

# HECTOR CLIMENTE-GONZALEZ

+44 07542 973135 ◊ Oxford, United Kingdom

[hclimente@gmail.com](mailto:hclimente@gmail.com) ◊ [linkedin.com/in/hclimente](https://www.linkedin.com/in/hclimente) ◊ [hclimente.eu](http://hclimente.eu)

## ABOUT ME

---

Applied machine learning researcher with 10+ years of experience in biology.  
Leveraging genetics, epigenetics and genomics to tackle open questions in human health.

## EDUCATION

---

|   |             |
|---|-------------|
| <b>Ph.D. in Bioinformatics</b> , Paris Sciences & Lettres University  | 2016 - 2020 |
| <b>Master in Bioinformatics</b> , Pompeu Fabra University             | 2012 - 2014 |
| <b>Bachelor of Biochemistry</b> , Autonomous University of Barcelona  | 2010 - 2013 |
| <b>Bachelor of Biotechnology</b> , Autonomous University of Barcelona | 2008 - 2012 |

## SELECTED WORK EXPERIENCES

---

**Senior Data Scientist** Mar 2023 - Now  
Novo Nordisk Ltd. Oxford, United Kingdom

↔ **Key skills:** Human genetics, Sequence learning, Pipeline development

1. Developing machine learning models for SNP interpretation, drug target discovery and disease risk prediction
2. Data mining large scale datasets (UK Biobank, GTEx or the whole GWAS Catalog)

**Special Postdoctoral Researcher** May 2020 - Feb 2023  
RIKEN AIP Kyoto, Japan

↔ **Key skills:** Deep learning, Nonlinear feature selection, Graph regularization

1. Developed novel deep learning architectures for DNA sequence learning and computer vision
2. Developed graph methods to study RNA-seq gene expression profiles in leukemia
3. Developed strategies for FDR-controlled feature selection

**Ph.D. student** Oct 2016 - Apr 2020  
Institut Curie & Mines ParisTech Paris, France

↔ **Key skills:** Machine learning, Human genetics, Statistical interactions, Kernel methods

1. Developed machine learning methods that leverage graphs to study the genetics of complex diseases
2. Applied them to discover cancer and autoimmune biomarkers in GWAS (single-SNP and epistasis)
3. Authored three software packages: martini (R, in Bioconductor), and pyHSICLasso and spada (Python, in PyPI)

## TECHNICAL SKILLS

---

|                    |   |
|--------------------|---|
| <b>Programming</b> | Python (numpy, pandas, scikit-learn), Deep learning (PyTorch), R, nextflow, C++, Bash |
| <b>Big data</b>    | Accelerated computing (CUDA, HPC), Cloud computing (Azure), Databases (SQL)           |
| <b>Omics</b>       | GWAS, Epigenetics (ATAC-seq), RNA-seq (single cell and bulk), MS-MS proteomics        |
| <b>DevOps</b>      | Virtual environments (Docker, conda), Testing, CI/CD, Open source, git                |

## PROFESSIONAL SKILLS

---

|                             |  |
|-----------------------------|--|
| <b>Communication</b>        | Lead writer of articles and proposals<br>Speaker in international conferences<br>Conductor of tech workshops                   |
| <b>Project management</b>   | Management of projects within and across teams<br>Coordination of AI-driven research project with external partner (Microsoft) |
| <b>Interpersonal skills</b> | Active listener, interested in sharing knowledge<br>Initiated study groups and conversations to improve our coding practices   |

## SELECTED COURSES AND CERTIFICATIONS

---

As an avid self-learner, I took dozens of courses on mathematics, machine learning, formal logic, cloud computing, ethics, project management, theater, and mindfulness. Selected ones:

- MicroMaster in Statistics and Data Science [MITx, Ongoing]
- Deep Learning specialization [Coursera, 2022]
- Oxford Machine Learning Summer School [AI for Global Goals, 2021]
- Data Science Summer School [École Polytechnique, 2017]

## LATEST PUBLICATIONS

---

I have co-authored 14 articles (6 of them as first author), which have been cited more than 800 times.

Climente-González, H., Azencott, C., & Yamada, M. (2023). Network-guided GWAS using stability selection. *STAR Protocols*, 17;4(1).

Poignard, B., Naylor, P. J., Climente-González, H., & Yamada, M. (2022). Feature screening with kernel knockoffs. *International Conference on Artificial Intelligence and Statistics, 1935–1974*.

Duroux, D.<sup>†</sup>, Climente-González, H.<sup>†</sup>, Azencott, C.-A., & Van Steen, K. (2022). Interpretable network-guided epistasis detection. *GigaScience*, 11, giab093.

<sup>†</sup> Equal contribution

Freidling, T., Poignard, B., Climente-González, H., & Yamada, M. (2021). Post-selection inference with HSIC-Lasso. *International Conference on Machine Learning*, 3439–3448.

Climente-González, H., Lonjou, C., Lesueur, F., GENESIS study group, Stoppa-Lyonnet, D., Andrieu, N., & Azencott, C.-A. (2021). Boosting GWAS using biological networks: A study on susceptibility to familial breast cancer. *PLOS Computational Biology*, 17(3), e1008819.

## AWARDS AND RECOGNITIONS

---

2022 Joined the [AIME registry](#)'s steering committee

2020 Competitive postdoctoral fellowship from RIKEN: Special Postdoctoral Researcher (SPDR)

2018 Scholarship to carry out a Ph.D. internship at RIKEN AIP

2016 Competitive Ph.D. fellowship from the European Union: Marie Skłodowska-Curie Actions COFUND

2013 Master internship fellowship from the Research Programme on Biomedical Informatics

2013 First class honors of the 2013 graduating class of the B.Sc. in Biochemistry