

# LUIS PEDRO COELHO

## Curriculum Vitæ

September 25, 2021

Institute of Science and Technology  
for Brain-inspired Intelligence (ISTBI)  
Fudan University, Shanghai, China  
email : [luis@luispedro.org](mailto:luis@luispedro.org)  
ORCID : [0000-0002-9280-7885](https://orcid.org/0000-0002-9280-7885)  
Erdős-Bacon Nr. : 7  
Citizenship : U.K. and Portugal

## Education

**2011** PhD in computational biology, Carnegie Mellon University

Dissertation topic: *Modeling the Space of Subcellular Location Patterns Using Images and Other Sources of Information*, advised by Prof. Robert F. Murphy.

**2006** MSc in computer science, Instituto Superior Técnico (Technical University Lisbon)

Dissertation topic: *Bayesian Network Parameter Estimation Using Noisy Observations or Soft Evidence*, advised by Prof. Arlindo Oliveira.

**2004** BSc in computer science, Instituto Superior Técnico (Technical University Lisbon)

Graduated top of my class.

## Professional Experience

**2018-present** Junior Principal Investigator (Fudan University)

**2013-2018** Postdoctoral researcher at European Molecular Biology Laboratory (EMBL)

Supervisor: Dr. Peer Bork

**2012** Postdoctoral researcher at Instituto de Medicina Molecular (Lisbon)

Supervisor: Dr. Musa Mhlanga

## Scholarships & Awards

**2013** JBS Authors' Choice Award

For most popular article published in The Journal of Biomolecular Screening in 2012.

**2012** Siebel Scholar

Awarded annually for academic excellence and demonstrated leadership to 85 top students from the world's leading graduate schools

**2007-2011** PhD. Scholarship from Portuguese Science Foundation

**2009** Joint CMU-U. of Pittsburgh PhD. in Computational Biology Research Excellence Award

**2008** Joint CMU-U. of Pittsburgh PhD. in Computational Biology Academic Excellence Award

**2006** Fulbright Fellow

**2005** Scholarship from Portuguese Science Foundation

**2001** Instituto Superior Técnico (IST) Academic Excellence Award

## Highlighted Publications

**2018** L. P. Coelho et al., *Similarity of the dog and human gut microbiomes in gene content and response to diet* in MICROBIOME (2018) [\[DOI\]](#)

**2015** S. Sunagawa\*, L. P. Coelho\*, S. Chaffron\* et al., *Structure and Function of the Global Ocean Microbiome* in SCIENCE [\[DOI\]](#)

**2013** L. P. Coelho\*, T. Peng\* et al., *Determining the subcellular location of new proteins from microscope images using local features* in BIOINFORMATICS [\[DOI\]](#)

## Service & Outreach

**2020–present** Member of the ISCB (International Society for Computational Biology) Equity, Diversity and Inclusion (EDI) Committee

**2016–present** Associate Editor for the Journal of Open Research Software

**2014–2017** Postdoc representative at EMBL

I helped organized the 2015 and 2017 EMBL Postdoc retreats.

**2014–2015** Organized Software Carpentry Workshop at EMBL

**2012–2013** Member of the Organization of the Lisbon Machine Learning School

**2010** Local Committee for Portuguese-American Postgraduate Society National Forum

I headed the local organising committee for the 2010 edition of this annual event. It took place in Pittsburgh and included, as speakers, cabinet-level Portuguese government officials, renowned researchers, artists, as well as participants from all around the US.

**2006–present** Peer Review

I have reviewed papers for BMC Bioinformatics, PLoS Computational Biology, Database, IEEE Transactions in Computational Biology and Bioinformatics, and IEEE Transactions on Medical Imaging. I have also reviewed grant proposals for the Czech Science Foundation and the French National Research Agency (ANR).

**2002–2004** Producer for IST Theatre Group

## Teaching & Mentoring

**2019–present** Co-teaches *Scientific Communication* at Fudan University, a one-semester graduate course

**2012–present** Certified Software Carpentry Instructor

I have given lectures on Software Carpentry in Germany, Denmark, Cyprus, Jordan, and Spain.

**2012–2013** Supervision of Students at Instituto Superior Técnico (IST, Lisbon)

Co-advised one Master thesis in Computer Science (Paulo Monteiro) and one in Biomedical Engineering (Joana Cruz). During 2012, I was co-mentor to bioinformaticians in Luisa M. Figueiredo's lab as part of a collaboration.

**2008–2011** Mentoring Junior Members of Murphy Lab

During my PhD. studies, I had the opportunity to directly supervise and mentor several junior members of the Murphy Lab. This experience includes working with **paid undergraduate programmers** (Nathan Herzing and Jephthah Liddie one **MSc. student** (Shannon Quinn), **undergraduate students** performing lab work for credit (Jimmy Xu), and **high school students** volunteering over the summer (Peter Webb and Robert Webb).

**2012** Programming for Scientists (short course)

**2012** Practical Tutorial in Next Generation Sequencing in “Omics for Parasite Biology” Course

**2011–2015** Tutorial in Lisbon Machine Learning Summer School

**2009** Programming for Scientists

I designed and taught a semester-long course on computer programming for scientists at Carnegie Mellon University.

**2008** Teaching Assistant for Laboratory Methods for Computational Biologists (CMU)

**2005** Introduction to Computers Course in Cacém

I co-developed and co-taught an introductory course in computer usage in Cacém, an underprivileged neighborhood near Lisbon. This was pro-bono work.

**2005** Teaching Assistant for Decision Support Systems (IST)

## Language Skills

I am bilingual in **English** and **Portuguese**. I speak and write fluent **German** (I attended a German high-school, obtaining an Abitur; and later spent a year as an exchange student at Technical University of Vienna) and speak fluent **French**. I know basic **Luxembourgish**.

## Publications

Google scholar profile: <http://scholar.google.com/citations?user=qTYua0cAAAAJ>

## First or corresponding author

(includes co-first/co-corresponding author publications.)

### Preprints submitted for publication

1. Shaojun Pan, Chengkai Zhu, Xing-Ming Zhao<sup>+</sup>, and **Luis Pedro Coelho**<sup>+</sup> *SemiBin: Incorporating information from reference genomes with semi-supervised deep learning leads to better metagenomic assembled genomes (MAGs)* in *BroRxiv* 2021 [\[DOI\]](#)
2. Senying Lai, Shaojun Pan, **Luis Pedro Coelho**<sup>+</sup>, Wei-Hua Chen,<sup>+</sup> Xing-Ming Zhao<sup>+</sup>, *metaMIC: reference-free Misassembly Identification and Correction of de novo metagenomic assemblies* at *BioRxiv* 2021 [\[DOI\]](#)

### Peer-reviewed journal publications

3. Célio Dias Santos-Junior, Shaojun Pan, Xing-Ming Zhao, **Luis Pedro Coelho**, *MACREL: antimicrobial peptide screening in genomes and metagenomes* at *PEERJ* 2020 [\[DOI\]](#)
4. **Luis Pedro Coelho**, Renato Alves, Paulo Monteiro, Jaime Huerta-Cepas, Ana Teresa Freitas, Peer Bork *NG-meta-profiler: fast processing of metagenomes using NGLess, a domain-specific language* in *MICROBIOME* vol. 7, 84, 2019 [\[DOI\]](#)  
**My contribution:** *I conceived of the concept, co-wrote the software and the manuscript*
5. **Luis Pedro Coelho**, Jens Kultima, Paul Costea, Coralie Fournier, Yuanlong Pan, Gail Czarnecki-Maulden, Matthew Hayward, Kristoffer Forslund, Patrick Descombes, Janet Jackson, Qinghong Li, and Peer Bork *Similarity of the dog and human gut microbiomes in gene content and response to diet* in *MICROBIOME*, vol 6:72, 2018 [\[DOI\]](#)  
**My contribution:** *I designed and implemented the analysis strategy, wrote the first version of the manuscript, and lead the subsequent incorporation of co-author suggestions.*
6. **Luis Pedro Coelho**, *Jug: Software for parallel reproducible computation in Python* in *JOURNAL OF OPEN RESEARCH SOFTWARE*, 2017 [\[DOI\]](#)  
**My contribution:** *I designed and implemented the software presented and wrote the manuscript.*
7. Sebastien Colin<sup>\*</sup>, **Luis Pedro Coelho**<sup>\*</sup>, Shinichi Sunagawa, Chris Bowler, Eric Karsenti, Peer Bork, Rainer Pepperkok, Colomban de Vargas, *Quantitative 3D-Imaging of Environmental Microbial Eukaryotes using Automated Fluorescence Microscopy* in *eLIFE*, vol. 6:e26066, 2017 [\[DOI\]](#).  
**My contribution:** *I designed and implemented the computational analysis necessary for the method presented in the paper and participated in the writing of the manuscript.*
8. Shinichi Sunagawa<sup>\*</sup>, **Luis Pedro Coelho**<sup>\*</sup>, Samuel Chaffron<sup>\*</sup>, Jens Roat Kultima, Karine Labadie, Guillem Salazar, Bardya Djahanschiri, Georg Zeller, Daniel R. Mende, Adriana Alberti, Francisco M. Cornejo-Castillo, Paul I. Costea, Corinne Cruaud, Francesco d'Ovidio, Stefan Engelen, Isabel Ferrera, Josep M. Gasol, Lionel Guidi, Falk Hildebrand, Florian Kokoszka, Cyrille Lepoivre, Gipsi Lima-Mendez, Julie Poulain, Bonnie T. Poulos, Marta Royo-Llonch, Hugo Sarmento, Sara Vieira-Silva, Céline Dimier, Marc Picheral, Sarah Searson, Stefanie Kandels-Lewis, Tara Oceans coordinators, Chris Bowler, Colomban de Vargas, Gabriel Gorsky, Nigel Grimsley, Pascal Hingamp, Daniele Iudicone, Olivier Jaillon, Fabrice Not, Hiroyuki Ogata, Stephane Pesant, Sabrina Speich, Lars Stemmann, Matthew B. Sullivan, Jean Weissenbach, Patrick Wincker, Eric Karsenti, Jeroen Raes, Silvia G. Acinas, Peer Bork, *Structure and Function of the Global Ocean Microbiome* in *SCIENCE* 348 (6237), 1261359, 2015 [\[DOI\]](#)  
**My contribution:** *As co-first author, I performed analysis of the data, including linking the metagenomics data*

to the environmental parameters, as well as the comparison to the human gut (this required extending existing software tools which were unable to scale to this dataset); and helped draft the manuscript.

9. **Luis Pedro Coelho**, Catarina Pato, Ana Friães, Ariane Neumann, Maren von Kückritz-Blickwede, Mário Ramirez, João André Carriço, *Automatic determination of NET (neutrophil extracellular traps) coverage in fluorescent microscopy images* in *BIOINFORMATICS* 31 (14): 2364–2370, 2015 [\[DOI\]](#)

**My contribution:** *I conceived and implemented the algorithm, and wrote the manuscript.*

10. **Luis Pedro Coelho**, Joshua D. Kangas, Armaghan Naik, Elvira Osuna-Highley, Estelle Glory-Afshar, Margaret Fuhrman, Ramanuja Simha, Peter B. Berget, Jonathan W. Jarvik, and Robert F. Murphy, *Determining the sub-cellular location of new proteins from microscope images using local features* in *BIOINFORMATICS*, 2013 [\[DOI\]](#)

**My contribution:** *I designed and implemented the proposed algorithm, collected microscopy data (in collaboration with other authors), and wrote the manuscript.*

11. **Luis Pedro Coelho** Mahotas: Open source software for scriptable computer vision, *JOURNAL OF OPEN RESEARCH SOFTWARE*, vol. 1, 2013 [\[DOI\]](#)

**My contribution:** *I implemented the underlying computer vision software and the manuscript.*

12. **Luis Pedro Coelho\***, Tao Peng\*, and Robert F. Murphy, *Quantifying the distribution of probes between subcellular locations using unsupervised pattern unmixing* in *BIOINFORMATICS*, vol. 26 (12), pp. i7–i12, 2010 [\[DOI\]](#)

**My contribution:** *I conceived of and implemented one of the methods presented and wrote the manuscript.*

13. **Luis Pedro Coelho**, Amr Ahmed, Andrew Arnold, Joshua Kangas, Abdul-Saboor Sheikh, Eric P. Xing, William W. Cohen, and Robert F. Murphy, *Structured Literature Image Finder: Extracting Information from Text and Images in Biomedical Literature* in *LECTURE NOTES IN BIOINFORMATICS*, vol. 6004, pp. 23–32, 2010 [\[DOI\]](#)

**My contribution:** *With the paper below, this forms a package on the SLIF project, where I enhanced the image analysis pipeline (for the analysis of figures in published scientific literature). I wrote the first draft of this manuscript and helped draft the one below.*

## Peer-reviewed Conference Proceedings

14. **Luis Pedro Coelho**, Aabid Shariff, and Robert F. Murphy; *Nuclear segmentation in microscope cell images: A hand-segmented dataset and comparison of algorithms* in *PROCEEDINGS OF IEEE INTERNATIONAL SYMPOSIUM IN BIOMEDICAL IMAGING*, 2009 [\[DOI\]](#)

**My contribution:** *I acquired (part of) the microscopy data, built the hand-segmented dataset, implemented the methods and wrote the first draft of the manuscript.*

15. **Luis Pedro Coelho** and Robert Murphy; *Identifying Subcellular Locations from Images of Unknown Resolution* in *BIOINFORMATICS RESEARCH AND DEVELOPMENT, Communications in Computer and Information Science*, vol. 13, pp. 235–242, 2008 [\[DOI\]](#)

**My contribution:** *I implemented the algorithm, ran the tests, and wrote the first draft of the paper.*

16. **Luis Pedro Coelho** and Arlindo Oliveira; *Dotted Suffix Trees: A Structure for Approximate Text Indexing* in *STRING PROCESSING AND INFORMATION RETRIEVAL, Lecture Notes in Computer Science*, vol. 4209, pp. 329–336, 2006 [\[DOI\]](#)

**My contribution:** *I developed and implemented the algorithm and wrote the first draft of the manuscript.*

## Comments and review articles

17. **Luis Pedro Coelho** *Voices of the new generation: science in a state of benign confusion* in *NATURE REVIEWS MOLECULAR CELL BIOLOGY* 2020 [\[DOI\]](#)

18. **Luis Pedro Coelho**, Estelle Glory-Afshar, Joshua Kangas, Shannon Quinn, Aabid Shariff, and Robert F. Murphy; *Principles of Bioimage Informatics: Focus on machine learning of cell patterns* in *LINKING LITERATURE, INFORMATION, AND KNOWLEDGE FOR BIOLOGY, Lecture Notes in Computer Science*, vol. 6004, pp. 8–18, 2010 [\[DOI\]](#)

## Co-authorships

### Peer-reviewed journal publications

19. Askarbek Orakov\*, Anthony Fullam\*, **Luis Pedro Coelho**, Supriya Khedkar, Damian Szklarczyk, Daniel R Mende, Thomas SB Schmidt† and Peer Bork† *GUNC: Detection of Chimerism and Contamination in Prokaryotic Genomes* at GENOME BIOLOGY 2021 [\[DOI\]](#)
20. Mohammad Bahram, Tarquin Netherway, Clémence Frioux, Pamela Ferretti, **Luis Pedro Coelho**, Stefan Geisen, Peer Bork, and Falk Hildebrand *Metagenomic assessment of the global diversity and distribution of bacteria and fungi* at ENVIRONMENTAL MICROBIOLOGY 2020 [\[DOI\]](#)
21. Sara Vieira-Silva, Gwen Falony, Eugeni Belda, Trine Nielsen, Judith Aron-Wisnewsky, Rima Chakaroun, Sofia K. Forslund, Karen Assmann, Mireia Valles-Colomer, Thi Thuy Duyen Nguyen, Sebastian Proost, Edi Prifti, Valentina Tremaroli, Nicolas Pons, Emmanuelle Le Chatelier, Fabrizio Andreelli, Jean-Phillippe Bastard, **Luis Pedro Coelho**, Nathalie Galleron, Tue H. Hansen, Jean-Sébastien Hulut, Christian Lewinter, Helle K. Pedersen, Benoit Quinquis, Christine Rouault, Hugo Roume, Joe-Elie Salem, Nadja B. Søndertoft, Sothea Touch, MetaCardis Consortium, Marc-Emmanuel Dumas, Stanislav Dusko Ehrlich, Pilar Galan, Jens P. Gøtze, Torben Hansen, Jens J. Holst, Lars Køber, Ivica Letunic, Jens Nielsen, Jean-Michel Oppert, Michael Stumvoll, Henrik Vestergaard, Jean-Daniel Zucker, Peer Bork, Oluf Pedersen, Fredrik Bäckhed, Karine Clément and Jeroen Raes *Statin therapy is associated with lower prevalence of gut microbiota dysbiosis* in NATURE 581, 310–315 (2020) [\[DOI\]](#)
22. Oleksandr M. Maistrenko\*, Daniel R. Mende\*, Mechthild Luetge, Falk Hildebrand, Thomas S. B. Schmidt, Simone S. Li, **Luis Pedro Coelho**, Jaime Huerta-Cepas, Shinichi Sunagawa, Peer Bork *Disentangling the impact of environmental and phylogenetic constraints on prokaryotic strain diversity* in ISME JOURNAL 14, 1247–1259 2020 [\[DOI\]](#)
23. Daniel R Mende, Ivica Letunic, Oleksandr M Maistrenko, Thomas S B Schmidt, Alessio Milanese, Lucas Paoli, Ana Hernández-Plaza, Askarbek N Orakov, Sofia K Forslund, Shinichi Sunagawa, Georg Zeller, Jaime Huerta-Cepas, **Luis Pedro Coelho**, Peer Bork *proGenomes2: an improved database for accurate and consistent habitat, taxonomic and functional annotations of prokaryotic genomes* in NUCLEIC ACIDS RESEARCH 48:D1 D621-D625, 2020 [\[DOI\]](#)
24. Federico M. Ibarbalz, Nicolas Henry, Manoela C. Brandão, Séverine Martini, Greta Busseni, Hannah Byrne, **Luis Pedro Coelho**, Hisashi Endo, Josep M. Gasol, Ann C. Gregory, Frédéric Mahé, Janaina Rigonato, Marta Royo-Llonch, Guillem Salazar, Isabel Sanz-Sáez, Eleonora Scalco, Dodji Soviadan, Ahmed A. Zayed, Adriana Zingone, Karine Labadie, Joannie Ferland, Claudie Marec, Stefanie Kandels, Marc Picheral, Céline Dimier, Julie Poulain, Sergey Pisarev, Margaux Carmichael, Stéphane Pesant, Tara Oceans Coordinators, Marcel Babin, Emmanuel Boss, Daniele Iudicone, Olivier Jaillon, Silvia G. Acinas, Hiroyuki Ogata, Eric Pelletier, Lars Stemmann, Matthew B. Sullivan, Shinichi Sunagawa, Laurent Bopp, Colomban de Vargas, Lee Karp-Boss, Patrick Wincker, Fabien Lombard, Chris Bowler#, and Lucie Zinger# *Global Trends in Marine Plankton Diversity across Kingdoms of Life* in CELL, 2019 [\[DOI\]](#)
25. Guillem Salazar\*, Lucas Paoli\*, Adriana Alberti, Jaime Huerta-Cepas, Hans-Joachim Ruscheweyh, Miguelangel Cuenca, Christopher M. Field, **Luis Pedro Coelho**, Corinne Cruaud, Stefan Engelen, Ann C. Gregory, Karine Labadie, Claudie Marec, Eric Pelletier, Marta Royo-Llonch, Simon Roux, Pablo Sánchez, Hideya Uehara, Ahmed A. Zayed, Georg Zeller, Margaux Carmichael, Céline Dimier, Joannie Ferland, Stefanie Kandels, Marc Picheral, Sergey Pisarev, Julie Poulain, Tara Oceans Coordinators, Silvia G. Acinas, Marcel Babin, Peer Bork, Chris Bowler, Colomban de Vargas, Lionel Guidi, Pascal Hingamp, Daniele Iudicone, Lee Karp-Boss, Eric Karsenti, Hiroyuki Ogata, Stéphane Pesant, Sabrina Speich, Matthew B. Sullivan, , Patrick Wincker, and Shinichi Sunagawa, *Gene Expression Changes and Community Turnover Differentially Shape the Global Ocean Metatranscriptome* in CELL, 2019 [\[DOI\]](#)

26. Jakob Wirbel\*, Paul Theodor Pyl\*, Ece Kartal, Konrad Zych, Alireza Kashani, Alessio Milanese, Jonas S Fleck, Anita Y Voigt, Albert Palleja, Ruby P Ponnudurai, Shinichi Sunagawa, **Luis Pedro Coelho**, Petra Schrotz-King, Emily Vogtmann, Nina Habermann, Emma Niméus, Andrew M Thomas, Paolo Manghi, Sara Gandini, Davide Serrano, Sayaka Mizutani, Hirotugu Shiroma, Satoshi Shiba, Tatsuhiro Shibata, Shinichi Yachida, Takuji Yamada, Levi Waldron, Alessio Naccarati, Nicola Segata, Rashmi Sinha, Cornelia M. Ulrich, Hermann Brenner, Manimozhiyan Arumugam, Peer Bork, Georg Zeller *Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer* in NATURE MEDICINE vol. 25, pp. 679–689 (2019)
27. Alessio Milanese\*, Daniel Mende\*, Lucas Paoli, Guillem Salazar, Hans-Joachim Ruscheweyh, Miguelangel Cuenca, Pascal Hingamp, Renato Alves, Paul Costea, **Luis Pedro Coelho**, Thomas Schmidt, Alexandre Almeida, Alex Mitchell, Robert Finn, Jaime Huerta-Cepas, Peer Bork, Georg Zeller, and Shinichi Sunagawa *Microbial abundance, activity and population genomic profiling with mOTUs2* in NATURE COMMUNICATIONS, vol. 10, pp. 1014, 2019 [\[DOI\]](#)
28. Thomas SB Schmidt\*, Matthew R Hayward\*, **Luis Pedro Coelho**, Simone S Li, Paul I Costea, Anita Y Voigt, Jakob Wirbel, Oleksandr M Maistrenko, Renato J Alves, Emma Bergsten, Carine de Beaufort, Iradj Sobhani, Anna Heintz-Buschart, Shinichi Sunagawa, Georg Zeller, Paul Wilmes, Peer Bork *Extensive Transmission of Microbes along the Gastrointestinal Tract* in eLIFE, vol. 8, pp. e42693, 2019 [\[DOI\]](#)
29. Luigi Caputi\*, Quentin Carradec\*, Damien Eveillard\*, Amos Kirilovsky\*, Eric Pelletier\*, Fabio Rocha Jimenez Vieira, Emilie Villar, Shruti Malviya, Eleonora Scalco, Silvia G. Acinas, Adriana Alberti, Jean-Marc Aury, Alexis Bertrand, Tristan Biard, Lucie Bittner, Martine Boccara, Jennifer R. Brum, Christophe Brunet, Anna Carratalà, Samuel Chaffron, Hervé Claustre, **Luis Pedro Coelho**, Sébastien Colin, Salvatore D’Aniello, Corinne Da Silva, Marianna Del Core, Hugo Doré, Stéphane Gasparini, Florian Kokoszka, Jean-Louis Jamet, Christophe Lejeusne, Cyrille Lepoivre, Magali Lescot, Gipsi Lima-Mendez, Fabien Lombard, Julius Lukeš, Nicolas Maillet, Elodie Martinez, MariaGrazia Mazzocchi, Mario B. Néou, Javier Paz-Yepes, Julie Poulain, Simon Ramondenc, Jean-Baptiste Romagnan, Simon Roux, Daniela Salvaggio Manta, Remo Sanges, Sabrina Speich, Mario Sprovieri, Shinichi Sunagawa, Vincent Taillander, Atsuko Tanaka, Leila Tirichine, Camille Trottier, Julia Uitz, Jana Veselá, Flora Vincent, Sheree Yau, Stefanie Kandels-Lewis, Sarah Searson, Céline Dimier, Marc Picheral, Tara Oceans Coordinators, Emmanuel Boss, Colomban de Vargas, Mick Follows, Nigel Grimsley, Lionel Guidi, Pascal Hingamp, Eric Karsenti, Paolo Sordino, Lars Stemmann, Matthew B. Sullivan, Adriana Zingone, Laurence Garczarek, Fabrizio d’Ortenzio, Pierre Testor, Fabrice Not, Maurizio Ribera d’Alcalà, Patrick Wincker, Chris Bowler, Daniele Iudicone, *Community-level response to natural perturbations in open ocean planktonic ecosystems* in GLOBAL BIOGEOCHEMICAL CYCLES, vol. 33, 2019 [\[DOI\]](#)  
**My contribution:** *I contributed to data analysis (of metatranscriptomics and imaging data).*
30. Albert Palleja\*, Kristian H. Mikkelsen\*, Sofia K. Forslund\*, Alireza Kashani, Kristine H. Allin, Trine Nielsen, Tue H. Hansen, Suisha Liang, Qiang Feng, Chenchen Zhang, Paul Theodor Pyl, **Luis Pedro Coelho**, Huanming Yang, Jian Wang, Morten F. Nielsen, Henrik Bjorn Nielsen, Peer Bork, Jun Wang, Tina Vilsbøll, Torben Hansen, Filip K. Knop\*, Manimozhiyan Arumugam\*, Oluf Pedersen\* *Gut resistome modulates resilience and recovery of gut microbiota following broad-spectrum antibiotic treatment* in NATURE MICROBIOLOGY vol. 3(11) pp. 1255–1265, 2018 [\[DOI\]](#)  
**My contribution:** *I performed statistical analyses of the microbiome data.*
31. Mohammad Bahram\*, Falk Hildebrand\*, Kristoffer Forslund, Jennifer L. Anderson, Nadejda A. Soudzilovskaia, Peter M. Bodegom, Sten Anslan, Johan Bengtsson-Palme, **Luis Pedro Coelho**, Helery Harend, Mia R. Maltz, Sunil Mundra, Pål Axel Olsson, Mari Pent, Sergei Pölme, Shinichi Sunagawa, Martin Ryberg, Leho Teder-soo, and Peer Bork *Towards the structure and function of the global topsoil microbiome* in NATURE, vol. 560, pp. 233–237, 2018 [\[DOI\]](#)  
**My contribution:** *I contributed to the analysis of metagenomics data, in particular building a gene catalogue of the soil microbiome.*
32. Katri Korpela, Paul Costea, **Luis Pedro Coelho**, Stephanie Kandels-Lewis, Gonneke Willemsen, Dorret I Boomsma, Nicola Segata, Peer Bork *Selective maternal seeding and environment shape the human gut microbiome* in



GENOME RESEARCH vol. 28(4), pp. 561–568, 2018 [\[DOI\]](#)

**My contribution:** *I contributed to the analysis and interpretation of the metagenomics data.*

33. Paul I. Costea, **Luis Pedro Coelho**, Shinichi Sunagawa, Robin Muench, Jaime Huerta-Cepas, Kristoffer Forslund, Falk Hildebrand, Almagul Kushugulova, Georg Zeller, Peer Bork *Subspecies in the global human gut microbiome* in MOLECULAR SYSTEMS BIOLOGY, vol. 13(12), pp. 960, 2017 [\[DOI\]](#)

**My contribution:** *I contributed to the analysis and interpretation of the metagenomics data and discussions of statistical methodology.*

34. Paul I Costea, Georg Zeller, Shinichi Sunagawa, Eric Pelletier, Adriana Alberti, Florence Levenez, Melanie Tramontano, Marja Driessen, Rajna Hercog, Ferris-Elias Jung, Jens Roat Kultima, Matthew R Hayward, **Luis Pedro Coelho**, Emma Allen-Vercoe, Laurie Bertrand, Michael Blaut, Jillian R M Brown, Thomas Carton, Stéphanie Cools-Portier, Michelle Daigneault, Muriel Derrien, Anne Druésne, Willem M de Vos, B Brett Finlay, Harry J Flint, Francisco Guarner, Masahira Hattori, Hans Heilig, Ruth Ann Luna, Johan van Hylckama Vlieg, Jana Junick, Ingeborg Klymiuk, Philippe Langella, Emmanuelle Le Chatelier, Volker Mai, Chaysavanh Manichanh, Jennifer C Martin, Clémentine Mery, Hidetoshi Morita, Paul W O'Toole, Céline Orvain, Kiran Raosaheb Patil, John Penders, Søren Persson, Nicolas Pons, Milena Popova, Anne Salonen, Delphine Saulnier, Karen P Scott, Bhagirath Singh, Kathleen Slezak, Patrick Veiga, James Versalovic, Liping Zhao, Erwin G Zoetendal, S Dusko Ehrlich, Joel Dore, and Peer Bork *Towards standards for human fecal sample processing in metagenomic studies* in NATURE BIOTECHNOLOGY vol. 35(11), pp. 1069–1076, 2017 [\[DOI\]](#).

**My contribution:** *I contributed in the computational analysis of the samples in the benchmark.*

35. Paul I. Costea\*, Robin Muench\*, **Luis Pedro Coelho**, and Peer Bork *metaSNV: a tool for metagenomic strain level analysis* in PLoS One, vol. 12(7), pp. e0182392, 2017 [\[DOI\]](#)

**My contribution:** *I contributed computer code for the proposed tool, ideas for the design of the benchmark, and participated in the manuscript preparation.*

36. Jaime Huerta-Cepas, Kristoffer Forslund, **Luis Pedro Coelho**, Damian Szklarczyk, Lars Juhl Jensen, Christian von Mering, and Peer Bork *Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper* in MOLECULAR BIOLOGY AND EVOLUTION, vol. 34(8), pp. 2115–2122, 2017 [\[DOI\]](#)

**My contribution:** *I contributed to the evaluation of the tool, in particular, in the evaluation of the suitability of the tool to metagenomics data*

37. J. R. Bedarf\*, F. Hildebrand\*, **L. P. Coelho**, S. Sunagawa, M. Bahram, F. Goeser, P. Bork, U. Wüllner *Microbial, functional and viral gut metagenome shifts in early onset L-DOPA naive PD patients* GENOME MEDICINE 9:39 2017 [\[DOI\]](#)

**My contribution:** *I built machine learning models based on metagenomics data.*

38. Jens Roat Kultima, **Luis Pedro Coelho**, Kristoffer Forslund, Jaime Huerta-Cepas, Simone S. Li, Marja Driessen, Anita Yvonne Voigt, Georg Zeller, Shinichi Sunagawa and Peer Bork, *MOCAT2: a metagenomic assembly, annotation and profiling framework* BIOINFORMATICS 32 (16): 2520–2523, 2016 [\[DOI\]](#)

**My contribution:** *I participated in the design and implementation of the proposed tool.*

39. Lionel Guidi\*, Samuel Chaffron\*, Lucie Bittner\*, Damien Eveillard, Abdelhalim Larhlimi, Simon Roux, Youssef Darzi, Stephane Audic, Léo Berline, Jennifer Brum, **Luis Pedro Coelho**, Julio Cesar Ignacio Espinoza, Shruti Malviya, Shinichi Sunagawa, Céline Dimier, Stefanie Kandels-Lewis, Marc Picheral, Julie Poulain, Sarah Searson, Tara Oceans Consortium Coordinators, Lars Stemmann, Fabrice Not, Pascal Hingamp, Sabrina Speich, Mick Follows, Lee Karp-Boss, Emmanuel Boss, Hiroyuki Ogata, Stephane Pesant, Jean Weissenbach, Patrick Wincker, Silvia G. Acinas, Peer Bork, Colomban de Vargas, Daniele Iudicone, Matthew B. Sullivan, Jeroen Raes, Eric Karsenti, Chris Bowler, and Gabriel Gorsky, *Plankton networks driving carbon export in the oligotrophic ocean* in NATURE, vol. 532(7600), pp. 465–470, 2016 [\[DOI\]](#)

**My contribution:** *I performed functional analysis of the metagenomics data.*

40. Renato Pinheiro-Silva, Lara Borges, **Luis Pedro Coelho**, Alejandro Cabezas-Cruz, James J. Valdés, Virgilio do Rosário, José de la Fuente, *Gene expression changes in the salivary glands of Anopheles coluzzii elicited by*

*Plasmodium berghei* infection in PARASITES & VECTORS 8:485, 2015 [DOI]

**My contribution:** *I supervised the first-author, who was, at the time a PhD student, in the statistical analysis of the RNA-seq data.*

41. Ana C. Pena, Mafalda R. Pimentel, Helena Manso, Rita Vaz-Drago, Daniel Pinto-Neves, Francisco Aresta-Branco, Filipa Rijo-Ferreira, Fabien Guegan, **Luis Pedro Coelho**, Maria Carmo-Fonseca, Nuno L. Barbosa-Morais, Luisa M. Figueiredo, *Trypanosoma brucei* histone H1 inhibits RNA polymerase I transcription and is important for parasite fitness in vivo in MOLECULAR MICROBIOLOGY, vol. 93(4), pp645–63, 2014 [DOI]

**My contribution:** *I supervised D.P-N., who was at the time, a bioinformatics MSc student, in the statistical analysis of the RNA-seq data.*

42. Kristoffer Forslund, Shinichi Sunagawa, **Luis P. Coelho**, Peer Bork, *Metagenomic insights into the human gut resistome and the forces that shape it* in BIOESSAYS, vol. 36(3), pp. 316–29, 2014 [DOI]

**My contribution:** *I performed data analysis.*

43. Peter Liehl, Vanessa Zuzarte-Luís, Jennie Chan, Thomas Zillinger, Fernanda Baptista, Daniel Carapau, Madlen Konert, Kirsten K Hanson, Céline Carret, Caroline Lassnig, Mathias Müller, Ulrich Kalinke, Mohsan Saeed, Angelo Ferreira Chora, Douglas T Golenbock, Birgit Strobl, Miguel Prudêncio, **Luis P Coelho**, Stefan H Kappe, Giulio Superti-Furga, Andreas Pichlmair, Ana M Vigário, Charles M Rice, Katherine A Fitzgerald, Winfried Barchet, and Maria M Mota, *Host-cell sensors for Plasmodium activate innate immunity against liver-stage infection* in NATURE MEDICINE vol. 20, pp. 47–53, 2014 [DOI]

**My contribution:** *I performed statistical analysis of the gene expression data.*

44. Shinichi Sunagawa, Daniel R Mende, Georg Zeller, Fernando Izquierdo-Carrasco Simon A Berger, Jens Roat Kultima, **Luis Pedro Coelho**, Manimozhiyan Arumugam, Julien Tap, Henrik Bjørn Nielsen, Simon Rasmussen, Søren Brunak, Oluf Pedersen, Francisco Guarner, Willem M de Vos, Jun Wang, Junhua Li, Joël Doré, S Dusko Ehrlich, Alexandros Stamatakis and Peer Bork, *Metagenomic species profiling using universal phylogenetic marker genes* in NATURE METHODS, 2013 [DOI]

**My contribution:** *I implemented the stand-alone version of the published metagenomics tool.*

45. Amr Ahmed, Andrew Arnold, **Luis Pedro Coelho**, Joshua Kangas, Abdul-Saboor Sheikk, Eric P. Xing, William W. Cohen, *Structured Literature Image Finder: Parsing Text and Figures in Biomedical Literature* in WEB SEMANTICS: SCIENCE, SERVICES AND AGENTS ON THE WORLD WIDE WEB, vol. 8, pp. 151–154, 2010 [DOI]

## Peer-reviewed Conference Proceedings

46. Taraz Buck, Arvind Rao, **Luis Pedro Coelho**, Margaret Fuhrman, Jonathan W. Jarvik, Peter B. Berget, and Robert F. Murphy; *Cell Cycle Dependence of Protein Subcellular Location Inferred from Static, Asynchronous Images* in CONFERENCE PROCEEDINGS OF THE IEEE ENGINEERING IN MEDICAL AND BIOLOGY SOCIETY, pp. 1016–1019, 2009 [DOI]

**My contribution:** *I helped the first author acquire relevant microscopy data.*

47. Amina Chebira, **Luis Pedro Coelho**, Aliaksei Sandryhaila, Stephen Lin, William G. Jenkinson, Jeremiah MacSleyne, Christopher Hoffman, Philipp Cuadra, Charles Jackson, Markus Püschel, Jelena Kovacevick; *An Adaptive Multiresolution Approach to Fingerprint Recognition* in PROCEEDINGS OF IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 2007 [DOI]

**My contribution:** *I developed and implemented one of the methods presented in the paper.*

## Review Articles

48. Aabid Shariff, Joshua Kangas, **Luis Pedro Coelho**, Shannon Quinn, and Robert F. Murphy; *Automated Image Analysis for High Content Screening and Analysis* in JOURNAL OF BIOMOLECULAR SCREENING, August 2010, pp. 726–734 [DOI]

## Books



1. **Luis Pedro Coelho**, Willi Richert; *Building Machine Learning Systems with Python*, Packt Publishing, 2013 (first ed.); 2015 (second ed.); 2018 (third ed.)

## Invited Talks

1. *High-throughput fluorescence microscopy of environmental samples obtains single-cell phenotypic measurements*, Paris Single Cell Day 2018, October 2018
2. *Metagenomics based investigations of microbial communities*, Symposium on Computational Biology at Fudan University, Shanghai, August 2018
3. *Life in Words and Pictures: sequencing and imaging for analysing microbial communities*, Dept. of Biomedical Research, University of Bern, November 2017.
4. *Statistical tools for analysing the microbiome in environmental and clinical applications*, Novo Nordisk Foundation Center for Basic Metabolic Research, October 2017
5. *Looking at the oceans with computer vision in Python*, invited keynote at PyCon Firenze (Italy), April 2016
6. *Structure and function of the ocean microbiome*, lecture at the course *Biology of ecological systems*, Ecole Nationale Supérieure (France), December 2015
7. *Life in words and pictures: Sequencing and microscopy for the analysis of microbial communities*, Luxembourg Centre for Systems Biomedicine, September 2015
8. *An analysis of 243 metagenomes collected by the Tara Oceans Projects*, Station Biologique de Roscoff, July 2015
9. *Machine Learning: What it can do for you (and what it cannot)*, BrightTalk online seminar, May 2015
10. *Webcast: Penalized Linear Regression in Python*, O'Reilly Webcasts, October 2014
11. *Python for Computer Vision in Biology and Beyond*, Python San Sebastian, keynote, September 2014
12. *Large Scale Analysis of Bioimages Using Python*, International Workshop on Technical Computing for Machine Learning and Mathematical Engineering, Leuven (Belgium), September 2014
13. *Mahotas and the Python Ecosystem for Bioimage Informatics Applications*, European Bioimage Analysis Symposium, Barcelona, October 2013
14. *Organizing the Proteome with Location and Function Topics*, Freiburg Institute for Advanced Studies, April 2013
15. *Modeling Subcellular Location from Images and Other Sources of Information*, Luxembourg Centre for Systems Biomedicine, July 2012
16. *Modeling Subcellular Location from Images and Other Sources of Information*, Priberam Machine Learning Lunch Seminar, June 2012
17. *Modeling Subcellular Location from Images and Other Sources of Information*, EAO Seminar (Instituto Gulbenkian da Ciência), Oeiras, March 2012
18. *Learning Subcellular Location from Images and Other Sources of Information*, KDBIO (Knowledge Discovery and Bioinformatics) seminar, Lisbon, February 2012
19. *Bioimage Informatics: Computer Vision for Biology*, EMBO Practical Course on Microscopy: from single molecules to animals, Pretoria, November 2011
20. *Studying the subcellular location space with bioimages and other data modalities*, University of Delaware, Computer and Information Sciences Department, September 2011

21. *Proteome-scale analysis and modeling of subcellular location*, 4th CeBiTec Symposium BioImaging, Bielefeld (Germany), 25–27 August 2009
22. *Unsupervised Mixture Pattern Unmixing*, University of Bielefeld International Graduate School of Bioinformatics and Genome Research, July 2008

## Selected Conference Talks

1. **Luis Pedro Coelho\***, Sebastien Colin\*, Shinichi Sunagawa, Eric Karsenti, Peer Bork, Rainer Pepperkok, and Colombar de Vargas *Environmental High-content Fluorescence Microscopy (e-HCFM) of Tara Oceans Samples Provides a View of Global Ocean Protist Biodiversity*, Ocean Sciences, New Orleans, 2016
2. Shinichi Sunagawa, **Luis Pedro Coelho**, Samuel Chaffron, Eric Karsenti, Jeroen Raes, Silvia Acinas, Peer Bork *Structure and function of the global ocean microbiome*, ASLO, Granada (Spain) 2015
3. **Luis Pedro Coelho** *Jug: Reproducible Research in Python*, BOSC, Berlin 2013
4. Rita Reis and **Luis Pedro Coelho**; *Using Theatre to Fight HIV/AIDS in Mozambique*, National Conference of the Association for Theatre in Higher Education, Chicago 2011
5. **Luis Pedro Coelho** and Robert F. Murphy; *Determining Resolvable Subcellular Location Categories as a Function of Image Resolution*, 24th ISAC Congress, Budapest 2008