CABANA Workshop: Chemoinformatics in Drug Discovery

Venue: National Laboratory in Genomics for Biodiversity (LANGEBIO), CINVESTAV UGA, Irapuato Gto., Mexico

Date: 15-18 October 2019

Scientific and event organisation: Fabien Plisson [LANGEBIO, CINVESTAV]
Jose Medina-Franco [UNAM]

Speakers:

Rodrigo Ochoa [UdeA, Medellin]

Fabien Plisson [LANGEBIO, CINVESTAV]

Jose Medina-Franco [UNAM]

Barbara Diaz, Student, Medina-Franco Group [UNAM]

Edgar Lopez, Student, Medina-Franco Group [UNAM]

Registration opens: July 22nd, 2019

Registration deadline: August 18th, 2019

Acceptance notification date: September 2nd, 2019

Participation: Open application with selection

Sponsors

CABANA GCRF

LANGEBIO, CINVESTAV

Overview

The aim of this workshop is to introduce researchers to the field of Chemoinformatics, especially in the areas of Structure-Based Drug Design (SBDD) and Ligand-Based Drug Design (LBDD). Trainees will cover the use of protein, ligand and drug databases, protein modelling, molecular docking and virtual screening for drug discovery applications. Managing and analysing virtual chemical libraries will also be covered, including the application of machine learning algorithms in drug discovery.

The workshop will be instructed by trainers from the University of Antioquia in Colombia, from CINVESTAV Unit for Advanced Genomics (LANGEBIO) in Mexico, as well as local expert researchers from the School of Chemistry at the National Autonomous University of Mexico (UNAM).

Audience

This workshop is suitable for researchers working in Latin American and Carribean countries. Early career researchers, including graduate students, postdocs and professors working, or planning to start working in Chemoinformatics, are warmly encouraged to apply.

This workshop will be located in a room with computers however we recommend participants to bring their own laptops. In case you don't own one, please let us know ahead of time. If you have any questions regarding your laptop, please contact us ahead of time. All course materials include materials and lectures.

The workshop will provide lunches and refreshment breaks, please let us know in advance about dietary requirements.

Prerequisites:

Undergraduate-level knowledge of chemical and protein structures are required. Familiarity with basic chemoinformatics concepts and notions of programming (R, Python and/or UNIX) are recommended. **Please note parts of this course will be taught in English, however, the trainers are fluent in Spanish and can offer language support where needed.**

A number of travel fellowships are available for this course - early-stage researchers and researchers from underrepresented groups are especially encouraged to apply for CABANA travel fellowships. You can apply for travel fellowships via the course application form.

Syllabus, Tools and Resources

During this course you will learn about:

- Learn about Protein Data Bank in Europe (PDBe) resources
- Learn about compound databases i.e. BIOFACQUIM, ChEMBL, ZINC
- Use Molecular visualization tools such as DataWarrior, PyMol and Chimera
- Apply homology modelling, molecular docking and virtual screening
- Apply machine learning algorithms and statistics for drug discovery
- Compute with Python and Jupyter Notebook

Learning Outcomes

After this course, you should be able to:

- Visualize proteins and ligands using molecular visualization tools
- Browse and investigate biomolecular structural information using PDBe
- Use molecular docking and virtual screening for hit identification
- Apply homology modelling to the prediction of protein structures
- Apply predictive models (regressions, classifications) to molecules

Programme

Time	Subject	Trainer
Day 1 – Tuesday 15th October		
08:30-09:00	Arrival and registration	
09:00-10:00	Welcome and Introductions	Organizers
10:00-11:00	Public Seminar at Langebio: “SBDD and LBDD approaches in drug discovery to neglected tropical diseases”	Rodrigo Ochoa
11:00-11:30	Coffee Break	
11:30-12:30	Introduction to Chemoinformatics	José Medina Franco
12:30-13:30	Lunch	
13:30-15:00	Introduction to PDBe (proteins) and UCSF Chimera	Rodrigo Ochoa
15:00-15:30	Coffee break	
15:30-16:30	Introduction to ZINC database	Rodrigo Ochoa
16.30-17:30	Introduction to ChEMBL database	Rodrigo Ochoa

Day 2 – Wednesday 16th October		
09:00-09:30	Intro to Structure-Based Drug Discovery	José Luis Medina Franco
9:30-11:00	Practices in Molecular Docking	Edgar López López
11:00-11:30	Coffee break	
11:30-13:00	Practices in Virtual Screening	Edgar López López
13:00-14:00	Lunch	
14:00-15:00	Homology Modelling	Rodrigo Ochoa
15:00-15:30	Coffee break	
15:30-17:00	Practical session with UCSF Chimera and Modeller	Rodrigo Ochoa
Day 3 – Thursday 17th October		
09:00-09:30	Intro to Ligand-Based Drug Discovery	Fabien Plisson
10:00-11:00	Physicochemical descriptors with DataWarrior and RDkit	Barbara Diaz Eufracio Edgar López López
11:00-11:30	Coffee break	
11:30-13:00	Unsupervised Learning – Clustering Practices in BIOFAQUIM vs the world	Fabien Plisson Barbara Diaz Eufracio
13:00-14:00	Lunch	
14:00-15:30	Unsupervised Learning – Dimension Reduction, BIOFAQUIM vs the world	Fabien Plisson Barbara Diaz Eufracio
16:00-22:00	Excursion to Guanajuato, event dinner @Mexico Lindo y Sabroso	
Day 4 – Friday 18th October		
09:00-10:00	Intro to Machine Learning algorithms	Fabien Plisson
10:00-11:00	Supervised Learning: Regressions	Fabien Plisson
11:00-11:30	Coffee break	
11:30-13:00	Practices in regressions with Python / Jupyter Notebook	Edgar López López Barbara Diaz Eufracio
13:00-14:00	Lunch	
14:00-15:00	Supervised Learning: Classifications	Fabien Plisson
15:00-15:30	Coffee break	

15:30-17:00	Practices in Classifications with Python / Jupyter Notebook	Edgar López López Barbara Diaz Eufracio
17:00-17:30	End of Course – Wrap-up and feedbacks	

Registration and payment

In order to be considered for a place on this course applicants must complete the online application form.

Incomplete applications will NOT be considered.

If you have any queries about the registration process please email Guilherme Oliveira and Piv Gopalasingam.

Please note that **the registration fee of $800 MXN includes:**

- Lunches and refreshment breaks
- Accommodation for 5 nights (14-18) at Hotel Ibis Irapuato
- Course materials
- Course dinner on Thursday, October 17th 2019
- Course certificate

Key locations

1. Hotel Ibis Irapuato, Cuarto Cinturón Vial 6802, Ejido de Juarez, 36500 Irapuato, Gto. 462 800 6200 https://maps.app.goo.gl/QbQjHATan1ruSoSU9
2. LANGEBIO, Libramiento Norte, León Km 9.6, 36821 Irapuato, Gto. 462 166 3000 https://maps.app.goo.gl/ZT71fkXtPDnNEXZu9
3. Central De Autobuses Irapuato, Av. 1º de Mayo 436, Tenerias de Irapuato, Irapuato, Gto. 462 626 3708 https://maps.app.goo.gl/H4aYYtgq18EU5DTC7
4. Guanajuato International Airport, Carr. Silao-León Km 5.5, Nuevo México, 36270 Silao, Gto. 472 748 2120 https://maps.app.goo.gl/6zYYqSveYS2VdzTPA

There will be a maximum of 21 participants on this course and registration is dependent on selection after successful completion of the application process.

Registration will close on August 18th 2019 at midnight (CDT)

Appendix:

Speaker information

Dr. Fabien Plisson (organizer), National Laboratory in Genomics for Biodiversity (LANGEBIO), CINVESTAV-IPN Advanced Genomic Unit, Irapuato, Gto. Mexico
http://langebio.cinvestav.mx/en/Dr-Fabien-Plisson

Dr. José Luis Medina-Franco (co-organizer), School of Chemistry, National Autonomous University of Mexico (UNAM), Building F, office 309, CDMX, Mexico https://quimica.unam.mx/perfil/jose-luis-medina-franco/ members of the DIFACQUIM research group MC. Barbara Itzel Diaz-Eufracio and QC. Edgar Lopez Lopez (www.difacquim.com)

Rodrigo Ochoa, PhD(c), Max Planck Tandem Group BioTD, University of Antioquia, Colombia
http://www.udea.edu.co/wps/portal/udea/web/inicio/investigacion/grupos-investigacion/max-planck/biophysics-tropical-diseases
CABANA secondee EMBL-EBI, Thornton's research group