Rodrigo Cánovas

Computer Scientist and Bioinformatician with eight years experience. Experience working and researching on succinct data strutures, data compression, pattern finding, multi-alignment, sequencing, error correction, single-nucleotide polymorphism detection, amongst others. Proven abilities to work in interdisciplinary areas of research and quickly learn, apply and adjust to new research projects and environments around the world. Currently, working in the area of genomic risk prediction of autoimmune diseases applying mathematical learning models.

Experience

2018–Current **Postdoctoral Researcher**, Cambridge Baker Systems Genomics Initiative at the Baker Heart and Diabetes Institute.
The work consists in research applications of genomic risk prediction to common disease, particularly autoimmunity.

2015–2017 **Postdoctoral Researcher**, Lab. of Computer Science, Robotics and Microelectronics of Montpellier (LIRMM) and The Institute of Computational Biology (IBC).
The research work consisted of conceiving and developing novel algorithms for analysing sequence data coming from High Throughput Sequencing technologies. This included exploring and improving existing tools while at the same time proposing alternative methods and ideas.

2010–2011 **Research Assistant**, University of Leuven-KU Leuven.
During this short research stay we developed a tool in C++ that offers a new solution to the problem of identifying DNA patterns in a given set of DNA strings - better know as the *motif finding problem*.

2006–2010 **Tutor and Lab Demonstrator**, *Introduction to Hardware*, University of Chile.
This course was focused on the design of digital circuits, and the architecture of modern computers. I was in charge of teaching students how to design, plan, and construct circuits which were tested later during experimental sessions, with the purpose of learning and understanding the hardware elements of a computer.

2008 **Software Development in C#**, EVERIS, Santiago/Chile.
Everis is a multinational consulting firm providing business and strategy solutions, application development, maintenance, and outsourcing services. I was part of a team in charge of developing a program to collect and analyze information of the company. My main role was to implement the algorithms needed to compute the statistical data.

2007 **Tutor and Lab Demonstrator**, *Fundamentals of Computer Science*, University of Chile.
This course was an introduction to the topic of computational complexity including finite automata, regular expressions, context-free grammars, Turing machines, etc. Students were required to complete assignments to explore the efficiency and differences of various computational problems.

Education

2011–2015 **Ph.D.** - **Engineering**, The University of Melbourne.

Thesis Title *Practical Compression for Multi-Alignment Genomic Files*

Supervisors Professor Alistair Moffat & Professor Andrew Turpin

Description This work explored how succinct data structures can be use to compress genomic files, proving the feasibility of random access to the compressed genetic data without requiring full decompression of large data files. The algorithms explored were implement on a Linux platform using C++, and are distributed under an open source license.

Source Code `https://github.com/rcanovas/libCSAM`

40 Lewisham Road – Windsor, VIC 3181 – Australia
(+61) 0476 633 439 • canovas.ba@gmail.com
https://rcanovas.github.io/

2008–2010 **MSc - Computer Science**, University of Chile.

Thesis Title *Compressed Data Structures for Suffix Trees*

Supervisors Professor Gonzalo Navarro

Description In this thesis we presented a new practical compressed suffix tree implementation, which supports various operations over the compressed stored tree without the need of fully decompressing the stored data. This work opened the door to a number of practical suffix tree applications, particularly relevant to Bioinformatics.

Source Code `https://github.com/rcanovas/CST-CN`

2007–2010 **Engineering in Computer Science**, University of Chile.

2002–2007 **Bachelor of Engineering (Computer Science)**, University of Chile.

Personal Qualities

- Ability to perform independent research
- Ability to work effectively within a team
- Ability to rapidly learn and apply new knowledge
- Commitment to interdisciplinary research
- Proven excellent interpersonal skills

Interests

- Bioinformatics
- Genomic Risk Prediction
- Machine Learning
- Genomic Data Analysis
- Design and Analysis of Algorithms
- Data Structure and Compression

Conferences, Seminars, Workshops and Research Stays

2019 Poster Presentation at the 14th GeneMappers Conference, "Genomic Risk Scores for Predicting Juvenile Idiopathic Arthritis", Manly, NSW, Australia

2019 Speaker at The Australian Polygenic Risk Symposium, "Genomic Risk Scores for Predicting Juvenile Idiopathic Arthritis", Sydney, Australia

2017 Poster Presentation at the 2017 conference on Intelligent Systems for Molecular Biology and the European Conference on Computational Biology, "Developing Computational Descriptors to Analyse and Classify Glycosylphosphatidylinositol (GPI) Proteins in Apicomplexan", ISMB/ECCB, Prague.

2017 Organizer of 2017 Annual Research Symposium of the Institut de Biologie Computationnelle

2016 Seminar on Computation over Compressed Structured Data, Dagstuhl, Germany.

2016 Summer school on Bioinformatics Data Structure, Helsinky, Finland.

2016 Workshop on Data Structures in Bioinformatics, Bielfeld, Germany.

2016 Data Compression Conference, DCC 2016, Cliff Lodge, Snowbird, UT, USA.

2014 Poster Presentation at the 2014 Computing and Information Systems Doctoral Colloquium Melbourne, Australia.

2013 Poster Presentation at the 2013 Computing and Information Systems Doctoral Colloquium Melbourne, Australia.

2013 Thirty-Sixth Australasian Computer Science Conference (ACSC 2013), Adelaide, Australia.

2010 String Processing and Information Retrieval, SPIRE 2010 and the 5th Workshop on Compression, Text, and Algorithms Los Cabos, Mexico.

2010 Experimental Algorithms, SEA 2010 Ischia Island, Napoli, Italy.

2010 Research stay at RMIT, invited by Simmon Puglissi (Melbourne, Australia).

2010 Algorithm Engineering and Experimentation, ALENEX 2010 and the ACM-SIAM Symposium on Discrete Algorithms, Austin, Texas.

2009 Poster Presentation in Workshop ICDB (Institute for Cell Dynamics and Biotechnology), Marbella, Chile.

2009 String Processing and Information Retrieval, SPIRE 2009 and the 4th Workshop on Compression, Text, and Algorithms, Saariselka, Finland.

Publications

Diego Arroyuelo, Rodrigo Cánovas, Gonzalo Navarro, and Rajeev Raman. "LZ78 Compression in Low Main Memory Space", To appear in Proc. SPIRE'17, 2017

Rodrigo Cánovas, and Eric Rivals. "Full Compressed Affix Tree Representations", In Proc. DCC'17, pages 102–111, 2017

Rodrigo Cánovas, Alistair Moffat, and Andrew Turpin. "CSAM: Compressed SAM Format", Bioinformatics, btw543, 2016.

Bastien Cazaux, Rodrigo Cánovas, and Eric Rivals. "Shortest DNA Cyclic Cover in Compressed Space", In Proc. DCC'16, pages 536–545, 2016.

Miguel Martínez-Prieto, Nieves Brisaboa, Rodrigo Cánovas, Francisco Claude, Gonzalo Navarro. "Practical compressed string dictionaries", Information Systems 56, 73–108, 2016.

Rodrigo Cánovas, Alistair Moffat, and Andrew Turpin. "Lossy compression of quality scores in genomic data", Bioinformatics, btu183, 2014.

Andres Abeliuk, Rodrigo Cánovas, and Gonzalo Navarro. "Practical Compressed Suffix Trees", Algorithms 6(2):319–351, 2013.

Rodrigo Cánovas and Alistair Moffat. "Practical Compression for Multi-Alignment Genomic Files", In Proc. 36th ACSC, pages 51–60, 2013.

Miguel Martínez-Prieto, Javier Fernández, and Rodrigo Cánovas. "Compression of RDF dictionaries", SAC 2012: 340–347, 2012.

Miguel Martínez-Prieto, Javier Fernández, and Rodrigo Cánovas. "Querying RDF dictionaries in compressed space", ACM SIGAPP Applied Computing Review 12(2):64–77, 2012.

Nieves Brisaboa, Rodrigo Cánovas, Francisco Claude, Miguel Martínez-Prieto, and Gonzalo Navarro. "Compressed String Dictionaries", In Proc. SEA'11, pages 136–147. LNCS 6630, 2011.

Rodrigo Cánovas and Gonzalo Navarro. "Practical Compressed Suffix Trees", In Proc. SEA'10, pages 94–105. LNCS 6049, 2010.

Diego Arroyuelo, Rodrigo Cánovas, Gonzalo Navarro, and Kunihiko Sadakane. "Succinct Trees in Practice", In Proc. ALENEX'10, pages 84–97, 2010.

Computer skills

Basic	Prolog, PHP, MySql, JavaScript
Intermediate	Perl, Java, Python, Linux (user/administrator level), MacOS and Windows (user level)
Advanced	C/C++/C++11, R, LaTeX

Languages

Spanish	**Native**
English	**Advanced**

40 Lewisham Road – Windsor, VIC 3181 – Australia
(+61) 0476 633 439 • canovas.ba@gmail.com
https://rcanovas.github.io/