	Conference Desk Open (until 19:00)
09:00	<b>Opening and Welcome</b> Olivier Inizan, INRA; Marilynne Summo, CIRAD
09:15	<b>Session 1: Galaxy Framework</b> Moderator: Jeremy Goecks, OHSU
09:15	<b>Keynote: The Galaxy Project: from the Planck Epoch to Montpellier</b> Anton Nekrutenko, Penn State University; James Taylor, Johns Hopkins University
10:00	<b>Bioconda and BioContainers - Enabling Universal Reproducibility in Galaxy</b> John Chilton, Penn State University
10:20	<b>Building an open, collaborative, online infrastructure for bioinformatics training</b> Bérénice Batut, University of Freiburg
10:40	Break
11:10	<b>Session 2: Galaxy Framework (continued)</b> Moderator: Dave Clements, Johns Hopkins University
11:10	<b>An Educational Framework for Galaxy Administration</b> Nate Coraor, Penn State University; Simon Gladman, Melbourne Bioinformatics / University of Melbourne
11:30	<b>Integrative analysis of hundreds of datasets in Galaxy</b> Geir Kjetil Sandve, University of Oslo
11:50	<b>Galaxy Genome Annotation Project: Galaxy and GMOD for Annotation, Teaching, and Genomic Databases</b> Eric Rasche, Texas A&M University
12:10	<b>Apollo in Galaxy: Increasing Opportunities for Collaborative Genome Annotation</b> Nathan Dunn, Lawrence Berkeley National Lab (LBNL)
12:30	<b>Galaxy at the IFB: From computing to training</b> Victoria Dominguez del Angel, Institut Français de Bioinformatique; Christophe Caron, INRA

## Conference: Thursday 29 June (continued)

Time	Einstein Auditorium (Level 0)
12:45	<b>Lunch (and Birds-of-a-Feather)</b>
13:45	<b>Session 3: Tools &amp; Workflows</b> Moderator: Nicola Soranzo, Earlham Institute
13:45	<b>Galaxy Community Update</b> Jeremy Goecks, Oregon Health & Science University; Dan Blankenberg, Penn State University
14:25	<b>The Year of IUC</b> Björn Grüning, University of Freiburg
14:40	<b>Workflow4Metabolomics: Towards an international computing infrastructure and a tools showcase for Metabolomics</b> Gildas Le Corguillé, ABiMS; Franck Giacomoni, PFEM
15:00	<b>South Green Galaxy: a suite of tools for plant genomics</b> Jean-François Dufayard, CIRAD UMR AGAP
15:20	<b>Performing Next Generation Phylogenetic Analyses with NGPhylogeny.fr</b> Damien Correia, Institut Français de Bioinformatique
15:40	<b>Sponsor, Poster and Demo Session</b>



premier

## Conference: Thursday 29 June (continued)

Time	Einstein Auditorium (Level 0)
17:00	<b>Session 4</b> Moderator: Bérénice Batut, University of Freiburg
17:00	<b>The Institute of Computational Biology: an overview</b> Eric Rivals, The Institute of Computational Biology
17:30	<b>Lightning Talks</b> <b>Improving Your Text Life</b> James Johnson, University of Minnesota Supercomputing Institute (MSI) <b>Galaxy-E : Towards an accessible, reproducible and transparent data analysis and management universe dedicated to Ecology</b> Valentin Chambon, Thimothée Virgoulay, MNHN CESCO & Staiton de Biologie Marine de Concarneau <b>The Galaxy Gateway on F1000Research</b> Holly Murray, F1000, Science Navigation Group <b>GalaxyCat - The online catalog of Galaxy tools across instances</b> Julien Seiler, Stéphanie Le Gras, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS <b>Galaxy at scale: Analyzing thousands of single cell transcriptomes</b> Mo Heydarian, Johns Hopkins University
18:00	<b>Done (and Birds-of-a-Feather)</b>
19:00	<b>Gala Dinner @ Chateau du Pouget</b>




# Conference: Friday 30 June

Time	Einstein Auditorium (Level 0)
08:00	Conference Desk Open (until 18:00)
09:00	<b>Session 5: Galaxy for Life Science and Medicine</b> Moderator: Christophe Antoniewski, Institut de Biologie Paris-Seine
09:10	<b>Keynote: Personalized Analysis of Cancer Data: From Genes to Pathways (and Back)</b> Eytan Domany, Weizmann Institute of Science
10:00	<b>Comparing hundreds of RNA-seq libraries using Galaxy and SeqResults</b> Brad Langhorst, New England Biolabs
10:20	<b>Approaches for small RNA-seq analysis in Galaxy</b> Mallory Freeberg, Johns Hopkins University
10:40	Break
11:10	<b>Session 6: Galaxy for Life Science and Medicine</b> Moderator: Thérèse Commes, INSERM
11:10	<b>COMBAT TB: an integrated environment for Tuberculosis data analysis and surveillance</b> Peter Van Heusden, South African National Bioinformatics Institute (SANBI)
11:30	<b>The RNA Workbench - reproducible, transparent, and accessible RNA research</b> Florian Eggenhofer, University of Freiburg
11:50	<b>Interaction Hub: a comprehensive Galaxy-based chromatin folding database, browser, and analysis platform</b> Michael Sauria, Johns Hopkins University
12:10	<b>Mechanizing our biologists with Galaxy and IRIDA platforms</b> Philip Mabon, Public Health Agency of Canada
12:40	Lunch (and Birds-of-a-Feather)

# Conference: Friday 30 June (continued)

Time	Einstein Auditorium (Level 0)
13:45	<b>Session 7: Galaxy in today's infrastructures</b> Moderator: Björn Grüning, University of Freiburg
13:45	<b>ELIXIR supports Galaxy users and developers</b> Frederik Coppens, ELIXIR Belgium & VIB
14:00	<b>Advancing metaproteomics research via community-based informatics development and dissemination</b> Pratik Jagtap, University of Minnesota
14:20	<b>Making Galaxy User Interface Pluggable with Webhooks</b> Martin Čech, Penn State University; Björn Grüning, University of Freiburg
14:40	<b>A reproducible data analysis environment for next-generation sequencing on public cloud computer</b> Manabu Ishii, RIKEN ACCC Bioinformatics Research Unit
15:00	<b>Developing analytical workflows for plant genomics through the integration of Galaxy and Tripal</b> Sean Buehler, University of Connecticut
15:20	<b>Sponsor, Poster and Demo Session</b>

Looking for something?  
<https://galaxyproject.org/search>

 Use ~ Community ~ Education ~ Deploy & Develop ~ Support ~

### Searching the Galaxy


Welcome to the comprehensive Galaxy search interface. This is a Google Custom Search Engine that will search all things Galaxy. This includes all the pages on the Hub, the codebase (Galaxy and related projects), the mailing lists, biostars, and more.

If you haven't already, enter your query into the box in the masthead, and results will appear below.


All results   Use   Hub   Corpus   Code   Admin/Dev   Discussion   Shared   Tools   Requests

About 8,910 results (0.42 seconds)


Did you mean: [gcc 2017](#)



**Galaxy Community Conference 2017 - Sciencesconf.org**  
GCC2017 will include keynote talks and exhibitors, and plenty of networking opportunities. There are also three days of pre-conference activities, including ...  
<https://gcc2017.sciencesconf.org/>  
Labeled [Admin/Dev](#) [Use](#)



**GCC2017 Call for Host Proposals**  
Sep 14, 2015 ... Galaxy's informal policy of rotating between North America and elsewhere every other year is now a formal policy: Hosts for GCC2017 need to ...  
<https://galaxyproject.org/news/gcc2017-call-for-host/>  
Labeled [Admin/Dev](#) [Hub](#) [Use](#)



**2017 Galaxy Community Conference (GCC2017)**  
GCC2017 is the 8th annual gathering of the Galaxy community. The conference will include keynotes and accepted talks, poster sessions, demos, ...  
<https://galaxyproject.org/events/2017-gcc/>  
Labeled [Admin/Dev](#) [Hub](#) [Use](#)

# Conference: Friday 30 June (continued)

Time	Einstein Auditorium (Level 0)
------	-------------------------------

16:35

## Session 8

Moderator: Virginie Rossard, INRA

### 16:35 Lightning Talks

**Using visual programming in Cervical cancer data of The Cancer Genome Atlas**

Thais Hosokawa, Barretos Cancer Hospital

**Analyzing Relationships among Clinical Factors of Cancer Patients by Regression Analysis**

Sugandima Vidanagamachchi, Tamara Waidyarathna, University of Ruhuna

**ToolDog - generating tool descriptors from the ELIXIR tool registry**

Kenzo-Hugo Hillion, Center of Bioinformatics, Biostatistics and Integrative Biology (C3BI), Institut Pasteur de Paris

**Pushing the limits of job flexibility on HPC**

Carrie Ganote, Indiana University

**Pulling the Galaxy's Strings**

Jeffrey Miller, University of Iowa

**Using Galaxy as a platform for continuous software development**

Geir Kjetil Sandve, Boris Simovski, University of Oslo

**Automated Generation of Complex Toolkits for Galaxy**

Daniel Blankenberg, Penn State / Galaxy Team

**Hack the Galaxy: Data Report**

Organizers

**Hack the Galaxy: Dev Report**

Organizers

**JBrowse & Circos Tool Updates**

Eric Rasche, Texas A&M University

**Parsec: Intergalactic DevOps and Analysis Automation**

Eric Rasche, Texas A&M University

### 17:42 GCC2018 Launch

Dave Clements, Johns Hopkins University

### 17:50 GCC2017 Closing

GCC2017 Organisers

18:00

Fini

# Lightning Talks

Topics for lightning talks were solicited before the meeting. They will be presented during Session 4, on Thursday and Session 8 on Friday. There may also be last minute openings, and if you wish to give a lightning talk, please send a proposal to the GCC2017-Scientific Program Committee <gcc\_2017\_scientific@groupe.fr> before the start of Session 2 (Thursday) or the start of Session 6 (Friday). The slides for all lightning talks will be made available on the conference web site, and the talks may be videotaped and also posted on the conference web site.

## Goals

This is your opportunity to give an impassioned and enthralling talk about something that you care about - but you only have 300 seconds. Make every one count, because your audience may include people suffering from limited attention spans this late in the proceedings.

## Timing

Lightning talks are *short*.

- We will tell you when there are 2 minutes left.
- At 1 minute left we will take over the presentation laptop and start switching to the next set of slides.
- At 0 minutes left the next talk will start, no matter what.

## Slides

- Your slides (as PDF or PowerPoint or ...) should be on the presentation computer before the session starts (talk to Dave Clements)

## INTEGRATING WORKFLOWS WITH PAPERS

Publish in the GigaScience Special Galaxy Series and benefit from:

- 15% off submissions to the Galaxy Series
- Quick publication— average time to first decision less than 4 weeks
- Free deposition and curation of your data in GigaDB database with no size limit
- All data is hosted and can be integrated into the journal's GigaGalaxy server
- A home & citeable DOI for data & workflows

(GIGA)<sup>TM</sup> Galaxy (GIGA)<sup>TM</sup> SCIENCE (GIGA)<sup>TM</sup> DB

OXFORD  
UNIVERSITY PRESS

bit.ly/galaxy\_series

华大基因  
BGI

See <http://bit.ly/gcc2017lightning> for the current list of lightning talks.

# Birds of a Feather Meetups

There is no better place than a Galaxy Community Conference to meet and learn from others doing high-throughput biology. GCC2017 includes *Birds of a Feather* (BoF) meetups that are informal gatherings based on participants' shared interests. BoFs are encouraged throughout GCC2017, particularly during the *flocking* sessions each day. These sessions are time set aside specifically for Birds of a Feather gatherings.

If you are interested in a BoF then *just show up*. If you want to organize a BoF, see <http://bit.ly/gcc2017createbof> for how to get one going. It's never too late to start a BoF, and once one is proposed the organizers will get the word out about it.

The BoFs below were planned when this program was printed. See [http://bit.ly/gcc2017\\_bofs](http://bit.ly/gcc2017_bofs) for the current list.

## Thursday BoFs

### Arts & Crafts

*Thursday & Friday, 29-30 June, Lunch*

Organizers: Saskia Hiltemann, Erasmus MC; Eric Rasche Texas A&M

GCC sure can be overwhelming sometimes! This BoF is a quiet place to do some stress free, science related, arts and crafts. A&C was a big hit at GCC2016.

<http://sched.co/BABe>

### Reproducibility best practices for computational analyses involving large numbers of datasets

*Thursday, 29 June, Lunch*

Organizer: Geir Kjetil Sandve, University of Oslo

The aim for this BoF is a very concrete one: to initiate a collaborative project towards a manuscript on best practices for reproducibility and transparency when doing computational analysis involving large numbers of datasets. Anyone curious or potentially interested in contributing to such a manuscript is encouraged to join!



**STAY**  
**Informed**  
with  
**genome**web

**DID YOU KNOW** THAT  
ACADEMIC AND NON-PROFIT  
RESEARCHERS HAVE FREE  
ACCESS TO **GENOMEWEB'S**  
PREMIUM NEWS CONTENT?

**Just register** with your work  
email address and you'll have  
access to the **entire site!**

**REGISTER HERE**  
[bit.ly/gcc17gw](http://bit.ly/gcc17gw)



Ensuring that omics analyses are truly transparent and reproducible is always a challenge, and even more so when analyses involve large numbers of datasets. The exact scope of a manuscript on this topic and organization of the work is open for discussion at the BoF session. How to make transparent analyses using Galaxy and dataset collections will clearly be central. How much focus to put on comparison with alternative interfaces is open for discussion. As for organization of the work, collaborative editing on an open document might serve as a main pillar, but this would also be a point of discussion at the BoF.

<http://sched.co/BGzq>

## Supporting the non-coder Galaxy community *Thurs, 29 June, Lunch*

*Organizer: Peter Van Heusden, SANBI*

How can we support Galaxy community members whose (main) contribution is not coding? I.e. they might be doing training (although GTN is generally covering these people), being "ambassadors" for Galaxy, doing user support. Some questions:

1. How can these people be recognised and
2. what communication channels can enhance this part of the community and
3. how can this community and the coder / developer community link?

<http://sched.co/BJ0D>

## GalaxyAdmins

*Thursday, 29 June, 18:05 - 18:45*

*Organizer: Dave Clements, Johns Hopkins University*

GalaxyAdmins is a group of people that are responsible for administering Galaxy instances. We meet online and at events like GCC2017, where a lot of us happen to be. GCC2017 will be the 5th in-person GalaxyAdmins meetup. Past meetups have resulted in action items for the coming year. This meetup will discuss plans for the coming year, GalaxyAdmins leadership, and whatever else you want to talk about.

<http://sched.co/B4DP>

## Genome Annotation in Galaxy

*Thursday, 29 June, 18:05 - 18:45*

*Organizers: Anthony Bretaudeau, INRA, BIPAA and GenOuest platforms; Nathan Dunn, Lawrence Berkeley Lab; Björn Grüning, University of Freiburg; Eric Rasche, Texas A&M*

Discuss genome annotation pipelines in Galaxy. Integration of tools, current / future offerings, needs of biologists. (<https://galaxy-genome-annotation.github.io/>).

<http://sched.co/B78V>

# Friday BoFs

## Arts & Crafts

Lunch

Organizers: Saskia Hiltermann, Erasmus MC; Eric Rasche Texas A&M

GCC sure can be overwhelming sometimes! This BoF is a quiet place to do some stress free, science related, arts and crafts. A&C was a big hit at GCC2016.

<http://sched.co/BCVY>

## Come Together: Exploring Proteomic Analysis within Galaxy

Organizers: Tim Griffin, Pratik Jagtap, James Johnson; University of Minnesota

Lunch

Mass spectrometry-based proteomic analysis is evolving along with new and emerging research disciplines. For example, advances in genomics and transcriptomics have offered opportunities in multi-omics exploration of biological systems. In particular, conversion of RNASeq data into protein FASTA format greatly aids the field of proteogenomics. Moreover, functional microbiome research is being greatly helped by newer developments in metaproteomics research. There is a need for improved Galaxy workflows for these multi-omics research areas and emerging methods such as data-independent analysis. The GCC offers a great forum for community to come together and discuss challenges, opportunities and possibilities.

<http://sched.co/BJ5t>

## Containerized Galaxy Deployments and Advanced Testing

Organizers: Johns Chilton, Penn State University; Björn Grüning, University of Freiburg

Lunch

During this buzzword heavy BoF, we hope to gather deployers interested in a broad range of topics from big picture choices (such as DevOps, Containers, and Clouds) to specific technologies (for instance Ansible, Condor, and Kubernetes) and specific Galaxy communities (including docker-galaxy-stable, CloudMan, and GalaxyKickStart). This last year has seen significant sharing and splintering of resources for containerized deployments - we hope this BoF can serve as a venue to coordinate plans for the next year to maximize reuse we can accomplish going forward.

Possible Topics Include:

- Discuss plans for major Galaxy virtualization projects for the next year - including at least docker-galaxy-stable but hopefully more such CloudMan and GalaxyKickStart.
- Discuss different options for deploying Galaxy to Kubernetes / Swarm / OpenShift and potentially coordinate reuse across methods.

- Discuss multi-container deployment options being developed as part of the docker-galaxy-stable project including planned development and testing of complete deployment recipes for Kubernetes, Swarm, and OpenShift.
- Discuss plans for major Galaxy-related Ansible efforts this year and develop plans to reduce duplication across projects.
- Discuss our efforts to automate testing of Galaxy deployment artifacts over the last year and how they can be leveraged. Discuss plans for future testing.

<http://sched.co/BCVj>

## Galaxy for Core Facilities

*Lunch*

*Organizer: Brad Langhorst, New England BioLabs*

Do you work with a sequencing core facility? How are you using galaxy (QC, sample tracking, etc)? What missing things do you need?

<http://sched.co/BJby>

## Trackster support group

*Lunch*

*Organizers: Frederik Coppens, ELIXIR Belgium & VIB; Mo Heydarian, Johns Hopkins University; Michael Sauria, Johns Hopkins University*

We will gather to discuss Trackster, Galaxy's very own genome browser. We'll discuss the benefits of Trackster and start an ongoing discussion on what features could be added to further increase the utility of this valuable feature.

<http://sched.co/BHI3>

# Social Events

We have numerous scheduled social events at GCC2017. However, GCC is one large social event for our community. When you're free, skip checking your email and introduce yourself to the person next to you. You have at least one thing in common - you both have excellent taste in choosing conferences! If you dig deeper, you'll likely find that you have much more in common and maybe shared interests. Many fruitful collaborations have been initiated at GCC. Strike up a conversation!

## Conference Dinner

The official conference dinner is Thursday evening. It will be held at Le Château du Pouget. Buses will make the journey from, and return to Le Corum. Don't miss it.

## Capture the Flag: Hack the Universe

Saskia Hilemann and Eric Rasche have put together a *Capture the Flag (CTF)* event for GCC2017. Hack the Universe is a competition where you try to hack into a Galaxy instance and learn about Galaxy, Galaxy Administration, vulnerabilities, and good security practices. CTF will start on the first night of GCC2017, and run until a week after the conference.

Anyone can participate! We *strongly* recommend you form teams involving at least one bioinformatician or biologist, and one computer person. *Get out there, make new friends!*

Interested? See <https://ctf.galaxians.org/> for more and visit Saskia and Eric at the Arts & Crafts BoF.

## BoFs

BoFs aren't focused on social interactions, but they are an excellent way to meet other participants that are eager to discuss a common interest. The Arts & Crafts BoF, however, is all about relaxing with your fellow participants.



INDIANA UNIVERSITY

Travel Fellowships / Galaxy Community Fund

# Keynotes & Accepted Talks

The full abstracts are available online.

## Session 1, Thursday, 29 June, morning

### **The Evolution of Galaxy: A rough Timeline**

<http://sched.co/B4z8>

Anton Nekrutenko 1, James Taylor 2

1. Penn State University
2. Johns Hopkins University

### **Bioconda and BioContainers - Enabling Universal Reproducibility in Galaxy**

<http://sched.co/B4zG>

John Chilton 1\*, Marius Van Den Beek 2, Björn Grüning 3, Galaxy Team

1. Department of Biochemistry and Molecular Biology, Penn State University, University Park, PA 16801 - United States
  2. Institut Curie, Institut Curie, 26 rue d'Ulm 75248 PARIS CEDEX 05 - France
  3. Uni-Freiburg (ALU), Bioinformatics Group Department of Computer Science, Albert-Ludwigs-University Freiburg Georges-Köhler-Allee 106 79110 Freiburg Germany - Germany
- \* : Corresponding author

### **Building an open, collaborative, online infrastructure for bioinformatics training**

<http://sched.co/B55d>

Bérénice Batut 1\*, Galaxy Training Network 2, Dave Clements 3\*, Björn Grüning 4\*

<http://sched.co/B55d>

1. University of Freiburg, Bioinformatics Group, Department of Computer Science
  2. Galaxy Training Network
  3. Johns Hopkins University
  4. University of Freiburg, Bioinformatics Group, Department of Computer Science
- \* : Corresponding author

## Session 2, Thursday, 29 June, late morning

### **An Educational Framework for Galaxy Administration**

<http://sched.co/B56f>

Martin Cech 1\*, Enis Afghan 1, Nate Coraor 1, Simon Gladman 2, Daniel Blankenberg 1, Björn Grüning 3, Ross Lazarus, Dave Clements 4

1. Department of Biochemistry and Molecular Biology, Penn State University, University Park, PA
  2. University of Melbourne
  3. Department of Computer Science, Albert-Ludwigs-University, Freiburg
  4. Department of Biology, Johns Hopkins University, Baltimore, MD
- \* : Corresponding author

### **Integrative analysis of hundreds of datasets in Galaxy**

<http://sched.co/B56o>

Geir Kjetil Sandve 1, Boris Simovski 1, Sveinung Gundersen, Diana Domanska, Christin Lund-Andersen

1. Department of Informatics [Oslo], Informatikkbygget Gaustadalléen 23 N-0371 Oslo Norway - Norway

## Galaxy Genome Annotation Project: Galaxy and GMOD for Annotation, Teaching, and Genomic Databases

<http://sched.co/B59G>

Eric Rasche 1, Björn Grüning 2, Nathan Dunn 3, Anthony Bretaudeau 4

1. Center for Phage Technology, Texas A&M University (CPT)
2. Department of Computer Science [Freiburg], Georges-Köhler-Allee 79, 79108 Freiburg, Germany - Germany
3. Lawrence Berkeley National Laboratory (LBNL)
4. BIPAA/GenOuest, Institut National de la Recherche Agronomique – INRA

## Apollo in Galaxy: Increasing Opportunities for Collaborative Genome Annotation

<http://sched.co/B57M>

Nathan Dunn 1\*, Monica Munoz-Torres 1, Deepak Unni 2, Eric Rasche 3, Eric Yao 4, Ian Holmes 4, Christine Elisk 2, Suzanna Lewis 1

1. Lawrence Berkeley National Lab (LBNL), Lawrence Berkeley National Lab 717 Potter Street Berkeley, CA 94710 - United States
2. Division of Plant Sciences, University of Missouri, 52 Agriculture Lab, Columbia, MO 65211 - United States
3. Center for Phage Technology, Texas A&M University (CPT), 2128 TAMU College Station, TX 77843-2128 - United States
4. Department of Bioengineering, University of California, Berkeley (UC Berkeley), 119 California Hall Berkeley, CA 94720-1500 - United States

\* : Corresponding author

## Galaxy at the IFB: From Compute to Training

<http://sched.co/B8XT>

Victoria Dominguez del Angel 1, Cristophe Caron 2

1. Institut Français de Bioinformatique
2. INRA

## Session 3, Thursday, 29 June, early afternoon

### Galaxy Community Update

<http://sched.co/B57W>

Jeremy Goecks 1, Dan Blankenberg 2

1. Oregon Health & Science University
2. Penn State University

### The Year of IUC

<http://sched.co/B57a>

Björn Grüning 1\*, Bérénice Batut 1, Marius Van Den Beek 2, Daniel Blankenberg 3, Dave Bouvier 3, Anthony Bretaudeau 4, Nate Coraor 3, John Chilton 3, Peter Cock 5, Saskia Hiltemann 6, Youri Hoogstrate 6, James Johnson 7, Greg Von Kuster 8, Lance Parsons 9, Eric Rasche 10, Nicola Soranzo 11

1. Uni-Freiburg (ALU), Bioinformatics Group Department of Computer Science Albert-Ludwigs-University Freiburg Georges-Köhler-Allee 106 79110 Freiburg Germany
2. Institut Curie, 26 rue d'Ulm 75248 PARIS CEDEX 05 - France
3. Penn State / Galaxy Team
4. Institut de Génétique, Environnement et Protection des Plantes (IGEPP), Institut national de la recherche agronomique (INRA) : UMR1349, AGROCAMPUS OUEST, UMR1349 IGEPP, F-35042 Rennes, - France
5. James Hutton Institute
6. ErasmusMC
7. University of Minnesota

8. Penn State University, University Park, PA
  9. Princeton University
  10. Center for Phage Technology, Texas A&M University (CPT), 2128 TAMU College Station, TX 77843-2128 - United States
  11. Earlham Institute
- \* : Corresponding author

## **Workflow4Metabolomics: Towards an international computing infrastructure and a tools showcase for Metabolomics**

<http://sched.co/B58O>

Gildas Le Corguillé 1, Franck Giacomoni 2, Pierrick Roger-Mele 3, Christophe Duperier 2, Mélanie Pétéra 2, Yann Guitton 4, Marie Tremblay-Franco 5, Jean-François Martin 5, Cécile Canlet 5, Alexis Delabrière 3, Etienne Thévenot 3, Christophe Caron 6

1. ABIMS, CNRS : FR2424, Université Pierre et Marie Curie [UPMC] - Paris VI, Station Biologique de Roscoff - France
2. PFEM, Institut national de la recherche agronomique (INRA) : UMR1019, Saint Genes Champanelle - France
3. Laboratory for Data Analysis and Systems Intelligence (CEA-LIST/LADIS), CEA-LIST, Gif-sur-Yvette - France
4. Laboratoire d'Etude des Résidus et Contaminants dans les Aliments - Oniris (LABERCA), LUNAM Université, Nantes - France
5. Toxalim, Institut national de la recherche agronomique (INRA), Université Paul Sabatier - Toulouse 3, Toulouse - France
6. Ingenium, Institut national de la recherche agronomique (INRA), Toulouse - France

## **South Green Galaxy: a suite of tools for plant genomics**

<http://sched.co/B595>

Jean-François Dufayard 1, Marilyne Summo 2, Gaëtan Droc 2, Mathieu Rouard 1, Bertrand Pitollat 1, Paul Pastor 1, Delphine Larivière 2, Frédéric De Lamotte 3, Stéphanie Bocs 1, Manuel Ruiz 1, Felix Homa 1, Guillaume Martin 1, Gautier SARAH 4, Sébastien RAVEL 5, Vincent Maillol 1, Jonathan Lorenzo 6, Alexis Dereeper 1

1. CIRAD UMR AGAP (AGAP)
2. Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD)
3. INRA UMR AGAP (AGAP)
4. Amélioration Génétique et Adaptation des Plantes Méditerranéennes et Tropicales (AGAP)
5. Biologie et génétique des interactions plantes-parasites pour la protection intégrée (BGPI)
6. Institut Français de Bioinformatique (IFB)

## **Performing Next Generation Phylogenetic Analyses with NGPhylogeny.fr**

<http://sched.co/B598>

Damien Correia 1, 2, 3\*, Vincent Lefort 4, Olivia Doppelt-Azeroual 2, Fabien Mareuil 2, Sarah Cohen-Boulakia 3\*, Olivier Gascuel 2, 4

1. Institut Français de Bioinformatique - UMS CNRS 3601 (IFB-CORE), Centre National de la Recherche Scientifique : UMS3601, IFB-core, Bâtiment 21, 2 av. de la Terrasse, 91190, Gif-sur-Yvette - France
2. Centre de Bioinformatique, Biostatistique et Biologie Intégrative (C3BI), Institut Pasteur [Paris], CNRS : USR3756, 25-28 rue du docteur Roux, 75724 Paris cedex 15 - France
3. Laboratoire de Recherche en Informatique (LRI), Université Paris-Sud - Paris 11, Centre National de la Recherche Scientifique : UMR8623, LRI - Bâtiments 650-660 Université Paris-Sud 91405 Orsay Cedex - France
4. Laboratoire d'Informatique de Robotique et de Microélectronique de Montpellier (LIRMM), Université de Montpellier : UMR5506, Centre National de la Recherche Scientifique : UMR5506, CC 477, 161 rue Ada, 34095 Montpellier Cedex 5 - France

\* : Corresponding author

## Session 4, Thursday, 29 June, late afternoon

### The Institute of Computational Biology: an overview

<http://sched.co/B59E>

Eric Rivals I

I. The Institute of Computational Biology

### Lightning Talks

#### Improving Your Text Life

<http://sched.co/B9jV>

James Johnson I

I. University of Minnesota Supercomputing Institute (MSI)

#### Galaxy-E : Towards an accessible, reproducible and transparent data analysis and management universe dedicated to Ecology

<http://sched.co/B5xf>

Valentin Chambon, Thimothée Virgoulay, Eloïse Trigodet, Marie Delannoy, Marianne Linares, Grégoire Loïs, Romain Julliard, Yvan Le Bras I, \*

I. MNHN CESCO & Staiton de Biologie Marine de Concarneau, Muséum National d'Histoire Naturelle (MNHN)

\* : Corresponding author

#### The Galaxy Gateway on F1000Research

<http://sched.co/B9id>

Holly Murray I

I. Science Navigation Group, Middlesex House, 34-42 Cleveland St - United Kingdom

#### GalaxyCat - The online catalog of Galaxy tools across instances

<http://sched.co/BCAV>

Julien Seiler I\*, Stéphanie Le Gras I

I. Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS :

UMR7104, Inserm : U964, université de Strasbourg Parc D'Innovation I Rue Laurent Fries - BP 10142 67404 ILLKIRCH CEDEX - France

\* : Corresponding author

#### Analyzing Relationships among Clinical Factors of Cancer Patients by Regression Analysis

<http://sched.co/B5yY>

Sugandima Vidanagamachchi I, Tamara Waidyarathna I\*

I. Department of Computer Science (University of Ruhuna)

\* : Corresponding author

## Session 5, Friday, 30 June, morning

### Personalized Analysis of Cancer Data: From Genes to Pathways (and Back)

<http://sched.co/B59d>

Eytan Domany I

I. Weizmann Institute of Science

### Comparing hundreds of RNA-seq libraries using Galaxy and SeqResults

<http://sched.co/B59n>

Brad Langhorst I

I. New England Biolabs (NEB)



## Approaches for small RNA-seq analysis in Galaxy

<http://sched.co/B5GD>

Mallory Freeberg I, James Taylor I

I. Biology Department, Johns Hopkins University (JHU), 3400 N Charles St - United States

## Session 6, Friday, 30 June, late morning

### COMBAT TB: an integrated environment for Tuberculosis data analysis and surveillance

<http://sched.co/B5GG>

Peter Van Heusden I\*, Thoba Lose I, Ziphozakhe Mashologu I, Alan Christoffels I\*

I. South African National Bioinformatics Institute (SANBI)

\*: Corresponding author

### The RNA Workbench - reproducible, transparent, and accessible RNA research

<http://sched.co/B5Ld>

Florian Eggenhofer I, Jörg Fallmann 2, Björn Grüning I, 3, Bérénice Batut I, Peter Stadler 2, 4, 5, Rolf Backofen I, 3, 6

1. University of Freiburg (ALU), Bioinformatics Group Department of Computer Science Albert-Ludwigs-University Freiburg Georges-Köhler-Allee 106 79110 Freiburg Germany - Germany
2. University of Leipzig, Bioinformatics Group, Department of Computer Science, and Interdisciplinary Center for Bioinformatics, University of Leipzig, Härtelstraße 16-18, 04107 Leipzig, Germany - Germany
3. Center for Biological Systems Analysis (ZBSA), Center for Biological Systems Analysis (ZBSA), University of Freiburg, Habsburgerstr. 49, 79104 Freiburg, Germany - Germany
4. Institute for Theoretical Chemistry (TBI), Institute for Theoretical Chemistry, University of Vienna, Währingerstrasse 17, 1090 Vienna, Austria - Austria
5. Max Planck Institute for Mathematics in the Sciences (MPI), Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany - Germany
6. Centre for Biological Signaling Studies [Freiburg] (BIOSS), Schänzlestr. 1879104 Freiburg - Germany

### Interaction Hub: a comprehensive Galaxy-based chromatin folding database, browser, and analysis platform

<http://sched.co/B5Le>

Michael Sauria I, James Taylor I

I. Department of Biology [Johns Hopkins University], 3400 N. Charles Street, Mudd 144 Baltimore, MD 21218 - United States

### Mechanizing our biologists with Galaxy and IRIDA platforms

<http://sched.co/B5Lf>

Philip Mabon I

I. Public Health Agency of Canada (PHAC), 1015 Arlington Street, Winnipeg, Manitoba R3E3R2 - Canada

## Session 7, Friday, 30 June, early afternoon

### ELIXIR supports Galaxy users and developers

<http://sched.co/B5Lh>

Frederik Coppens I, 2

1. ELIXIR-Belgium
2. VIB

## Advancing metaproteomics research via community-based informatics development and dissemination

<http://sched.co/B5Li>

Pratik Jagtap 1\*, Björn Grüning, James Johnson, Alessandro Tanca, Bart Mesuere, Judson Herve, Carolin Kolmeder, Jeremy Fischer, Thomas Doak, Thilo Muth, Dave Clements 2, Praveen Kumar, Subina Mehta, Thomas Mcgoan, Clemens Blank, Bernhard Renard, Josh Elias, Joel Rudney, Timothy Griffin

1. University of Minnesota (Galaxy-P), 7-166 MCB, 420 Washington Ave SE, Minneapolis, MN 55455 - United States
  2. Johns Hopkins University
- \* : Corresponding author

## Making Galaxy User Interface Pluggable with Webhooks

<http://sched.co/B5Lk>

Evgeny Anatskiy 1, Anup Kumar 1, Joachim Wolff 1, Clemens Blank 1, Eric Rasche 2, Martin Cech 3, Björn Grüning 1

1. Bioinformatics Group at Albert-Ludwigs-Universität, Freiburg
2. Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX
3. Department of Biochemistry and Molecular Biology, Penn State University, University Park, PA

## A reproducible data analysis environment for next-generation sequencing on public cloud computer

<http://sched.co/B5Lo>

Manabu Ishii 1, Matsushima Akihiro 1, Mika Yoshimura 1, Hiroki Danno 1, Itoshi Nikaido 1

4. RIKEN ACCC Bioinformatics Research Unit, 2-1 Hirosawa, Wako, Saitama 351-0198 - Japan

## Developing analytical workflows for plant genomics through the integration of Galaxy and Tripal

<http://sched.co/B5x9>

Sean Buehler 1, Nic Herndon 1, Emily Grau 1, Ming Chen 2, Abdullah Almsaeed 2, Connor Wytko 3, Brian Soto 3, Sook Jung 3, Shawna Spoor 3, Kuangching Wang 4, Chun-Huai Cheng 3, Nick Watts 4, Lacey Sanderson 5, Jill Wegrzyn 6, Doreen Main 3, Alex Feltus 4, Margaret Staton 7, Stephen Ficklin 3, Nathan Henry 7

1. University of Connecticut
2. University of Tennessee Institute of Agriculture
3. Washington State University
4. Clemson University
5. University of Saskatchewan
6. University of Connecticut
7. University of Tennessee

## Session 8, Friday, 30 June, late afternoon

### Lightning Talks

#### Using visual programming in Cervical cancer data of The Cancer Genome Atlas

<http://sched.co/B5yV>

Thais Hosokawa 1, José Fregnani 1, Rui Reis 1, Adriane Evangelista 1

1. [Barretos Cancer Hospital \(BCH\)](#), Barretos, São Paulo, Brazil Galaxy at scale: Analyzing thousands of single cell transcriptomes

## **Galaxy at scale: Analyzing thousands of single cell transcriptomes**

<http://sched.co/BH73>

Mo Heydariyan I, Enis Afgan I, James Taylor I\*

I. Johns Hopkins University

\*: Corresponding author

## **ToolDog - generating tool descriptors from the ELIXIR tool registry**

<http://sched.co/B5y9>

Kenzo-Hugo Hillion I\*, Ivan Kuzmin 2, Hedi Peterson 2, Jon Ison 3, Hervé Ménager I\*

1. Center of Bioinformatics, Biostatistics and Integrative Biology (C3BI), Institut Pasteur de Paris

2. Center of Bioinformatics, Biostatistics and Integrative Biology

3. Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark

\*: Corresponding author

## **Pushing the limits of job flexibility on HPC**

<http://sched.co/B5z0>

Carrie Ganote I

I. Indiana University (IU), 107 S. Indiana Avenue Bloomington, IN 47405-7000 - United States

## **Pulling the Galaxy's Strings**

<http://sched.co/B5z4>

Jeffrey Miller I\*, Brenna Miller I

I. The University of Iowa

\*: Corresponding author

## **Using Galaxy as a platform for continuous software development**

<http://sched.co/B5z9>

Boris Simovski I, Geir Kjetil Sandve I\*, Sveinung Gundersen I

I. Department of Informatics, University of Oslo [Oslo] (UiO), Informatikkbygget Gaustadalléen 23 N-0371 Oslo Norway - Norway

\*: Corresponding author

## **Automated Generation of Complex Toolkits for Galaxy**

<http://sched.co/B5y5>

Daniel Blankenberg I

I. Penn State / Galaxy Team

## **Hack the Galaxy: Data Report**

<http://sched.co/BFR9>

Hack the Galaxy: Data Organizers

## **Hack the Galaxy: Dev Report**

<http://sched.co/BFRA>

Hack the Galaxy: Dev Organizers

## **JBrowse and Circos Tool Updates**

<http://sched.co/BH74>

Eric Rasche I, Saskia Hiltmann 2,

1. Center for Phage Technology, Texas A&M University (CPT)

2. Bioinformatics Dept., Erasmus University Medical Center (Erasmus MC)

## **Parsec: Intergalactic DevOps and Analysis Automation**

<http://sched.co/BH78>

Eric Rasche I

I. Center for Phage Technology, Texas A&M University (CPT)

# Posters

Odd numbered posters will be presented Thursday at 15:40, and even-numbered posters on Friday at 15:20.

Poster titles and authors are listed here. The full abstracts of posters can be found online at <https://gcc2017.sched.com/overview/type/Posters>.

## **P01: Galaxy-docker-ansible**

<http://sched.co/B8RO>

Ruben Vorderman I

I. Leiden University Medical Center (LUMC)

## **P02: Pulling Galaxy's Strings**

<http://sched.co/B8WC>

Jeffrey Miller I

I. The University of Iowa (UI), 230 S. Madison St. Iowa City, IA 52242 - United States

## **P03: Automated customized deployment Galaxy portal instance in IFB Cloud Computing**

<http://sched.co/B8RR>

Sandrine Perrin I\*, Björn Grüning 2, Bryan Brancotte I, Christophe Blanchet I\*

1. Institut Français de Bioinformatique (IFB), Institut Français de Bioinformatique/IFB
  2. Biological Systems Analysis (ZBSA), Department of Computer Science, Albert-Ludwigs-University, Center for Biological Systems, University of Freiburg, Freiburg, - Germany
- \* : Corresponding author

## **P04: Integration of openBIS - improving metadata management in Galaxy**

<http://sched.co/B8XV>

Aarif Mohamed Nazeer Batcha I\*, Guokun Zhang I\*, Ulrich Mansmann I

1. Department of Medical Informatics, Biometry and Epidemiology, University of Munich (IBE, LMU), Marchioninistrasse 15, 81377 Munich - Germany
- \* : Corresponding author

## **P05: The EDAM ontology and its integration into Galaxy**

<http://sched.co/B8RV>

Hervé Ménager I, Jon Ison 2, Matúš Kalaš 3, Veit Schwaemmle 4, Edam Contributors

1. Bioinformatics and Biostatistics Hub of the C3BI, Institut Pasteur, Institut Pasteur de Paris, 25 rue du Dr Roux, 75015 Paris - France
2. Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark
3. Computational Biology Unit (CBU), University of Bergen
4. Protein Research Group, Department for Biochemistry and Molecular Biology, University of Southern Denmark

## **P06: A Public Galaxy platform @Pasteur used as an execution engine for web services**

<http://sched.co/B8Xm>

Fabien Mareuil I, Olivia Doppelt-Azeroual I, Hervé Ménager I

1. Hub Bioinformatique et Biostatistique, Institut Pasteur [Paris], C3BI 25-28, rue du Docteur Roux, 75724 Paris cedex 15 - France

### **P07: myFAIR Analysis: Personal FAIR Data Management and Analysis**

<http://sched.co/B8Rc>

Saskia Hiltemann 1, Rick Jansen 1, David Van Zessen 1, John Hays 2, Andrew Stubbs 1\*

1. Bioinformatics Dept., Erasmus University Medical Center (Erasmus MC)
2. MMIZ, Erasmus University Medical Center (Erasmus MC)

\* : Corresponding author

### **P08: GEGA: An application for secure access and analysis of EGA data with Galaxy**

<http://sched.co/B9ip>

David Van Zessen 1, Alexander Sned 2, Youri Hoogstrate 1, Dylan Spalding 2, Andrew Stubbs 1\*

1. Bioinformatics Dept., Erasmus University Medical Center (Erasmus MC)
2. European Bioinformatics Institute [Hinxton], EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK - United Kingdom

\* : Corresponding author

### **P09: Wrapping tools: on the shoulders of the GCC community**

<http://sched.co/B8TT>

Valentin Marcon 1, Nathalie Choise 2, Veronique Jamilloux 2, Gwendoline Andres, Isabelle Luyten 2, Joelle Amselem 2, Mikael Loaec 2, Françoise Alfama-Depauw 2, Sarah Maman-Haddad 3, Melanie Petera 4, Luc Jouneau 5, Sandrine Laguerre 6, Olivier Inizan 1

1. MaLAGE (MaLAGE), Institut National de la Recherche Agronomique - INRA (FRANCE), Centre de Jouy-en-Josas Domaine de Vilvert F78352 JOUY-EN-JOSAS Cedex - France
2. Unité de Recherche Génomique Info (URGI), Institut National de la Recherche Agronomique : UR1164, INRA, Centre de recherche de Versailles, bat.18 Route de Saint Cyr 78000 Versailles - France
3. Génétique Physiologie et Systèmes d'Elevage (GenPhySE), Institut national de la recherche agronomique [Toulouse], Chemin de Borde Rouge CS 52627 31326 Castanet Tolosan - France
4. Unité de Nutrition Humaine (UNH), Institut national de la recherche agronomique (INRA)
5. Unité de recherche Virologie et Immunologie Moléculaires (VIM), Institut National de la Recherche Agronomique : UR0892, UR0892 INRA Centre de Recherche de Jouy-en-Josas F-78352 Jouy-en-Josas cedex - France
6. l'Ingénierie des Systèmes Biologiques et des Procédés (LISBP), Institut National des Sciences Appliquées de Toulouse, 135 Avenue de Rangueil, 31400 Toulouse - France

### **P10: ToolDog - generating tool descriptors from the ELIXIR tool registry**

<http://sched.co/B8Xp>

Kenzo-Hugo Hillion 1\*, Ivan Kuzmin 2, Hedi Peterson 2, Jon Ison 3, Hervé Ménager 1\*

1. Center of Bioinformatics, Biostatistics and Integrative Biology (C3BI), Institut Pasteur de Paris
2. Institute of Computer Science, University of Tartu
3. Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark

\* : Corresponding author

### **P11: Utility scripts for data manipulation in Galaxy**

<http://sched.co/B8Tc>

Simon Gladman 1\*, Madison Flannery 1

1. Melbourne Bioinformatics, The University of Melbourne (MelBioinfo), 187 Grattan St, Carlton, Victoria, 3053 - Australia

\* : Corresponding author

## **P12: Integration of Alvis within the OpenMinTeD platform as Galaxy utilities**

<http://sched.co/B8Xs>

Jean-Baptiste Bohuon \*, Mouhamadou Ba, Robert Bossy I, Claire Nédellec I

1. Mathématique et Informatique Appliquées du Génome à l'Environnement (MaIAGE), Institut national de la recherche agronomique (INRA), MaIAGE, INRA, F78350 Jouy-en-Josas - France - France

\* : Corresponding author

## **P13: Prioritizing SNPs using the Neo4j Galaxy Interactive Environment**

<http://sched.co/B8Ti>

Lose Thoba I, Ziphozakhe Mashologu I\*, Peter Van Heusden I\*, Alan Christoffels I\*

1. South African National Bioinformatics Institute (SANBI)

\* : Corresponding author

## **P14: Du Novo: a simple, reference-free tool for turning duplex sequencing data into extremely accurate reads**

<http://sched.co/B8Y4>

Nicholas Stoler I\*, Barbara Arbeithuber 2, Wilfried Guiblet I, Marcia Su 3, Kateryna Makova 2, Anton Nekrutenko I\*

1. Department of Biochemistry and Molecular Biology, Penn State University (BMB), University Park, PA 16802 - United States
2. Department of Biology, Penn State University, University Park, PA 16802 - United States
3. The Jackson Laboratory For Genomic Medicine (JAX), Farmington, CT 06032 - United States

\* : Corresponding author

## **P15: The MiModD suite of tools for genetic variant mapping and identification**

<http://sched.co/B8Tl>

Wolfgang Maier I\*, Mark Seifert I, Katharina Moos, Gregory Minevich, Oliver Hobert 2, Ralf Baumeister I

1. University of Freiburg
2. Department of Biological Sciences, Columbia University

\* : Corresponding author

## **P16: Galaxy-P rides the Jetstream: Cloud-based multi-omic informatics**

<http://sched.co/B8YY>

Timothy Griffin I\*, Matthew Chambers 2, James Johnson I, McGowan Thomas I, Thomas Doak 3, Jeremy Fischer 3, Praveen Kumar I, Pratik Jagtap I

1. University of Minnesota
2. Vanderbilt University
3. Indiana University

\* : Corresponding author

## **P17: ProteoRE, a Galaxy-based infrastructure for interpreting and exploring mass spectrometry-based proteomics data**

<http://sched.co/B8U9>

Lien Nguyen I, Maud Lacombe I, Sandra Dérozier 2, Lisa Perus I, Olivier Rué 2, Florence Combes I, Christophe Caron 2, Virginie Brun I, Valentin Loux 2, Yves Vandenbrouck I

1. Institut de Biosciences et Biotechnologies de Grenoble (BIG), Université de Grenoble, Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA) - Grenoble, Centre National de la Recherche Scientifique : FR3425, 17, rue des Martyrs 38054 Grenoble cedex 9 - France
2. Unité Mathématiques et Informatique Appliquées du Génome à l'Environnement (MaIAGE), Institut National de la Recherche Agronomique, Université Paris-Saclay, Domaine de Vilvert 78350 Jouy en Josas Cedex - France

**P18: Implementation of an interface for exploring molecular diversity by proteogenomics**

<http://sched.co/B8Yy>

Yannick Cogne 1, Christine Almunia 1, Olivier Pible 1, Duarte Gouveia 2, Arnaud Chaumot 2, Olivier Geffard 2, Jean Armengaud 1\*

1. Innovative technologies for Detection and Diagnostics (Li2D), CEA Marcoule, BP17171, F-30207 BAGNOLS-SUR-CEZE - France
2. Milieux aquatiques, écologie et pollutions (UR MALY), CEMAGREF, 5 rue de la Doua, CS70077, 69626 Villeurbanne Cedex, France - France

\* : Corresponding author

**P19: A reproducible data analysis environment for next-generation sequencing on public cloud computer**

<http://sched.co/B8UA>

Manabu Ishii 1, Akihiro Matsushima 1, Mika Yoshimura 1, Hiroki Danno 1, Itoshi Nikaido 1

1. RIKEN ACCC Bioinformatics Research Unit, 2-1 Hirosawa, Wako, Saitama 351-0198 - Japan

**P20: Development and implementation of a transversal NGS & bioinformatics platform at the Belgian Institute of Public Health: Deployment of a Galaxy instance for routine use in a public health setting**

<http://sched.co/B8Z6>

Julien Van Braekel 1, Bert Bogaerts 1, Raf Winand 1, Quiang Fu 1, Sigrid De Keersmaecker 1, Nancy Roosens 1, Kevin Vanneste 1\*

1. Belgian Institute of Public Health (WIV-ISP), Rue Juliette Wytsman 14 1050 Ixelles – Belgium

\* : Corresponding author

**P21: Deployment of genome databases for insects using Galaxy Genome Annotation**

<http://sched.co/B8VM>

Anthony Breteau 1\*

1. Institut de Génétique, Environnement et Protection des Plantes (IGEPP), Institut National de la Recherche Agronomique : UMR1349, Université de Rennes 1 : UMR1349, Agrocampus Ouest : UMR1349, Domaine de la Motte au Vicomte BP 3532735653 Le Rheu - France

\* : Corresponding author

**P22: Enhancing the Multi-omics Visualization Platform (MVP) Plug-in for Galaxy-based Applications**

<http://sched.co/B5xC>

Thomas McGowan 1\*, James Johnson 1, Pratik Jagtap 2\*

1. University of Minnesota Supercomputing Institute (MSI), 599 Walter Library, 117 Pleasant St. SE Minneapolis, MN 55455 - United States
2. Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota (BMBB), - 155 Jackson Hall, 321 Church Street SE, Minneapolis, MN 55455 - United States

\* : Corresponding author

**P23: Detecting somatic de-novo Transposable Element insertions with Galaxy**

<http://sched.co/B8VS>

Marius Van Den Beek 1, Nick Riddiford, Katarzyna Siudeja, Allison Bardin

1. Institut Curie, 26 rue d'Ulm 75248 PARIS CEDEX 05 - France

**P24: Annotating microRNA isoforms (isomiRs) of non-model organisms to analyze expression levels using a Galaxy workflow**

<http://sched.co/B8Z7>

Jochen Bick 1, Susanne Ulbrich 1, Stefan Bauersachs 1

1. ETH Zurich Animal Physiology

**P25: Galaxy and the full spectrum of needs in a small-scale cancer study**

<http://sched.co/B8VU>

Christin Lund-Andersen 1, Ståle Nygård 2, 3, Stein Larsen 4, Eivind Hovig 1, 3, Kjersti Flatmark 1, 3, 4

1. Department of Tumor Biology, Institute for Cancer Research, The Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway
2. Bioinformatics Core Facility, Institute for Medical Informatics, Oslo University Hospital.
3. University of Oslo, Oslo, Norway
4. Department of Gastroenterological Surgery, The Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway

**P26: The Burden of Occupational Cancer in Korea**

<http://sched.co/B8Z8>

Mia Son 1

1. Kangwon National University (KNU)

**P27: ASaiM: a Galaxy-based framework to analyze raw sequencing data from microbiota**

<http://sched.co/B8Vj>

Bérénice Batut 1\*, Clémence Defois 2, Kévin Gravouil 2, 3, 4, Jean-François Brugère 2, Eric Peyretailade 2, Pierre Peyret 2

1. Bioinformatics Group, Department of Computer Science, University of Freiburg, Germany
2. Microbiologie Environnement Digestif Santé - Clermont Auvergne (MEDIS), INRA Clermont-Ferrand-Theix : UMR0454, Université Clermont Auvergne : UMR0454, INRA, Centre de Recherches de Clermont-Ferrand - Theix / 63122 Saint-Genès Champanelle - France
3. Microorganismes : Génome et Environnement - Clermont Auvergne (LMGE), Université Clermont Auvergne : UMR6023, Centre National de la Recherche Scientifique : UMR6023, Campus des Cézeaux / 24, avenue des Landais BP 80026 / 63170 Aubière - France
4. Laboratoire d'Informatique, de Modélisation et d'Optimisation des Systèmes - Clermont Auvergne (LIMOS), SIGMA Clermont, Université Clermont Auvergne : UMR6158, Centre National de la Recherche Scientifique : UMR6158, ISIMA / Campus des Cézeaux BP 10025 / 63173 AUBIERE cedex - France

\*: Corresponding author

**P28: ChemFlow, chemometrics for everybody**

<http://sched.co/B8Z9>

Virginie Rossard 1, Eric Latrille 1, Fabien Gogé 2, Jean-Michel Roger 3, Martin Ecarnot 1, Jean-Claude Boulet 1

1. INRA
2. IRSTEA

**P29: Developing a bioinformatics pipeline for optimization of sperm epigenome analysis in mice and men to be used for the identification of epigenetic signatures in sperm associated with environmental perturbation**

<http://sched.co/B8Vl>

Deepak Tanwar 1, Romain Lambrot 1, Keith Siklenka 2, Mahmoud Aarabi 3, Donovan Chan 3, Clifford Librach 4, Sergey Moskvovtsev 4, Jacquetta Trasler 3, Sarah Kimmins 1, 2

1. Department of Animal Science, McGill University, Ste Anne-de-Bellevue, QC - Canada
2. Department of Pharmacology and Therapeutics, McGill University, Montreal, QC - Canada
3. Research Institute of the McGill University Health Centre at the Montreal Children's Hospital, Montreal, QC - Canada
4. Department of Obstetrics and Gynaecology, University of Toronto, Toronto, ON - Canada



**P30: Pushing the limits of job flexibility on HPC**

<http://sched.co/B9jp>

Carrie Ganote I, Sheri Sanders, Phil Blood, Thomas Doak, Bhavya Nalagampalli Papudeshi, Asma Bankapur \*, Brian Haas, Tim Tickle, Cicada Brokaw

I. Indiana University (IU), 107 S. Indiana Avenue Bloomington, IN 47405-7000 - United States

\* : Corresponding author

**P31: In silico toxicity assessment in cultured rat intestinal cells deduced from cellular impedance measurements and transcriptomics data**

<http://sched.co/BIAZ>

Pooja Gupta I

I. Institute for Mathematics, Freie Universität Berlin (FU Berlin)

**P32: An odyssey into Galaxy: Quality control and quality assurance of NGS-based laboratory diagnostic tests offered by Newborn Screening Ontario**

<http://sched.co/BKkj>

Lemuel J. Racacho I,\*

I. [Newborn Screening Ontario \(NSO\)](#)

\* : Corresponding author

# Accepted Demos

Odd numbered demos will be presented Thursday at 15:40, and even-numbered demos on Friday at 15:20. Demos with corresponding posters will be presented concurrently with and adjacent to the poster. Full demo abstracts are online.

## D01: Integration of Linked Data into Galaxy using AskOmics

<http://sched.co/B8ZC>

Xavier Garnier 1, 2, Olivier Dameron 2, Olivier Filangi 1, 3, Fabrice Legeai 1, 4, Anthony Bretaudeau 1, 3

1. Institut de Génétique, Environnement et Protection des Plantes (IGEPP), Institut National de la Recherche Agronomique : UMR1349, Université de Rennes 1 : UMR1349, Agrocampus Ouest : UMR1349, Domaine de la Motte au Vicomte BP 3532735653 Le Rheu - France
2. DYLISS (INRIA - IRISA), INRIA, Université de Rennes 1, CNRS : UMR6074, Campus de Beaulieu 35042 Rennes cedex - France
3. Plateforme bio-informatique Genouest (Genouest), IRISA/INRIA Plateforme GenOuest Campus de Beaulieu 35042 Rennes cedex, France - France
4. GenScale, IRISA/INRIA équipe Genscale Campus de Beaulieu 35042 Rennes cedex, France - France

## D02: Improve Genome Annotation Accuracy through Manual Annotation in Apollo

<http://sched.co/B8aC>

Nathan Dunn 1\*, Monica Munoz-Torres 1, Deepak Unni 2, Eric Rasche 3, Eric Yao 4, Ian Holmes 4, Christine Elsie 2, Suzanna Lewis 1

1. Lawrence Berkeley National Lab (LBNL), Lawrence Berkeley National Lab 717 Potter Street Berkeley, CA 94710 - United States
  2. Division of Plant Sciences, University of Missouri, 52 Agriculture Lab, Columbia, MO 65211 Columbia, MO 65211 - United States
  3. Center for Phage Technology, Texas A&M University (CPT), 2128 TAMU College Station, TX USA 77843-2128 - United States
  4. Department of Bioengineering, University of California, Berkeley (UC Berkeley), 119 California Hall Berkeley, CA 94720-1500 - United States
- \* : Corresponding author

## D03: De novo transcriptome annotation visualization and filtering

<http://sched.co/B8ZY>

Eduardo De Paiva Alves 1

1. Centre for Genome-Enabled Biology and Medicine, University of Aberdeen (CGEBM), 23 St Machar Dr, Aberdeen, UK AB24 3UU - United Kingdom

## D04: KnetMiner: an application suite to integrate, search, and interactively explore large knowledge networks

<http://sched.co/BJ5F>

Ajit Singh 1, Monika Mistry 1, Marco Brandizi 1, Chris Rawlings 1, Keywan Hassani-Pak 1

1. Rothamsted Research, Harpenden, AL5 2JQ - United Kingdom

## Organising Committee

Christophe Caron  
INRA

Dave Clements  
Johns Hopkins University

Gildas Le Corguillé  
Station Biologique Roscoff

Jean-François Dufayard  
CIRAD

Olivier Inizan  
INRA

Patricia Laplagne  
IFB

Marilyne Summo  
CIRAD

Nicole Vasilevsky  
OHSU

## Hack the Galaxy: User

Frederic Coppens  
VIB

Mallory Freeberg  
Johns Hopkins University

Jennifer Hillman-Jackson  
Penn State University

Yvan Le Bras  
MNHM

Stephanie Le Gras  
INSERM

Christian Schudoma  
Earlham Institute

## Scientific Committee

Christophe Antoniewski, Chair  
UPMC

Sarah Cohen-Boulakia  
IRI

Alexis Dereeper  
IRD

Thérèse Combes  
Université de Montpellier

Anton Nekrutenko  
Penn State University

Björn Grüning  
University of Freiburg

Jeremy Goecks  
OHSU

Jean-François Gibrat  
IFB

Vincent Lefort  
LIRMM

Nicola Soranzo  
Earlham Institute

James Taylor  
Johns Hopkins University

## Hack the Galaxy: Dev

Anthony Bretaudeau  
INRA

Björn Grüning  
University of Freiburg

Marius van den Beek  
Institut de Biologie Paris Seine

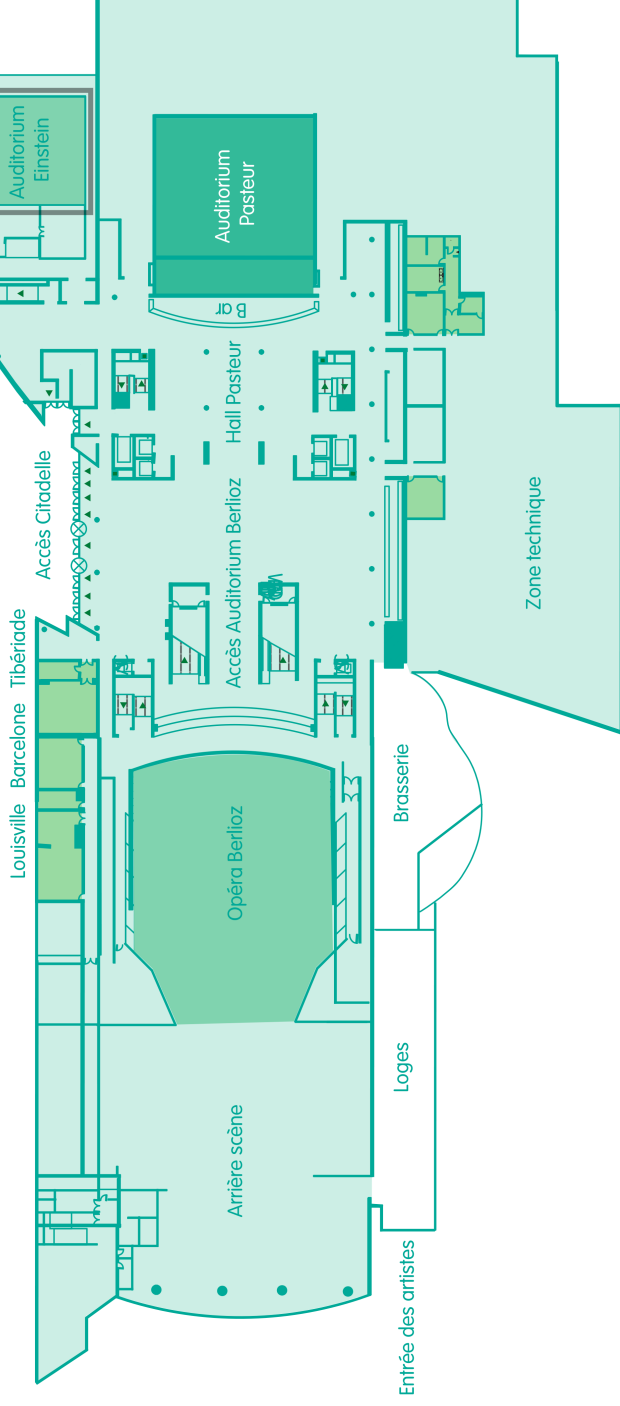
Julien Seiler  
IGBMC



LE CORUM  
BY MONTPELLIER EVENTS

## Einstein Auditorium Meeting Venue

### Level 0

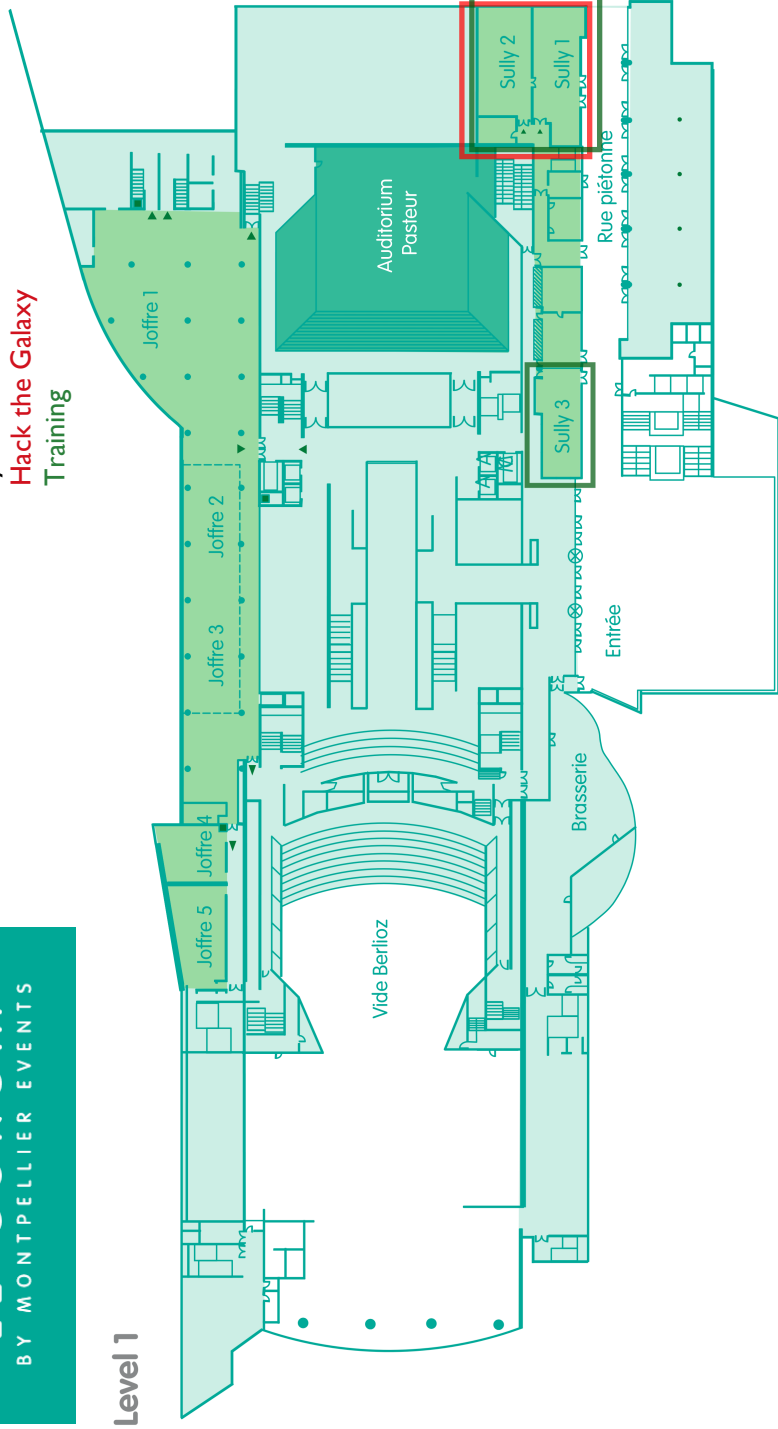




LE CORUM  
BY MONTPELLIÉR EVENTS

Sully 1, 2, and 3  
Hack the Galaxy  
Training

Level 1

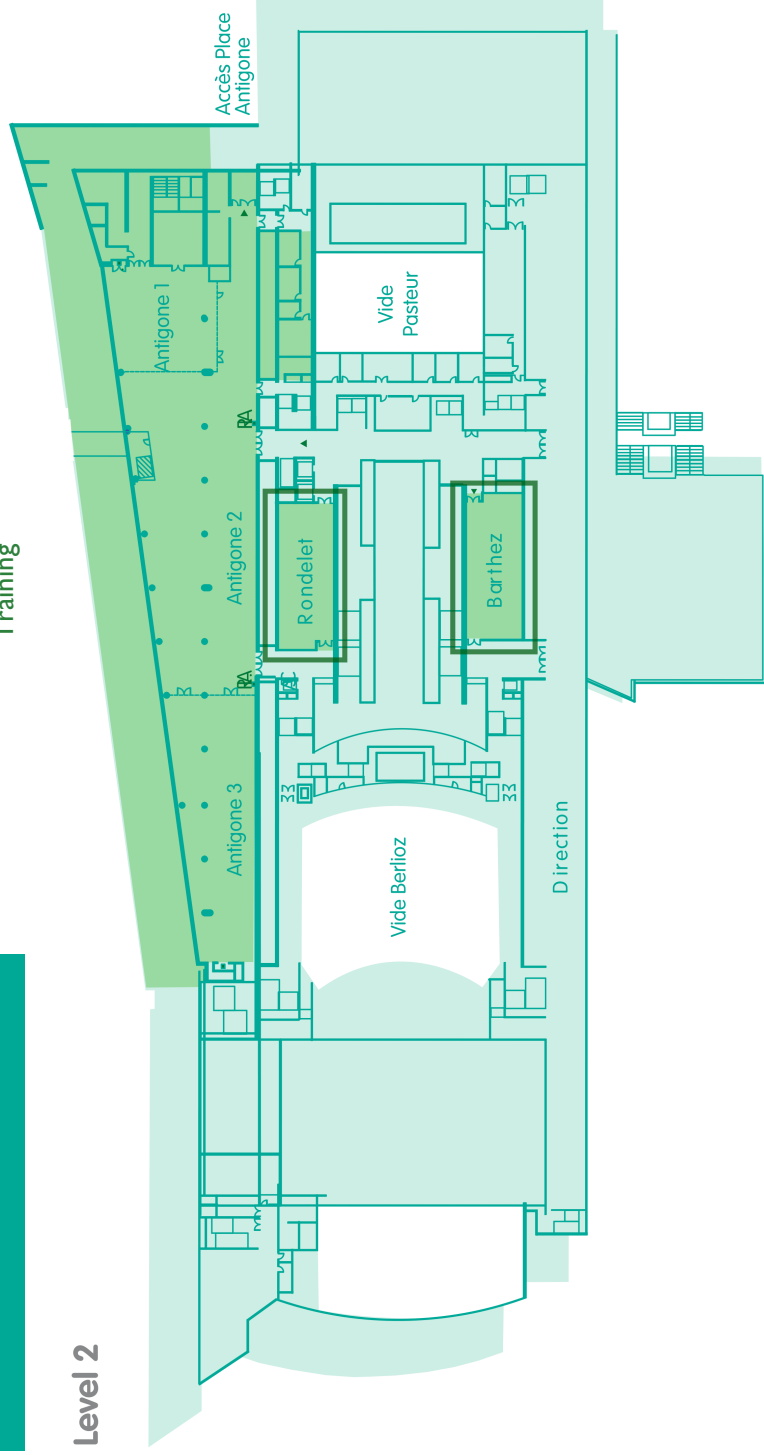




LE CORUM  
BY MONTPELLIER EVENTS

## Barthez & Rondelet Training

### Level 2





# Portland 2018

---

**GCC2018 & BOSC 2018**  
are joining up

---

June 25-30, Reed College, Portland, Oregon  
[gccbosc2018.sched.com](http://gccbosc2018.sched.com)

---

See you next year!