Diogo M. Camacho

Lead, Predictive BioAnalytics

Wyss Institute for Biologically Inspired Engineering @ Harvard University

+1 617 945 4383 | diogo.camacho@wyss.harvard.edu | diogocamacho | diogocamacho | DiogoMCamacho

Computational Systems Biology, Machine Learning, Bioinformatics

Qualifications and research interests

Highly effective computational system biologist, with graduate and post-doctoral work focusing on network inference and machine learning. Industry experience in development and implementation of computational tools for multi omics data analysis (including next-generation sequencing, metabolomics, proteomics), drug discovery, and target identification. I am interested in the application of machine learning/deep learning tools and techniques in the context of drug discovery, disease biology characterization, large data analytics for biology, while focused on bridging the gap between the computational and experimental labs through highly engaging and fruitful collaborations.

Technical Skills

Machine learning, multi omics data analytics, R/Bioconductor, keras, perl, Latex, python, MATLAB, awk, bash. Adept user of OS X/macOS, Unix. Familiarity with cloud computing architectures (AWS) and high performance computing environments.

Education

Virginia Polytechnic Institute and State University *Blacksburg, VA*
Ph.D. in Genetics, Bioinformatics, and Computational Biology *2007*

Faculdade de Ciencias da Universidade de Lisboa *Lisboa, Portugal*
B. Sc. in Biochemistry *2002*

Experience

Wyss Institute @ Harvard University *Boston, MA*
Lead, Predictive BioAnalytics Initiative, Advanced Technology Team *July 2016 - Present*

I am currently the group lead of the Predictive BioAnalytics Initiative at the Wyss Institute, where I lead a team of computational biologists, computer scientists, and software engineers to address challenges in machine learning and biomedical data sciences, both in-house or thtough external collaborations. Some of the functions associated with the role include:

- Development and implementation of research strategy for the Initiative, focusin on enabling ML/DL/AI capabilities
- Managing and mentoring staff scientists, post-doctoral fellows, graduate students, and interns
- Writing and managing federal grant applications and grant awards with DARPA, NIH
- Hands-on development of algorithms and computational approaches for dissemination internally and with corporate partners via R Shiny applications
- Responsible for the implementation of diverse tools for analysis of high throughput data (transcriptomics, RNA-seq, metabolomics, proteomics, 16S rDNA sequencing) in support of diverse grant work

Evelo Biosciences *Cambridge, MA*
Senior Scientist, Computational Systems Biology Lead *January 2015 - April 2016*

As the lead Computational Systems Biologist at Evelo Biosciences, I was involved in the build out of the computational capabilities of the company, in support of preclinical development of microbiome-focused therapeutics for oncology. Some of the functions of the role included:

- Implementation of diverse tools for analysis of high throughput data (transcriptomics, RNA-seq, metabolomics) as well as a 16S rDNA sequencing analysis pipeline

- Responsible for the identification of novel therapeutic opportunities based on in-house and publicly available data
- Implementation of analytical software tools to be used by bench scientists in R Shiny
- Responsible for the interface with IT provider to delineate and expand computational capabilities of the company, from general to research needs

Symbiota (now Indigo Ag) *Cambridge, MA*
Consultant, Computational Systems Biology Lead *December 2014 - January 2015*

In a short engagement with Symbiota, I developed an R Shiny application for pathway enrichment for plant transcriptomics data.

Ember Therapeutics *Cambridge, MA*
Principal Scientist, Computational Systems Biology Lead *January 2014 - December 2014*

At Ember Therapeutics I was the Computational Systems Biology lead, where I oversaw the implementation and deployment of novel analytical tools for the identification of novel target opportunities for increased energy expenditure, not only from publicly available data but also from data coming from internal efforts. Some of the functions of the role included:

- Responsible for the implementation of diverse tools for analysis of high throughput data (transcriptomics, RNA-seq, metabolomics)
- Implementation of a knowledge based and data driven construction of screening libraries for recombinant proteins, small peptides, and RNAi therapeutic efforts
- Implementation of analytical strategies (QC, statistical analyses, hit-calling) for high throughput screens
- Responsible for the interface with IT provider to delineate and expand computational capabilities of the company, from general to research needs

Pfizer, Inc *Cambridge, MA*
Senior Research Scientist, Computational Sciences Center of Emphasis *January 2011 - January 2014*

I was a member of the Computational Sciences Center of Emphasis at Pfizer, providing computational support across different preclinical programs at the organization, from cardiovascular disease to pain management and drug repositioning efforts. Some of the functions of the role included:

- Development and implementation of a network analysis tool for the characterization of differential networks in healthy and diseased populations under the scope of metabolic diseases
- Implementation of network inference tools for the analysis of large-scale data sets
- Development and implementation of a methodology for metabolite set enrichment analysis for metabolomics data
- Performed data analysis in transcriptomics, proteomics and metabolomics for different partners within the organization with particular emphasis in Cardiovascular and Metabolic diseases

Howard Hughes Medical Institute @ Boston University *Boston, MA*
Post-doctoral Fellow *July 2007 - January 2011*

Post-doctoral training with Dr. James Collins at Boston University, focusing on the application of machine learning and network inference approaches in biomedicine. Some efforts included:

- Developed a network inference algorithm for identification of regulatory architectures of pathways
- Identified a novel mechanism of action for antifungal drugs using transcriptomics and metabolomics data
- Identified and characterized the small RNA regulatory network in bacterial systems using gene expression data
- Performed data analysis of gene expression data and metabolomics data in bacterial and fungal systems
- Consulted in the implementation and usage of inference algorithms across platforms for biotechnology and pharmaceutical companies in the Boston area

Awards and Grants

Synergistic Discovery and Design (DARPA) *$2,000,000*
co-PI *9/1/17*

Publications

- Bojar, D, Powers, RK, **Camacho, DM**, Collins, JJ (2020), Deep learning reveals glycan-mediated host-microbe interactions, *Cell Host & Microbe* (*submitted*)
- Bojar, D, **Camacho, DM**, Collins, JJ (2019), Using natural language processing to learn the language of glycans, *Nature communications* (*submitted*)
- Valeri, J, Collins, Ramesh, P, KM, Alcantar, M, Lepe, BA, Lu, TK, **Camacho, DM** (2019), Sequence-to-function deep learning frameworks for engineered riboregulators, *Nature communications* (*submitted*)
- Jalili-Firoozinezhad, S, Gazzaniga, FS, Calamari, EL, **Camacho, DM**, Fadel, C, Nestor, B, Cronce, MJ, Tovaglieri, A, Levy, O, Gregory, KE, Breault, DT, Cabral, JMS, Kasper, DL, Novak, R, Ingber, DE (2019), A complex human gut microbiome cultured in an anaerobic intestine-on-a-chip, *Nature Biomedical Engineering*, 3, 520-531 [PubMed]
- Tovaglieri, A, Sontheimer-Phelps, A, Geirnaert, A, Prantil-Baun, R, **Camacho, DM**, Chou, DB, Jalili-Firoozinezhad, S, de Wouters, T, Kasendra, M, Super, M, Cartwright, M, Richmond, CA, Breault, DT, Lacroix, C, Ingber, DE (2019), Species-specific enhancement of enterohemorrhagic *E. coli* pathogenesis mediated by microbiome metabolites, *Microbiome*, 7, 43 [PubMed]
- **Camacho, DM**, Collins, KM, Powers, RK, Costello, JC, Collins, JJ (2018), Next-generation machine learning for biological networks, *Cell*, 173, 1581-1592 [PubMed]
- Musah, S, Dimitrakakis, N, **Camacho, DM**, Church, GM, Ingber, DE (2018), Directed differentiation of human induced pluripotent stem cells into mature kidney podocytes and establishment of a Glomerulus Chip, *Nature Protocols*, 13, 1662-1685 [PubMed]
- Paandey, SP, Winkler, JA, Li, H, **Camacho, DM**, Collins, JJ, Walker, GC (2014), Central role for RNase YbeY in Hfq-dependent and Hfq-independent small-RNA regulation in bacteria, *BMC Genomics*, 15, 121 [PubMed]
- Galagan, JE, Minch, K, Peterson, M, Lyubetskya, A, Azzizi, E, Sweet, L, Gomes, A, Rustad, T, Dolganov, G, Glotova, I, Abeel, T, Mawhinney, C, Kennedy, A, Allard, R, Brabant, W, Krueger, A, Jaini, S, Honda, B, Yu, W-H, Hickey, M, Zucker, J, Garay, C, Weiner, B, Sisk, P, Stolte, C, Winkler, J, Van de Peer, Y, Iazzetti, P, **Camacho, D**, Dreyfuss, J, Liu, Y, Dorhoi, A, Mollenkopf, H-J, Drogaris, P, Lamontagne, J, Zhou, Y, Piquenot, J, Park, ST, Raman, S, Kaufmann, S, Mohney, R, Chelsky, D, Moody, B, Sherman, D, Schoolnik, G (2013), The Mycobacterium tuberculosis regulatory network and hypoxia, *Nature*, 499, 178-183 [PubMed]
- Belenky, P, **Camacho, D**, Collins, JJ (2013), Fungicidal drugs induce a common oxidative-damage cellular death pathway, *Cell Reports*, 3, 350-358 [PubMed]
- Marbach, D, Costello, JC, Kuffner, R, Vega, N, Prill, RJ, **Camacho, DM**, Allison, KR, the DREAM5 Consortium, Kellis, M, Collins, JJ, Stolovitzky, G (2012), Wisdom of crowds for robust gene network inference, *Nature Methods*, 9, 796-804 [PubMed]
- Dwyer, DJ, **Camacho, DM**, Callura, JM, Kohanski, MA, Collins, JJ (2011), Antibiotic-induced bacterial cell death exhibits physiological and biochemical hallmarks of apoptosis, *Molecular Cell*, 46, 561-572 [PubMed]
- Modi, SR, **Camacho, DM**, Kohanski, MA, Collins, JJ (2011), Functional characterization of bacterial sRNAs using a network biology approach, *Proc. Natl. Acad. Sci. USA*, 108, 15522-15527 [PubMed]
- **Camacho, DM**, Collins, JJ (2009), Systems biology strikes gold, *Cell*, 137, 24-26 [PubMed]
- **Camacho, D**, Vera-Licona, P, Laubenbacher, R, Mendes, P (2007), Comparison of existing reverse engineering methods by use of an in silico system, *Ann. N. Y. Acad. Sci.*, 1115, 73-89 [PubMed]
- Mendes, P, **Camacho, D**, de la Fuente, A (2005), Modelling and simulation for metabolomics data analysis, *Biochem. Soc. Trans.*, 33, 1427-1429 [PubMed]
- **Camacho, D**, de la Fuente, A, Mendes, P (2005), The origin of correlations in metabolomics data, *Metabolomics*, 1, 53-63 [Link]
- Martins, AM, **Camacho, D**, Shuman, J, Sha, W, Mendes, P, Shulaev, V (2004), A systems biology study of two distinct growth phases of *Saccharomyces cerevisiae* cultures, *Curr. Genomics*, 5, 649-663 [Link]

Patents

- Soenksen, LR, **Camacho, DM**, Collins, KM, Angenent-Mari, NM, Garruss, AS, Collins, JJ, Church, GM, Method for deep learning for RNA synthetic biology (*submitted*)
- Collins, JJ, Belenky, P, **Camacho, DM** (2015), Compositions and methods for treating fungal infections, WO2014130922A1 [Link]

References

References will be provided upon request.