Jaime A Castro-Mondragon

BIOINFORMATICIAN

Oslo, Norway

✉ jamondra@uio.no | 🏠 jaimicore.github.io | ⬡ jaimicore | in jaime-castro-mondragon

I am a computational biologist with 11 years of experience in research, sequence analysis, methods development, and data visualisation. I have worked in both the academic and private sector. My main expertise is the development of algorithms to analyze multivariate data and apply these methods in a wider range of biomedical fields such as vaccine design or gene expression.

Working Experience

Akershus Universitetssykehus (Ahus) — *Oslo, Norway*
BIOINFORMATICIAN — *2025 - Present*

- Development, and maintenance of bioinformatic workflows (prediction of single nucleotide variants, structural variants, copy number alterations and gene fusions).
- Development of reports to be shared with molecular biologist, pathologists and oncologists with the summary of the mutational status of cancer samples.

Nykode Therapeutics AS — *Oslo, Norway*
BIOINFORMATICIAN (SCIENTIST II) — *2022 - 2025*

- Design and analysis of RNA-seq experiments to detect statistically significant changes in gene expression in human and mouse cells transfected with a vaccine candidate.
- Development, establishment, and maintenance of bioinformatic analysis pipelines (prediction of MHC-binding peptides, protease cleavage sites, RNA secondary structure).
- Train ML models to highlight protein features explaining experimental observations.
- Computational design of vaccine candidates (protein sequences) in 5 projects.
- Design of a NoSQL (JSON) protein sequence database.

University of Oslo — *Oslo, Norway*
POST-DOCTORAL RESEARCHER — *2017 - 2022*

- Co-lead the JASPAR database team (20 people) in the 2020 and 2022 releases (data QC, manual curation of results, development of motif analysis pipeline (R/python/Snakemake) and added motif clusters to the website).
- Development of a multi-omics (WES, WGS, RNA-seq, sRNA-seq, CNAs) integration method (written in R/Snakemake) to calculate the likely association of single-nucleotide variants with changes in gene expression in gene regulatory networks.
- +30 verified peer-reviews for scientific journals.
- Mentoring of 2 PhD and 3 master students.
- Supervisor: Anthony Matelier.

Toolkit

What	Level	Category
Python	Proficient	Programming language
R	Proficient	
Perl	Proficient	
Bash	Intermediate	
HTML / CSS	Intermediate	
D3	Intermediate	
JavaScript	Beginner	
SQL	Beginner	Databases
Snakemake	Proficient	Pipeline Development
GNU make	Intermediate	
nextflow	Beginner	
Docker	Intermediate	Containerization
Apptainer (Singularity)	Intermediate	
Git	Proficient	Version Control
aws S3	Beginner	Cloud
CI/CD	Beginner	Automation

Technical Skills: clustering, dimensionality reduction, exploratory data analysis (EDA), data wrangling, multivariate analysis, statistical tests, Markov chains, ML (Random Forest, Linear Regression, Logistic Regression, model evaluation), Deep Learning (neural networks), Data Curation, pattern discovery, pattern matching.

Tools: jupyter notebooks, Rstudio, emacs, VS code, google docs, Linux (Ubuntu).

Education

Aix-Marseille Universite
Marseille, France
PhD in bioinformatics
2014 - 2017
- Development of a computational methods to cluster transcription factor motifs. Development of a computational methods to detect enrichment of sequences. Design and implementation of software to create negative controls in DNA sequences. Supervisor: PhD Jacques van Helden. 6 publication derived from this thesis.

Institute de biology de l'Ecole Normale Superieure
Paris, France
Master in systems biology
2013 - 2014
- Development of a method to cluster transcription factor binding motifs. Supervisor: PhD Jacques van Helden.

Universidad Nacional Autonoma de Mexico
Cuernavaca, Mexico
Bachelor in genomic sciences
2009 - 2013
- Detection of transcription factor binding motifs for uncharacterized transcritpion factors using sequence homology. Supervisors: PhD Julio Collado-Vides and PhD Alejandra Medina-Rivera.

Languages

Spanish	English	French	Norwegian
Native	Fluent	Fluent	Intermediate (A2)

Publications

See the complete list of publications (including preprints) and citations in my google scholar profile: *pcevKk0AAAA*

ORCID profile: 0000-0003-4069-357X

* = equal contributions (first-author)

^ = equal contributions (second-author)

= co-corresponding authors

Peer-reviewed

1. Taboada-Castro H, Hernandez-Alvarez AJ, **Castro-Mondragon JA**, and Encarnacion-Guervara S. *RhizoBindingSites v2.0 Is a Bioinformatic Database of DNA Motifs Potentially Involved in Transcriptional Regulation Deduced From Their Genomic Sites*. Bioinformatics and Biology Insights. (2024)

2. Rauluseviciute I, Launay T, Barzaghi G, Nikumbh S, Lenhard B, Krebs AR, **Castro-Mondragon JA**, and Mathelier A. *Identification of transcription factor co-binding patterns with non-negative matrix factorization.* Nucleic Acids Research (2024)

3. Rauluseviciute I*, Riudavets-Puig R*, Blanc-Mathieu R^, **Castro-Mondragon JA**^, Ferenc K^, Kumar V^, Berhanu Lemma R^, Lucas J^, Cheneby J, Baranasic D, Khan A, Fornes O, Gundersen S, Johansen M, Hovig E, Lenhard B#, Sandelin A#, Wasserman WW#, Parcy F#, and Mathelier A#. *JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles.* Nucleic Acids Research (2024)

4. **Castro-Mondragon JA**, Ragle Aure M, Lingjærde OC, Langerød A, Martens JWM, Børresen-Dale AL, Kristensen VN, and Mathelier A. *Cis-regulatory mutations associate with transcriptional and post-transcriptional deregulation of gene regulatory programs in cancers.* Nucleic Acids Research (2022)

5. Saha S*, Spinelli L*, **Castro-Mondragon JA**, Kervadec A, Lynott M, Kremmer L, Roder L, Sallouha J, Torres M, Brun C, Vogler G, Bodmer R, Colas AR#, Ocorr K#, and Perrin L#. *Genetic architecture of natural variations of cardiac performance in flies*. eLife (2022)

6. Santana-Garcia W*, **Castro-Mondragon JA***, Padilla-Galvez M, Nguyen NTT, *Elizondo-Salas A, Ksouri N, Gerbes F, Thieffry D, Vincens P, Contreras-Moreira B#, van Helden J#, Thomas-Chollier M#, Medina-Rivera A#*. RSAT 2022: regulatory sequence analysis tools*. Nucleic Acids Research (2022)

7. **Castro-Mondragon JA***, Riudavets-Puig R*, Rauluseviciute I*, Berhanu Lemma R, Turchi L, Blanc-Mathieu R, Lucas J, Boddie P, Khan A, Manosalva Perez N, Fornes O, Leung TY, Aguirre A, Hammal F, Schmelter D, Baranasic D, Ballester B, Sandelin A#, Lenhard B#, Vandepoele K, Wasserman WW#, Parcy F#, and Mathelier A#. *JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles*. Nucleic Acids Research (2021)

8. Santiago-Algarra D, Souaid C, Singh H, Dao TML, Hussain S, Medina-Rivera A, Ramirez-Navarro L, **Castro-Mondragon JA**, Sadouni N, Charbonnier G, Spicuglia S. *Epromoters function as a hub to recruit key transcription factors required for the inflammatory response*. Nature Communications (2021)

9. Riudavets-Puig R, Boddie P, Khan A, **Castro-Mondragon JA**, and Mathelier A. *UniBind: maps of high-confidence direct TF-DNA interactions across nine species*. BMC Genomics (2021)

10. Ragle Aure M, Fleischer T, Bjørklund S, Ankill J, **Castro-Mondragon JA**, OSBREAC (Oslo Breast Cancer Research Consortium), Børresen-Dale AL, Tost J, Sahlberg KK, Mathelier A, Tekpli X#, Kristensen VN#. *Crosstalk between microRNA expression and DNA methylation drives the hormone-dependent phenotype of breast cancer*. Genome Medicine (2021)

11. Ksouri N, **Castro-Mondragon JA**, Montardit-Tarda F, van Helden J, Contreras-Moreira B#, and Gogorcena Y#. *Tuning promoter boundaries improves regulatory motif discovery in non-model plants: the peach example*. Plant Physiology (2021)

12. Taboada-Castro H, **Castro-Mondragon JA**, Aguilar-Vera A, Hernandez-Alvarez AJ, van Helden J, and Encarnacion-Guervara S. *RhizoBindingSites, a Database of DNA-Binding Motifs in Nitrogen-Fixing Bacteria Inferred Using a Footprint Discovery Approach*. Frontiers in Microbiology (2020)

13. Fornes O*, **Castro-Mondragon JA***, Khan A*, van der Lee Robin, Zhang X, Richmond PA, Modi BP, Correard S, Gheorghe M, Baranasic D, Santana-Garcia Walter, Tan G, Cheneby J, Ballester B, Parcy F, Sandelin A#, Lenhard B#, Wasserman WW#, and Mathelier A#. *JASPAR 2020: update of the open-access database of transcription factor binding profiles*. Nucleic Acids Research (2019)

14. Nguyen NTT*, Contreras-Moreira B*, **Castro-Mondragon JA**, Santana-Garcia W, Ossio R, Robles-Espinoza CD, Bahin M, Collombet S, Vincens P, Thieffry D, van Helden J#, Medina-Rivera A#, Thomas-Chollier M#. *RSAT 2018: regulatory sequence analysis tools 20th anniversary*. Nucleic Acids Research (2018)

15. Taboada H*, Meneses N*, Dunn MF*, Vargas-Lagunas C, Buchs N, **Castro-Mondragon JA**, Heller M, and Encarnacion S. *Proteins in the periplasmic space and outer membrane vesicles of Rhizobium etli CE3 grown in minimal medium are largely distinct and change with growth phase*. Microbiology (2018)

16. Khan A*, Fornes O*, Stigliani A*, Gheorghe M, **Castro-Mondragon JA**, van der Lee R, Bessy A, Chèneby J, Kulkarni S, Tan G, Baranasic D, Arenillas D, Sandelin A#, Vandepoele K, Lenhard B#, Ballester B, Wasserman W#, Parcy F, Mathelier A#. *JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework*. Nucleic Acids Research (2017)

17. Dao LM*, Galindo-Albarran AO*, **Castro-Mondragon JA**^, Andireu-Soler C^, Medina-Rivera A^, Souadi C, Charbonnier G, Griffon A, Vanhille L, Stephen S, Alomairi J, Soler C, Stephen T, Martin D, Torres M, Fernandez N, Soler E, van Helden J, Puthier D, Spicuglia S. *Genome-wide characterization of mammalian promoters with distal enhancer functions*. Nature Genetics (2017)

18. **Castro-Mondragon JA**, Jaeger S, Thieffry D, Thomas-Chollier M#, and van Helden J#. *RSAT matrix-clustering: dynamic exploration and redundancy reduction of transcription factor binding motif collections*. Nucleic Acids Research (2017)

19. **Castro-Mondragon JA***, Rioualen C*, Contreras-Moreira B, van Helden J. *RSAT::Plants: Motif Discovery in ChIP-Seq Peaks of Plant Genomes*. Plant Synthetic Promoters - Springer Protocol (2016)

20. Contreras-Moreira B#, **Castro-Mondragon JA**, Rioualen C, Cantalapiedra CP, van Helden J. *RSAT::Plants: Motif Discovery Within Clusters of Upstream Sequences in Plant Genomes*. Plant Synthetic Promoters - Springer Protocol (2016)

21. Gama-Castro S*, Salgado H*, Santos-Zavaleta A, Ledezma-Tejeida D, Muñiz-Rascado L, García-Sotelo JS, Alquicira-Hernández K, Martínez-Flores I, Pannier L, **Castro-Mondragon JA**, Medina-Rivera A, Solano-Lira H, Bonavides-Martínez C, Pérez-Rueda E, Alquicira-Hernández S, Porrón-Sotelo L, López-Fuentes A, Hernández-Koutoucheva A, Del Moral-Chávez V, Rinaldi F, Collado-Vides J. *RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond*. Nucleic Acids Research (2015)

22. Medina-Rivera A*, Defrance M*, Sand O*, Herrmann C, **Castro-Mondragon JA**, Delerce J, Jaeger S, Blanchet C, Vincens P, Caron C, Staines DM, Contreras-Moreira B, Artufel M, Charbonnier–Khamvongsa L, Hernandez C, Thieffry D, Thomas-Chollier M#, van Helden J#. * RSAT 2015 : Regulatory Sequence Analysis Tools*. Nucleic Acids Research (2015)

23. Rogel MA, Bustos P, Santamaría RI, González V, Romero D, Miguel AC, Lozano L, **Castro-Mondragon JA**, Martínez-Romero J, Ormeño-Orrillo E, Martínez-Romero E. *Genomic basis of symbiovar mimosae in Rhizobium etli*. BMC Genomics (2014)

Preprints

No preprints at this moment.

Reviewed manuscripts

Web of Science (prev. Publons) profile: 1499198

+30 reviewed manuscripts in the following journals: Bioinformatics, Genetics, Molecular Plant, Nucleic Acids Research, Genome biology, Genome Biology and Evolution, Frontiers in Genetics, eLife.

Personalia

Date of birth

13/02/90

Citizenship

Mexican

Languages

Spanish (native speaker), English (fluent), French (fluent), Norwegian (intermediate, A2 level)

References

Professor Jacques van Helden

PHD SUPERVISOR, COLLABORATOR
• Jacques.van-Helden@france-bioinformatique.fr (https://orcid.org/0000-0002-8799-8584)

Aix-Marseille Universite, Marseille, France

PhD Alejandra Medina-Rivera

BACHELOR THESIS SUPERVISOR, COLLABORATOR
• amedina@liigh.unam.mx (https://liigh.unam.mx/amedina/)

Laboratorio Internacional de Investigacion sobre el Genoma Humano, Queretaro, Mexico

PhD Anthony Mathelier

POSTDOC SUPERVISOR, COLLABORATOR
• anthony.mathelier@ncmm.uio.no (https://mathelierlab.com/)

Norway Centre of Molecular Medicine, Oslo, Norway

PhD Andreas Midbøe Hoff

HEAD OF IN SILICO AND ARTIFICIAL INTELLIGENCE DEPARTMENT
• amhoff@nykode.com (linkedin.com/in/andreas-midbøe-hoff-853366b0)

Nykode Therapeutics AS, Oslo, Norway