IGB 3115, 1206 W Gregory Dr, Urbana, IL 61801
+1 (650) 868 2314
mhernaez@illinois.edu
mikelhernaez

Mikel Hernaez

Position

Carl R. Woese Institute for Genomic Biology *University of Illinois at Urbana-Champaign, IL, USA*
Director, Computational Genomics ***05/2017–Present***

Education

Stanford University *CA, USA*
Postdoctoral researcher in Electrical Engineering ***09/2013–12/2016***

- Developed compression tools for genomic data
- Performed extensive analysis on the impact of lossy compression of genomic data on variant calling
- Explored new approaches for performing alignment and variant calling on genomic data
- Characterized long non-coding RNAs across human cancer using unsupervised learning
- P.I.: Prof. Tsachy Weissman (tsachy@stanford.edu)
- Collaborated with:
  - Prof. Euan Ashley, M.D. (euan@stanford.edu), Stanford Medical School
  - Prof. Olgica Milencovic (milenkov@illinois.edu), Electrical and Computer Eng. Dept.
  - Prof. Olivier Gevaert (ogevaert@stanford.edu), Stanford Data Science in Biomedicine Dept.

TECNUN, University of Navarra *Spain*
PhD in Electrical Engineering, *GPA: Summa Cum Laude* ***09/2010–12/2012***

- Topics: Conducted research in information theory and coding, communication systems and signal processing, with special focus on iterative channel codes (LDPC, turbo codes...)
- Dissertation: Joint Network-Channel Coding Schemes for Relay Networks
  - Advisors: Prof. Pedro Crespo (pcrespo@tecnun.es) and Javier Del Ser (jdelser@tecnalia.es)

TECNUN, University of Navarra *Spain*
Telecommunications Engineering Master Degree ***09/2003–02/2009***

- Ranking position: Top 10
- Master thesis: Concatenated LDGM Codes for the Transmission of Correlated Sources over Gaussian Broadcast Channels *(GPA: 10/10)*
  - Advisor: Prof. Pedro M. Crespo (pcrespo@ceit.es)

Lulea Tekniska Universitet *Sweden*
Erasmus Program ***08/2007–01/2008***

Interests

Data Compression, Bioinformatics, Machine learning, Information Theory and Coding, Signal Processing.

Experience

Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign *IL, USA*
Director of Computational Genomics ***2017–***

- Advice leadership on the computational biology matters.
- Co-PI of the Mayo Grand Challenge project: AI collaboration between Mayo Clinic and the University of Illinois. Lead PI of the genomic data compression front.
- Develop statistical methods for RNA-Seq data in collaboration with the Carle College of Medicine and Stanford University.

Stanford University *CA, USA*
Group of Prof. Tsachy Weissman ***2013–2016***

I worked on the design and development of new algorithms to improve the distribution and storage of genomic data, to facilitate its access, and to boost the inferential power of analysis performed on it. My approach combines tools from information theory, statistics, and machine learning.

- Contributions:
  - Proposed, in collaboration with MIT and EPFL, a methodology for analysis of genomic data compression on Variant Calling that set the bases for the Standardization process of genomic information
  - Designed lossless and lossy compressors for genomic data
  - Developed a denoiser to reduce noise present in genomic data

ENIGMEDIA *Spain-USA*
Director of Research ***03/2013 - 09/2013***

- Supervisor: CEO & Founder Gerard Vidal (gerard@enigmedia.com)
- Worked on encrypted real-time communications based on chaos-based stream-ciphers

NAUI SYSTEMS *CA, USA*
Consultant ***04/2013 - 06/2013***

- Helped with the possibility of implementing a new coding solution for RAM memories

STANFORD UNIVERSITY-TECNUN *USA-Spain*
Visiting Researcher ***Summer 2012***

- Supervisor: Golan Yona (golan.yona@stanford.edu)
- Conducted research in biological relational databases under the BIOZON project (http://biozon.org)

CEIT, Centre of Studies and Technical Research of Gipuzkoa *Spain*
Research Assistant in the Electrical Engineering Department ***2009–2012***

- Supervisor: Prof. Pedro M. Crespo (pcrespo@ceit.es)
- Proposed several practical coding schemes for relay channels using LDPC and Turbo codes
- Set up of a point-to-point wireless communication system for pedagogic purposes

TECNUN, University of Navarra *Spain*
Collaborator Student in the Electrical Engineering Department ***2005***

- Worked on Cadence circuit design and lay-out

Teaching experience

TECNUN, University of Navarra *CA, USA*
Lecturer ***09/2012 - 01/2013***

- Information Theory and Coding
- Communication Systems
- Fundamentals of Computers course

TECNUN, University of Navarra *Spain*
Teaching Assistant ***2011–2012***

- Information Theory and Coding
- Communication Systems
  - Set up of a point-to-point wireless communication system for pedagogic purposes

TECNUN, University of Navarra *Spain*
Advisor of a Master Thesis ***2010–2011***

- Topic: Implementation of a Software Development Kit for Communications System.
- Resulted in a IEEE publication.

Research Grants

Awarded an **Strategic Research Initiative (SRI) grant** from the University of Illinois at Urbana-Champaign (UIUC) *2018*

Awarded an **Chan Zuckerberg Initiative (CZI) grant**, under the Human Cell Atlas *2018*

Awarded an **Mayo Grand Challenge grant**, from Mayo Clinic *2017*

Awarded an **NIH grant**, under the BD2K initiative, in collaboration with the University of Illinois at Urbana-Champaign (UIUC) *2015*

Journal Papers

- L. Rogusky, I. Ochoa, **M. Hernaez**, S. Deorowicz, *FaStore - a space-saving solution for raw sequencing data*, **In review, bioRxiv 168096**, doi: https://doi.org/10.1101/168096, 2017.

- J. Voges, J. Oesterman, **M. Hernaez**, *CALQ: compression of quality values of aligned sequencing data*, **Bioinformatics, to appear**, 2017.

- I. Ochoa, **M. Hernaez**, R. Goldfeder, T. Weissman and E. Ashley, *Effect of lossy compression of quality scores on variant calling*, **Briefings in Bioinformatics**, 2016.

- S. Deorowicz, S. Grabowski, I. Ochoa, **M. Hernaez** and T. Weissman, *Comment on: "ERGC: An efficient referential genome compression algorithm"*, **Bioinformatics**, btv704, 2015.

- G. Malysa, **M. Hernaez**, I. Ochoa, M. Rao, K. Ganesan and T. Weissman, *QVZ: lossy compression of quality values*, **Bioinformatics**, btv330, 2015.

- I. Alustiza, **M. Hernaez**, P. Crespo, *Design of a new scheme for multi-hop wireless networks*

using decode-and-forward strategy, **EURASIP Journal on Wireless Communications and Networking**, (1), 1-8, 2015

- I. Ochoa, **M. Hernaez** and T. Weissman, *Aligned genomic data compression via improved modeling*, **Journal of bioinformatics and computational biology**, Vol. 12, No. 6, 2014.

- I. Ochoa, **M. Hernaez** and T. Weissman, *iDoComp: a compression scheme for assembled genomes*, **Bioinformatics**, btu698, 2014.

- **M. Hernaez**, P.M. Crespo, J. Del Ser, *On the Design of a Novel Joint Network-Channel Coding Scheme for the Multiple Access Relay Channel*, **IEEE Journal on Selected Areas in Communications**, Vol. 31, No. 8, 1157-1167, August 2013.

- **M. Hernaez**, P.M. Crespo, J. Del Ser *A Flexible Channel Coding Approach for Short-Length Codewords*, **IEEE Communications Letters**, Vol. 16, No. 9, 1508-1511, September 2012.

- I. Ochoa, P. Crespo and **M. Hernaez**, *LDPC Codes for Non-Uniform Memoryless Sources and Unequal Energy Allocation*, **IEEE Communications Letters**, Vol. 14, No. 9, 2010.

- **M. Hernaez**, P. M. Crespo, J. Del Ser, J. Garcia-Frias, *Serially-Concatenated LDGM Codes for Correlated Sources over Gaussian Broadcast Channels*, **IEEE Communications Letters**, Vol 13, No. 10, 788-790, October 2009.

- I. Ochoa, P. Crespo, J. Del Ser and **M. Hernaez**, *Turbo Joint Source-Channel Coding of Non-Uniform Memoryless Sources in the Bandwidth-Limited Regime*, **IEEE Communications Letters**, Vol. 14, No. 4, 2010.

Conference Papers

- C. Alberti, N. Daniels, **M. Hernaez**, J. Voges, R. L. Goldfeder, A. A. Hernandez-Lopez, M. Mattavelli, B. Berger, *An Evaluation Framework for Lossy Compression of Genome Sequencing Quality Values*, **Data Compression Conference (DCC)**, 2016 (Accepted).

- I. Ochoa, **M. Hernaez**, R. Goldfeder, T. Weissman and E. Ashley, *Denoising of Quality Scores for Boosted Inference and Reduced Storage*, **Data Compression Conference (DCC)**, 2016.

- **M. Hernaez**, I. Ochoa and T. Weissman, *A cluster-based approach to compression of Quality Scores*, **Data Compression Conference (DCC)**, 2016.

- I. Ochoa, **M. Hernaez** and T. Weissman, *Aligned genomic data compression via improved modeling*, **GIW ISCB-Asia**, Japan, December 2014.

- **M. Hernaez**, G. Vidal, *Communication Services Empowered with a Classical Chaos Based Cryptosystem*, **Financial Cryptography 2013**, Okinawa, Japan 2013

- **M. Hernaez**, P.M. Crespo, J. Del Ser, *A Decode-and-Forward Scheme for Multihop Wireless*

Networks, **IEEE Vehicular Technology Conference (VTC2013-Fall)**, Las Vegas, Sept. 2013

- I. Alustiza, **M. Hernaez**, X. Insasusti and P.M. Crespo, *Teaching Information Theory via a Simulation Tool for Communications Systems*, **IEEE Collaborative Learning & New Pedagogic Approaches in Engineering Education (IEEE EDUCON)**, Berlin (Germany), March 2013.

- **M. Hernaez**, P.M. Crespo, *A novel Scheme for Message-Forwarding in Ad-Hoc Wireless Networks*, **IEEE Vehicular Technology Conference (VTC2011-Spring)**, Budapest (Hungary), May 2011

- **M. Hernaez**, P.M. Crespo, J. del Ser, *Joint Non-Binary LDPC-BICM and Network Coding with Iterative Decoding for the Multiple Access Relay Channel*, **IEEE Vehicular Technology Conference (VTC2011-Spring)**, Budapest (Hungary), May 2011

- I. Ochoa, P. Crespo, J. Del Ser and **M. Hernaez**, *Turbo Joint Source-Channel Coding of Cycle-Stationary Sources in the Bandwidth-Limited Regime*, The 2nd International Conference on Mobile Lightweight Wireless Systems (MOBILIGHT), Spain, May 2010.

Scholarships and Awards

Awarded an **NIH grant**, under the BD2K initiative, in collaboration with the University of Illinois at Urbana-Champaign (UIUC) *2015*

Postdoctoral research funded by the **Stanford Data Science Initiative** *2015-2016*

Enigmedia named the best new company of the Basque Country (Spain) *2013*

University of Navarra Fellowship for graduate studies *2009–2011*

Master Thesis funded by **Telefonica Fellowship.** *2008*

US Patents

- I. Ochoa and **M. Hernaez**, *A Universal Compressor for Genomic Re-Sequencing Data*, Provisional US patent filled by Stanford's OTL - The Office of Technology Licensing, June 2014.

Service Activities

Workshop Organization Committees:

- Chair of special session on "Omics Data Compression and Storage: Present and Future" at ISMB (International Society for Computational Biology), Chicago, 2018 (acceptance rate 20%).
- Chair of special session on "Bioinformatics" at the 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton), October, 2018.
- Chair of special session on "Bioinformatics" at the 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), October, 2017.

Professional Organizations:

- International Society of Computational Biology (ISCB): Member
- Stanford Compression Forum: Organizer of the first and second edition (2015 - 2016)
- International Organization for Standardization (ISO): Active participant in the initiative to define and establish a compression standard for genomic data (under the MPEG working group).
- Stanford Data Science Initiative (SDSI): Active member and grantee (2014 - 2016)

- Center for Science of Information (CSoI), NSF Science and Technology Center: Active member and grantee (2013-2015)

Additional Information

Reviewer: Bioinformatics, Nature Technical Reports, Nature Biotechnology, BMC Bioinformatics, IEEE Communications Letters, several conference proceedings.

Languages: Native: Spanish, Proficiency: English, Low-Intermediate: German, French and Basque.

Computer skills: Programming Languages: C/C++, Python, Applications: R, MatLab, LATEX, MS Office, CVX, Java Operating Systems: Linux, UNIX, Windows.

Student member: Institute of Electrical and Electronics Engineers (IEEE)