Vincent Mallet

vincent.mallet96@gmail.com
Google Scholar
Github

Education

PhD, *CBIO, Mines -Paristech/Structural Bioinformatics, Institut Pasteur*. **2019 - Present**
Equivariant neural networks methods for protein structure annotation.
Supervised by Jean-Philippe Vert and Michael Nilges. Expected graduation September 2022.
visiting student with Emanuele Rodolà : methods for learning on manifolds.

M.Sc. Computer Science, *McGill University*. **2018 - 2019**
Master of research in the Computational Biology Group supervised by Jérome Waldispühl
Coursework : Machine Learning, Deep Learning, Bioinformatics
GPA : 4.0

M.Eng. Bioinformatics, *Ecole Polytechnique*. **2015 - 2018**
French high-ranked engineering school
Coursework includes : Parallel computing, Big Data, Machine Learning, Bioinformatics
GPA : 3.71, ranked 120^{th}/560

Preparatory Program, *Louis le Grand*. **2013 - 2015**
A two-year post-secondary course in math and physics leading to nationwide competitive entrance examinations to the *Grandes Ecoles* for scientific studies
Admitted to *Ecole Polytechnique*, ranked 39^{th} at the national entrance exam
GPA : 4.0

Research Record

Publications

Vincent Mallet, Luis Checa Ruano, Alexandra Moine Franel, Michael Nilges, Karen Druart, Guillaume Bouvier, and Olivier Sperandio. InDeep: 3D fully convolutional neural networks to assist in silico drug design on protein-protein interactions. *Bioinformatics*, 12 2021. btab849

Vincent Mallet and Jean-Philippe Vert. Reverse-complement equivariant networks for dna sequences. *Advances in Neural Information Processing Systems*, 34, 2021

Vincent Mallet, Carlos Oliver, Jonathan Broadbent, William L Hamilton, and Jérôme Waldispühl. RNAglib: A python package for RNA 2.5D graphs. *Bioinformatics*, 12 2021. btab844

Carlos Oliver, **Vincent Mallet***, Pericles Philippopoulos, William L Hamilton, and Jérôme Waldispühl. VeRNAl: A Tool for Mining Fuzzy Network Motifs in RNA. *Bioinformatics*, 11 2021. btab768

Vincent Mallet, Michael Nilges, and Guillaume Bouvier. quicksom: Self-Organizing Maps on GPUs for clustering of molecular dynamics trajectories. *Bioinformatics*, 11 2020. btaa925

Jacques Boitreaud, **Vincent Mallet***, Carlos Oliver, and Jérome Waldispuhl. OptiMol: Optimization of binding affinities in chemical space for drug discovery. *Journal of Chemical Information and Modeling*, October 2020

Carlos Oliver, **Vincent Mallet**, Roman Sarrazin Gendron, Vladimir Reinharz, William L Hamilton, Nicolas Moitessier, and Jérome Waldispuhl. Augmented base pairing networks encode RNA-small molecule binding preferences. *Nucleic Acids Research*, 48(14):7690–7699, July 2020

Preprints

Vincent Mallet, Carlos G Oliver, Nicolas Moitessier, and Jerome Waldispuhl. Leveraging binding-site structure for drug discovery with point-cloud methods. *arXiv preprint arXiv:1905.12033*, 2019

Attendance to Conferences

2022: ICLR, JOBIM, Iktos Conference

2021: NeurIPS, RECOMB, AI4Health PRAIRIE Winter School

2020: NeurIPS, AI4Health PRAIRIE Winter School

2018: NeurIPS, Riboclub Annual Meeting, Machine Learning for Drug Design Summer School

Funding

2022: ELISE Mobility Award

2019: PhD funding from the INCEPTION Program

2019: Granted AMX PhD funding

2018: Research Grant from the Computational Biology Group

Teaching Experience

2022: 2 hours class in CS Master on geometric learning for drug design, McGill University

2022: 30 hours including classes and practical sessions on Foundations of AI, CRI (Centre de Recherche Interdisciplinaire)

2021-2020: Did twice a 6 hours research seminar on generative models for drug design at Master ISDD, Université de Paris

Research Experience

PhD, *CBIO, Mines Paristech / Structural Bioinformatics, Institut Pasteur*. **Present**

Equivariant networks for structural biology
Enforce SE(3) equivariance in networks whose input lies in 3D space : Cryo-EM data, protein interactions pattern, histopathology. Investigate also the equivariance property for genomic data

Small molecule design for drug design
Use generative models to generate de-novo drug candidates. Include information about the target structure in the optimisation

2.5D graphs to model RNA structure
Represent the RNA interactions as networks of interactions and investigate what this representation enables : drug design, motifs discovery... Develop tools and deep learning layers to deal with this graph representation

Master Thesis, *McGill Computational Biology Group*. **2019**

Protein pocket-ligand matching project
Use learnt representations for small molecules and target structure to learn a mapping from protein pockets structures to their potential binders.
Investigate the role of rotational invariance.

RNA-pocket-ligand interactions Learn a mapping between RNA pockets and potential binders. The representation used for RNA is the 2.5D graphs one. Show the relevance of this representation.

Research Project, *McGill-MILA*. **Fall 2018**

Developed methods for social network data analysis using NLP and sentiment analysis
Scraped Allociné and used data from IMDb, Wikipedia and Twitter

Research Internship, *McGill Computational Biology Group*. **Spring 2018**

Developed binding site comparison and embedding for proteins based on graphs vs spatial point sets. Assessed the efficiency on the DUDE Database
Received excellence award for this internship from Polytechnique

Project Work

Optimisation of Alignments, *Python, Biopython, Scipy*. **Winter 2018**

Studied variations on the Needleman Wunsch algorithm
Used the structural information to improve the alignments
Learnt optimal parameters from reference data

Reinforcement Learning, *Python*. **Winter 2018**

Built the environment and the agent in Crazy Taxi from scratch

Kaggle Challenges, *Python*. **2017**

Porto-seguro-safe-driver-prediction : top 23%
Kkbox-churn-prediction-challenge : top 10%

Team project, *Android Studio, Python, C*. **2017 - 2018**

Built a mobile Android App with a GPS. Connected the smartphone to a Raspberry Pi Installed a e-ink driver on the Raspberry. The images on the e-ink display reflected part of the sunlight on the windshield, projecting an image and creating a head-up display.

Fanfare Simulator, *C++*. **2017**

Developed a platform to play music on their computer together in real time
Team work of 10 people
In charge of the sound interface using PortAudio

Website development, *PHP, HTML/CSS/JS*. **2017**

Created a website similar to TripAdvisor

Professional Experience

Developer recruiter, *Ignition Program*, Paris. **June 2017 - August 2017**

In charge of developing Ignition recruitment of developers. Trained the employees and used web market and web design. Enhanced the traffic and the activity tenfold in two months

Lieutenant, *93rd Regiment of Artillery*, Varces. **2015 - 2016**

Internship in the mountaineering brigade. Commanded a platoon of 55 people. Organised and led training of 30 new recruits

Technical Skills

Languages: Python 3 (Advanced), Java, C++, Bash (intermediate), R, Web (beginner)

Libraries: Data Science : Scikit-learn, Pandas, Scipy, Seaborn, BeautifulSoup, Networkx
Deep learning : PyTorch, DGL, Keras, TensorFlow {1,2}, Sonnet, JAX, Haiku
Bioinformatics : Biopython, RDKit, Pymol
NLP : NLTK, Gensim

Databases: Bioinformatics : DUDE, PDB, ZINC, ChemBL, Fr3D, EMDB
Natural Language Processing : Allociné, IMDb, Wikipedia, Twitter

Developed packages: `rnaglib`, `quicksom`

Languages

French: Native

English: Fluent (TOEFL : 117)

German, Italian: Intermediate

Awards

2018: Jury's congratulations for my research internship

2017: Outstanding Investment Award

2016: Bronze medal of National Defence

Interests

Associations:
In charge of scenography for *Phasm*, an association organising festivals and concerts.
President of an association in charge of fixing materials worth 150,000 euros (pinballs, simulators). Included hardware and maintenance tasks
Vice President of the mountaineering association. Organisation of cross-country skiing, ice climbing and hiking events

Sports: Climbing, Running, Fencing, Skiing