List of publications

Peer reviewed

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

I launched the gene ontology analysis and prepared some figures, this is a continuation of paper 4.

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

I launched the motif analysis to identify TF motifs and TF binding sites in promoter regions of Rhizobiales genomes.

11.   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#. (2019) *JASPAR 2020: update of the open-access database of transcription factor binding profiles* Nucleic Acids Research

I developed the motif discovery pipeline and led the motif curation process. We created a novel collection with motifs with good quality but without literature support.

10.   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#. (2018) *RSAT 2018: regulatory sequence analysis tools 20th anniversary* Nucleic Acids Research

The algorithms that I previously developed during my PhD (motif clustering + local enrichment) were released in this version

9.   Taboada H*, Meneses N*, Dunn MF*, Vargas-Lagunas C, Buchs N, **Castro-Mondragon JA**, Heller M, and Encarnacion S# (2018) *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.

I contributed with the functional annotation.

8.   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. (2017) *JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework* Nucleic Acids Research

I adapted the motif clustering algorithm I developed to visualize hundreds of motifs as a radial tree and added colors to represent TF families.

7.   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 (2017) *Genome-wide characterization of mammalian promoters with distal enhancer functions* Nature Genetics

I developed a motif enrichment algorithm (RSAT position-scan) to find TFBSs that deviate an expected uniform distribution around TSS of promoters with potential enhancer function.

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

My main PhD project. An interactive motif clustering algorithm. We showed that combining different motif-comparison metrics we achieved a better partition of the clusters.

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

I wrote a pipeline for motif analysis (motif discovery + clustering, detection of TF binding sites, design of negative controls) in plant genomes.

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

I designed the negative controls (e.g., shuffled sequences, random genomic regions)

3.   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-Mondragón 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 (2015) *RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond.* Nucleic Acids Research

I designed a radial tree representation of the TF binding motifs of *E coli K12* grouped by TF families.

2.   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# (2015) *RSAT 2015 : Regulatory Sequence Analysis Tools*. Nucleic Acids Research

In this release, we added the first (draft) version of the motif clustering algorithm.

1.   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# (2014) *Genomic basis of symbiovar mimosae in Rhizobium etli.* BMC Genomics

I contributed to the genome annotation of this organism by finding open-reading frames.

Preprint

3.   Riudavets Puig R, Boddie P, Khan A, **Castro-Mondragon JA**, and Mathelier A# (2020) *UniBind: maps of high-confidence direct TF-DNA interactions across nine species.* bioRxiv.

I contributed to the development of TFBS enrichment tool and its interactive output.

 2.   **Castro-Mondragon JA**, Ragle Aure M, Lingjærde OC, Langerød A, Martens JWM, Børresen-Dale AL, Kristensen VN, and Mathelier A# (2020) *Cis-regulatory mutations associate with transcriptional and post-transcriptional deregulation of the gene regulatory program in cancers.* bioRxiv.

Integration of multiple-omics (WGS, RNA-seq, sRNA-seq, CNA). We studied 8 cancer cohorts and used a bayesian network to associate TFBS mutations with changes in gene expression in miRNA networks. Beside the analysis, this is a pipeline that could be used by external users.

1.   Ragle Aure M, Fleischer T, Bjørklund S, Ankill J, **Castro-Mondragon JA**, OSBREAC (Oslo Breast Cancer Research Consortium), Børresen-Dale AL, Sahlberg KK, Mathelier A, Tekpli X#, Kristensen VN# (2020) *Crosstalk between microRNA expression and DNA methylation drive the hormone-dependent phenotype of breast cancer.* bioRxiv.

I contributed with the normalization and processing of the sRNA-seq data.