Pierre Baldi

Department of Computer Science
Institute for Genomics and Bioinformatics
Center for Machine Learning and Intelligent Systems
University of California, Irvine
Irvine, CA 92697-3435
(949) 824-5809
(949) 824-9813 FAX
Assistant: Janet Ko (janetmk@uci.edu)
pfbaldi@ics.uci.edu
www.ics.uci.edu/~pfbaldi
www.igb.uci.edu

Professional Experience

- November 2018 to present: Fellow, Center for the Neurobiology of Learning and Memory, UCI
- July 2018 to present: Distinguished Professor, Department of Computer Science, UCI
- February 2017 to present: Joint appointment in the Department of Mathematics, UCI
- February 2015 to present: Joint appointment in the Department of Statistics, UCI
- October 2006 to June 2018t: Chancellor's Professor, UCI
- September 2006 to present: Founding Associate-Director, Center for Machine Learning and Intelligent Systems
- January 2001 to present: Founding Director Institute for Genomics and Bioinformatics.
- June 2001 to present: Professor Department of Computer Science, School of Information and Computer Sciences, University of California, Irvine. [Joint appointments in the Department of Biological Chemistry, College of Medicine, Department of Biomedical Engineering School of Engineering]
- June 2001 to July 2019: Joint appointment in the Department of Developmental and Cell Biology, School of Biological Sciences.
- January 2002 to December 2005: Application Layer Leader (Digitally Enabled Medicine) for the California Institute for Telecommunications and Information Technology [Calit2[2]]
- July 1999 to May 2001: Associate Professor, Department of Information and Computer Science, University of California, Irvine. [Joint appointment in the Department of Biological Chemistry, College of Medicine and the Department of Developmental and Cell Biology, School of Biological Sciences]
- 1991 to June 1999: Chairman and CEO, Net-ID, Inc
- January 1999: Visiting Professor, Department of Computer Science, University of Florence
- 1995 to 1996: Member of the Professional Staff, Division of Biology, California Institute of Technology
- 1988 to 1995: Member of the Technical Staff in the Nonlinear Science and Information Processing Group at the Jet Propulsion Laboratory, and Visiting Research Associate, Division of Biology, California Institute of Technology

- Summer 1988: Visiting Research Mathematician, Department of Mathematics, University of California, San Diego
- 1986-1988: Visiting Lecturer, Department of Mathematics, University of California, San Diego

Education

- 1986: Ph.D. Mathematics, California Institute of Technology
- 1983: MS Computer Science and Engineering, ENSTA, Paris
- 1981: D.E.A Mathematics, University of Paris VII
- 1980: MS Mathematics, University of Paris VII
- 1980: MS Psychology, University of Paris X

Honors

- 2019 Prominent Artificial Intelligence Journal Paper Award for the paper "The dropout learning algorithm" (published in 2014).
- 2019 Named Top 100 AI Leaders in Drug Discovery and Advanced Healthcare in the world by the Deep Knowledge Analytics
- 2018 Distinguished Professor, UCI
- 2017 Member of NOvA (neutrino experiment consortium)
- 2015 Elected short-term associated member of ATLAS at CERN
- 2014 Google Faculty Research Award
- 2013 Fellow International Society for Computational Biology (ISCB)
- 2012 Fellow Association Computing Machinery (ACM)
- 2011 Fellow Institute of Electrical and Electronics Engineers (IEEE)
- 2010 Eduardo R. Caianiello Prize for Scientific Contributions to the Field of Neural Networks
- 2009 Dean's Award for Research
- 2008 Fellow American Association Advancement of Science (AAAS)
- 2007 Fellow Association Advancement Artificial Intelligence (AAAI)
- 2006 Chancellor's Professor
- 2006 Microsoft Faculty Research Award
- 2006 IEEE Senior Member
- 2001 Certificate of Recognition for Neural Network Invention by NASA
- 1999 Laurel Wilkening Faculty Innovation Award at UCI
- 1998 Certificate of Service for the Caltech Alumni Fund for the 1997-98 fund raising effort.
- Recipient of the 1993 Lew Allen Award at JPL
- 1992 Certificate of Service with Distinction as Associate Editor of the *IEEE Transactions on Neural Networks*
- 1985 Bohnenblust Prize, Caltech

PUBLICATIONS

Refereed Journals

- J323. G. Urban, C. Magnan, and **P. Baldi**. SSpro/ACCpro 6: Almost Perfect Prediction of Protein Secondary Structure and Relative Solvent Accessibility Using Profiles, Deep Learning, and Structural Similarity. *Bioinformatics*, in press, (2022).
- J322. A. Tavakoli, A. Mood, D. Van Vranken, and **P. Baldi**. Quantum Mechanics and Machine Learning Synergies: Graph Attention Neural Networks to Predict Chemical Reactivity. Journal of Chemical Information and Modeling, in press, (2021).
- J321. Shogo Sato, Kenneth A. Dyar, Jonas T. Treebak, Astrid Linde Basse, Mirena Schönke, Siwei Chen, Muntaha Samad, **Pierre Bardi**, Dominik Lutter, Juleen R. Zierath, Paolo Sassone-Corsi. Atlas of Exercise Metabolism Reveals Time-Dependent Systemic Metabolic Homeostasis. *Cell Metabolism*, in press, (2021).
- J320. Michael James Fenton, Alexander Shmakov, Ta-Wei Ho, Shih-Chieh Hsu, Daniel Whiteson, and **Pierre Baldi**. Permutationless many-jet event reconstruction with symmetry preserving attention networks. Physical Review D, in press, (2021).
- J319. K. Lin, G. Urban, M. C. Yang, L.C Lee, D.W. Lu, W. L. M. Alwardand, and **P. Baldi**. Accurate Identification of the Trabecular Meshwork under Gonioscopic view in Real Time using Deep Learning. *Ophthalmology Glaucoma*, in press, (2021).
- J318. **P. Baldi**. A Call for a Public Open Database of All Chemical Reactions. Journal of Chemical Information and Modeling, in press, (2021).
- J317. Amal Alachkar *, Roudabeh Vakil Monfared, Wedad Alhassen, Tri Minh Truong, Michael Angelo Maglalang Gonzales, Vincent Vachirakorntong, Siwei Chen, **Pierre Baldi**, Olivier Civelli. Transcriptome Profiling of Dysregulated GPCRs Reveals Overlapping Patterns across Psychiatric Disorders and Age-Disease Interactions. *Cells*, in press, (2021).
- J316. B. Shahbaba, L. Li, F. Agostinelli, M. Saraf, G. Elias, **P. Baldi**, and N. Fortin. Hippocampal Ensembles Represent Sequential Relationships Among Discrete Nonspatial Events. *Nature Communications*, in press, (2021). Also: https://www.biorxiv.org/content/10.1101/840199v1
- J315. Julian Collado, Kevin Bauer, Edmund Witkowski, Taylor Faucett, Daniel Whiteson, and **Pierre Baldi**. Learning to Isolate Muons. *Journal of High Energy Physics*, 2021, 200 (2021). https://doi.org/10.1007/JHEP10(2021)200.
- J314. Siwei Chen, Wedad Alhassen, Roudabeh Vakil Monfared, Benjamin Vachirakorntong, Surya Nauli, **Pierre Baldi**, and Amal Alachkar. Dynamic Changes in Brain Cilia Transcriptomes across the Human Lifespan. *International Journal of Molecular Sciences*, 22, 19, 10387, https://doi.org/10.3390/ijms221910387, (2021).
- J313. Stephen McAleer; Alexander Fast; Yuntian Xue; Magdalene J. Seiler; William C. Tang; Mihaela Balu; **Pierre Baldi**; Andrew W. Browne. Deep Learning–Assisted Multiphoton Microscopy to Reduce Light Exposure and Expedite Imaging in Tissues With High and Low Light Sensitivity. *Translational Vision Science & Technology*, October 2021, Vol.10, 30, (2021). doi:https://doi.org/10.1167/tvst.10.12.30.
- J312. E. Lashgari, J. Ott, **P. Baldi**, and U. Maoz. An End-to-End CNN with Attentional Mechanism Applied to Raw EEG in a BCI Classification Task. *Journal of Neural Engineering*, in press, (2021).
- J311. **P. Baldi**, W. Alhassen, S. Chen, H. Nguyen, M. Khoudari, A. Alachkar. Large-Scale Analysis Reveals Spatiotemporal Circadian Patterns of Cilia Transcriptomes in the Primate Brain. Journal of Neuroscience Research, First published: 26 July 2021 https://doi.org/10.1002/jnr.24919, (2021).
- J310. Carolina M. Greco1, Kevin B. Koronowski, Jacob G. Smith, Jiejun Shi, Siwei Chen, Muntaha Samad, Patrick-Simon Welz, Valentina M. Zinna, Thomas Mortimer, Kohei Shimaji, Tomoki Sato, Paul Petrus, Arun Kumar, Mireia Vaca Dempere, Cassandra Van, Kenneth A. Dyar, Dominik Lutter, Marcus M. Seldin, Wei Li, **Pierre Baldi**, Pura Muñoz-Cánoves, Salvador Aznar Benitah, and Paolo Sassone-Corsi1. Integration of Feeding

Behaviour by the Liver Circadian Clock Reveals Network Dependency of Metabolic Rhythms. *Science Advances*, in press, (2021).

- J309. Chen, Siwei; Lee, Justine; Truong, Tri Minh; Alhassen, Sammy; **Baldi, Pierre**; Alachkar, Amal. Age-Related Neurometabolomic Signature of Mouse Brain. ACS Chemical Neuroscience, in press, (2021).
- J308. Julian Collado, Jessica N. Howard, Taylor Faucett, Tony Tong, **Pierre Baldi**, and Daniel Whiteson. Learning to identify electrons. Physical Review D, 103, 11, 116028, (2021).
- J307. J. Ott, D. Bruyetter, C. Arbuckle, D. Balsz, S. Hecth, L. Shubitz, and **P. Baldi**. Detecting Pulmonary Coccidioidomycosis with Deep Convolutional Neural Networks. *Machine Learning with Applications*, in press, (2021).
- J306. **P. Baldi** and R. Vershynin. A Theory of Capacity and Sparse Neural Encoding. *Neural Networks*, in press, (2021).
- J305. L. Hertel, and **P. Baldi**. Reproducible Hyperparameter Optimization. *Journal of Computational and Graphical Statistics*, in press, (2021).
- J304. Rianne Campbell, Siwei Chen, Joy Beardwood, Alberto Lopez, Lilyana Pham, Ashley Keiser, Jess Childs, Dina Matheos, Vivek Swarup, **Pierre Baldi**, and Marcelo Wood. Cocaine induces paradigm-specific changes to the transcriptome within the Ventral Tegmental Area. *Neuropsychopharmacology*, in press, (2021).
- J303. S. Chen, S. Alhassen, **P. Baldi**, G. Abott, A. Alachkar. Intergenerational Stress Transmission is Associated with Brain Metabotranscriptome Remodeling and Mitochondrial Dysfunction. *Communications Biology*, in press, (2021).
- J302. Griffin Mooers, Michael Pritchard, Tom Beucler, Jordan Ott, Galen Yacalis, **Pierre Baldi**, Pierre Gentine Assessing the Potential of Deep Learning for Emulating Cloud Superparameterization in Climate Models with Real-Geography Boundary Conditions. *Journal of Advances in Modeling Earth Systems*, 13, 5, First published: 23 April 2021 https://doi.org/10.1029/2020MS002385 ,(2021).
- J301. A. Tavakoli, F. Agostinelli, and **P. Baldi**. SPLASH: Learnable Activation Functions for Improving Accuracy and Adversarial Robustness. *Neural Networks*, 140, 1-12, (2021).
- J300. Kadish, Dora; Mood, Aaron; Tavakoli, Mohammadamin; Gutman, Eugene; **Baldi, Pierre**; Van Vranken, David. Methyl Cation Affinities of Canonical Organic Functional Groups. *The Journal of Organic Chemistry*, in press, (2021).
- J299. Tom Beucler, Michael Pritchard, Stephan Rasp, Jordan Ott, **Pierre Baldi**, and Pierre Gentine. Enforcing analytic constraints in neural networks emulating physical systems. *Physical Review Letters*, in press, (2021). Also: arXiv: http://arxiv.org/abs/1909.00912.
- J298. Wedad Alhassen, Siwei Chen, Marquis Vawter, Brianna Kay Robbins, Henry Nguyen, Thant Nyi Myint, Yumiko Saito, Anton Schulmann, Surya M. Nauli, Olivier Civelli, **Pierre Baldi**, and Amal Alachkar. Patterns of Cilia Gene Dysregulations in Major Psychiatric Disorders. *Progress in Neuropsychopharmacology & Biological Psychiatry*, page 110255, (2021).
- J297. Y. Lu, J. Collado, D. Whiteson, and **P. Baldi**. SARM: Sparse Autoregressive Models for Scalable Generation of Sparse Images in Particle Physics. *Physical Review D*, 103, 3,036012, (2021).
- J296. B. Abi, A. Abed Abud, R. Acciarri, M.A. Acero, G. Adamov, M. Adamowski, D. Adams, P. Adrien, M. Adinolfi, Z. Ahmad, et al. (DUNE Collaboration). First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN Neutrino Platform. Journal of Instrumentation, Volume 15, December 2020.
- J295. C. Greco, M. Cervantes, J. Fustin, K. Ito, N. Ceglia, M. Samad, J. Shi, K. Koronowski, I. Forne, S. Ranjit, J. Gaucher, K. Kinouchi, R. Kojima, E. Gratton, W. Li, **P. Baldi**, A. Imhof, H. Okamura, P. Sassone-Corsi. S-Adenosyl-L-Homocysteine Hydrolase Links

Methionine Metabolism to the Circadian Clock and Chromatin Remodeling. *Science Advances*, 6, 51, eabc5629, (2020).

- J294. Thrift, William; Ronaghi, Sasha; Samad, Muntaha; Wei, Hong; Nguyen, Dean; Cabuslay, Antony; Groome, Chloe; Santiago, Peter; **Baldi, Pierre**; Hochbaum, Allon; Ragan, Regina Deep Learning Analysis of Vibrational Spectra of Bacterial Lysate for Rapid Antimicrobial Susceptibility Testing. *ACS Nano*, 14, 11, 15336-15348, Publication Date (Web): October 23, 2020, DOI: 10.1021/acsnano.0c05693, (2020).
- J293. Mehran J. Umerani, Preeta Pratakshya, Atrouli Chatterjee, Juana A. Cerna, Hoshin Kim, Gregor Ilc, Matic Kovacic, Christophe Magnan, Benedetta Marmiroli, Barbara Sartori, Andrew W. Bartlett, Erica M. Leung, Zhijing Feng, Kyle L. Naughton, Brenna Norton-Baker, Long Phan, James Long, Alex Allvato, Jessica E. Leal-Cruz, Qiyin Lin, **Pierre Baldi**, Sigrid Bernstorff, Janez Plavec, Yara Yingling, Alon A. Gorodetsky Structure, Self-Assembly, and Properties of a Truncated Reflectin Variant, *Proceedings of the National Academy of Sciences*, DOI number 10.1073/pnas.2009044117, (2020).
- J292. Paola Tognini, Muntaha Samad, Kenichiro Kinouchi, Yu Liu, Jean-Christophe Helbling, Marie-Pierre Moisan, Kristin L. Eckel-Mahan, **Pierre Baldi**, and Paolo Sassone-Corsi. Reshaping Circadian Metabolism in the Suprachiasmatic Nucleus and Prefrontal Cortex by Nutritional Challenge. *PNAS*, 117, 47, 29904--29913, first published November 10, 2020; https://doi.org/10.1073/pnas.2016589117, (2020).
- J291. L. Hertel, Julian Collado, Peter Sadowski, Jordan Ott, **Pierre Baldi**. Sherpa: Robust Hyperparameter Optimization for Machine Learning. *SoftwareX*, 12, (2020). Also: arXiv:2005.04048.
- J290. C. Lee, M. Samad, I. Hofer, **P. Baldi**, and M. Cannesson. Development and Validation of an Interpretable Neural Network for Prediction of Postoperative In-hospital Mortality, *npj Digital Medicine*, 4, 1, 1--9, (2021).
- J289. Gregor Urban, Nate Feil, Ella Csuka, Kiana Hashemi, Chloe Ekelem, Franchesca Choi, Natasha Atanaskova Mesinkovska, and **Pierre Baldi**. Combining Deep Learning with Optical Coherence Tomography Imaging to Determine Scalp Hair and Follicle Counts. *Lasers in Surgery and Medicine*, 53:171–178 (2021). First published: 22 September 2020, **https://doi.org/10.1002/lsm.23324**.
- J288. Debora Napoli, Leonardo Lupori, Raffaele Mazziotti, Giulia Sagona, Sara Bagnoli, Muntaha Samad, Erika Kelmer Sacramento, Joanna Kirkpatrick, Elena Putignano, Siwei Chen, Eva Terzibasi Tozzini, Paola Tognini, Jessica Kwok, **Pierre Baldi**, Alessandro Cellerino, and Tommaso Pizzorusso. MiR-29 coordinates age-dependent plasticity brakes in the adult visual cortex. EMBO Reports, 21, 11, 1-19, DOI: 10.15252/embr.202050431, (2020).
- J287. K. J. Debski1, N. Ceglia, A. Ghestem, A. I. Ivanov, G. E. Brancati, S. Bröer, A. M. Bot1, J. A. Müller, S. Schoch, A. Becker, W. Löscher, M. Guye, P. Sassone-Corsi, K. Lukasiuk, **P. Baldi**, C. Bernard. The circadian dynamics of the hippocampal transcriptome and proteome is altered in experimental epilepsy. *Science Advances*, 6, 41, eaat5979, (2020).
- J286. G. Urban, M. Torrisi, C. Magnan, G. Pollastri, and **P. Baldi**. Protein Profiles: Biases and Protocols *Computational and Structural Biotechnology Journal*, 18, 2281--2289, (2020).
- J285. Pietro DiLena and **Pierre Baldi**. Fold recognition by scoring protein map similarities using the congruence coefficient. *Bioinformatics*, in press, (2020).
- J284. J. Ott, M. Pritchard, N. Best, E. Linstead, M. Curcic, and **P. Baldi**. A Fortran-Keras Deep Learning Bridge for Scientific Computing. *Scientific Programming*, (2020). Received Article of the Year 2020 Award.

- J283. Y. Liu, **P. Baldi**, P. Sassone-Corsi, E. Borrelli. Cocaine-mediated Circadian Reprogramming in the Striatum Through Dopamine-driven PPARγ Activation. *Nature Communications*, 11, 1, 1--14, (2020).
- J282. Carolina Magdalen Greco, Stefano Garretto, Emilie Montellier, Yu Liu, Siwei Chen, **Pierre Baldi**, Paolo Sassone-Corsi1 and Jacopo Lucci. A non-pharmacological approach in the gut triggers distal metabolic rewiring capable of ameliorating diet-induced metabolic dysfunction encmpasserd by metabolic syndrome. *Scientific Reports*, 10, 1, 1--13, (2020).
- J281. Siwei Chen, Wedad Alhassen, Ryan Yoshimura, Angele De Silva, Geoffrey W. Abbott, **Pierre Baldi**, Amal Alachkar. Metabolomic and Transcriptomic Signatures of Prenatal Excessive Methionine in Mice Support Nature Rather than Nurture in the Pathogenesis of Schizophrenia. *Communications Biology*, 3, 1, 1-12, (2020).
- J.280 G. Urban, S. Porhemmat, M. Stark, B. Feeley, K. Okada, and **P. Baldi**. Classifying Shoulder Implants in X-ray Images using Deep Learning. *Computational and Structural Biotechnology Journal*, in press, (2020).
- J. 279 J. Ott, E. Linstead, N. LaHaye, and **P. Baldi**. Learning in the Machine: To Share or Not to Share? *Neural Networks*, 126, 235-249, (2020). Available online March 25, 2020. https://doi.org/10.1016/j.neunet.2020.03.016.
- J278. A. Mood, A. Tavakoli, E. Gutman, D. Kadish, **P. Baldi**, and D. VanVranken. Methyl Anion Affinities of the Canonical Organic Functional Groups. *The Journal of Organic Chemistry.* DOI: 10.1021/acs.joc.9b03187 • Publication Date (Web): 29 Jan 2020.
- J277. Ira S. Hofer, Christine Lee, Eilon Gabel, **Pierre Baldi,** and Maxime Cannesson. Development and Validation of a Deep Neural Network Model to Predict Postoperative Mortality, Acute Kidney Injury, and Reintubation using a single feature set. *npj Digital Medicine*, 3, 1, 1--10, (2020). DOI : 10.1038/s41746-020-0248-0, NPJDIGITALMED-00456, (2020).
- J276. Margit Juhasz, SiWei Chen, Arash Khosrovi-Eghbal, Chloe Ekelem, Yessica Landaverde, **Pierre Baldi**, Natasha Atanaskova Mesinkovska. Characterizing the skin and gut microbiome of alopecia areata patients. *SKIN The Journal of Cutaneous Medicine*, 4,1, DOI: 10.25251/skin.4.1.4 (2020).
- J275. Giorgio Ramadori, Rafael M. Ioris, Zoltan Villanyi, Raquel Firnkes, Olesya O. Panasenko, George Allen, Georgia Konstantinidou, Ebru Aras, Xavier Brenachot, Tommasina Biscotti, Anne Charollais, Michele Luchetti, Fedor Bezrukov, Alfredo Santinelli, Muntaha Samad, **Pierre Baldi**, Martine A. Collart, and Roberto 6 Coppari. FKBP10 regulates protein translation to sustain lung cancer growth. *Cell Reports*, in press, (2019).
- J274. L. Li, N. Nayak, J. Bian, **P. Baldi**. Efficient neutrino oscillation parameter inference using Gaussian processes. *Physical Review D*, 101, 012001 – Published 2 January 2020, DOI:https://doi.org/10.1103/PhysRevD.101.012001, (2020).
- J273. Jonathan Gaucher, Kenichiro Kinouchi, Nicholas Ceglia, Emilie Montellier, Shahaf Peleg, Carolina Magdalen Greco, Andreas Schmidt, Ignasi Forne, Selma Masri, **Pierre Baldi**, Axel Imhof, Paolo Sassone-Corsi. Distinct Metabolic Adaptation of Liver Circadian Pathways to Acute and Chronic Patterns of Alcohol Intake. *Proceedings of the National Academy of Sciences USA*, USA, December 10, 2019, 116, (50), 25250-25259; https://doi.org/10.1073/pnas.1911189116 , (2019)., (2019).
- J272. M.A. Acero et al. (NovA collaboration). First measurement of neutrino oscillation parameters using neutrinos and antineutrinos by NovA. *Phys. Rev. Lett.* 123, 151803 – Published 11 October 2019.
- J271. Christine Lee, Christopher Wray, Vatche Agopian, Gregor Urban, Pierre Baldi, Maxime Cannesson, Brent Ershoff. Training and Validation of Deep Neural Networks for the Prediction of 90-Day Post-Liver Transplant Mortality Using UNOS Registry

Data.*Transplantation Proceedings*, 52.1, 246--258, https://doi.org/10.1016/j.transproceed.2019.10.019, (2020).

- J270. Heather Karner; Chiu-Ho Webb; Sarah Carmona; Yu Liu; Benjamin Lin; Micaela Erhard; Dalen Chen; **Pierre Baldi**; Robert C Spitale; Sha Sun. Functional conservation of lncRNA JPX despite sequence and structural divergence. *Journal of Molecular Biology*, 432, 283—300, https://doi.org/10.1016/j.jmb.2019.09.002, (2020).
- J269. F. Agostinelli, S. McAleer, A. Shmakov, and **P. Baldi**. Solving the Rubik's cube with deep reinforcement learning and search. *Nature Machine Intelligence*, https://doi.org/10.1038/s42256-019-0070-z, (2019).
- J268. **P. Baldi** and R. Vershynin. Polynomial threshold functions, hyperplane arrangements, and random tensors. *SIAM Journal on Mathematics of Data Science (SIMODS)*, 1, 4, 699-729, URL: https://epubs.siam.org/toc/sjmdaq/1/3 , DOI: 10.1137/19M1257792, (2019).
- J267. Yu-Han Chen, Ching-Chieh Su, Wu Deng, Leslie Lock, Peter Donovan, Matthew A Kayala, **Pierre Baldi**, Hsiao-Chen Lee, Yumay Chen, Ping H. Wang. Mitochondrial Akt Signaling Modulated Reprogramming of Somatic Cells. *Scientific Reports,* in press, (2019).
- J266. **P. Baldi** and R. Vershynin. Neuronal Capacity. *Journal of Statistical Mechanics, Theory and Experiment*. Special Issue on Machine Learning, 124012, (2019). This is an updated version of the article with the same name published at the NeurIPS conference. Online at stacks.iop.org/JSTAT/2019/124012. https://doi.org/10.1088/1742-5468/ab3285.
- J265. Alejandro Sanchez, Carlos Castro, Dora-Luz Flores, Everardo Gutierrez, and **Pierre Baldi**. Gap Junction Channels of Innexins and Connexins: Relations and Computational Perspectives. *International Journal of Molecular Sciences*, in press, (2019).
- J264. **P. Baldi** and R. Vershynin. The capacity of feedforward neural networks. *Neural Networks*, 116, August 2019, Pages 288-311, (2019). Available online 22 April 2019. https://doi.org/10.1016/j.neunet.2019.04.009. Also: Arxiv 1901.00434.
- J263. P. Baldi and and B. Shahbaba. Bayesian Causality. *The American Statistician*. 1—9, (2019). Published Online: 26 Aug 2019, in press, (2019).
- J262. Shogo Sato, Astrid Linde Basse, Milena Schönke, Siwei Chen, Muntaha Samad, Ali Altıntaş, Rhianna C. Laker, Emilie Dalbram, Romain Barrès, **Pierre Baldi**, Jonas Thue Treebak, Juleen R. Zierath, and Paolo Sassone-Corsi. Time of Exercise Specifies the Impact on Muscle Metabolic Pathways and Systemic Energy Homeostasis. *Cell Metabolism*, 30, 1, 92--110, published online on April 18th and in print in the July 2019 issue, (2019).
- J261. Kevin B. Koronowski, Kenichiro Kinouchi, Patrick-Simon Welz, Valentina Maria Zinna, Jiejun Shi, Muntaha Samad, Siwei Chen, C. Magnan, Jason Kinchen, Wei Li, **Pierre Baldi**, Salvador Aznar Benitah, and Paolo Sassone-Corsi1. Defining the Independence of the Liver Circadian Clock. *Cell*, 177, 6, 1448--1462, (2019).
- J260. Ebru Aras, Giorgio Ramadori, Kenichiro Kinouchi, Yu Liu, Rafael M. Ioris, Xavier Brenachot, Sanda Ljubicic, Christelle Veyrat-Durebex, Silvia Mannucci, Mirco Galié, **Pierre Baldi,** Paolo Sassone-Corsi, and Roberto Coppari. Light entrains diurnal changes in insulin sensitivity of skeletal muscle via ventromedial hypothalamic neurons. *Cell Reports*, (2019).
- J259. X. Li, Q. Zhao, W. Wei, Q. Lin, C. Magnan, M. Emami, L. E. Wearick da Silva, T. W. Viola, P. Marshall, J. Edmunds, S. Nainar, C. Broberg Vågbø, L. Leighton, E. Zajackowski, K. Ke, R. Grassi-Oliveira, M. Bjørås, **P. Baldi**, R. C. Spitale, and T. W. Bredy. The DNA modification N6-methyl-2'-deoxyadenosine (m6dA) drives activity-induced gene expression and is required for fear extinction. *Nature Neuroscience*, 22, 4, 534, (2019).
- J258. **P. Baldi**, J. Bian, L. Hertel, et al. Improved energy reconstruction in NOvA with regression convolutional neural networks. *Physical Review D*, 99, 012011 – Published 24 January, (2019).

- J257. L. Li, A. J. Holbrook, B. Shahbaba, and **P. Baldi**. Neural Network Gradient Hamiltonian Monte Carlo. *Computational Statistics*, 1-19, DOI 10.1007/s00180-018-00861-z, (2019).
- J256. Kenichiro Kinouchi, Christophe Magnan, Nicholas Ceglia, Yu Liu, Marlene Cervantes, **Pierre Baldi**, Selma Masri, and Paolo Sassone-Corsi. Fasting Imparts a Switch to Alternative Circadian Pathways in Liver and Muscle. *Cell Reports*, in press, (2018).
- J255. David Westergaard, Pope Moseley, Freja Karuna Hemmingsen Sørup, **Pierre Baldi**, Soren Brunak. Population-Wide Analysis of Differences in Disease Progression Patterns in Men and Women. *Nature Communications,* 10, 666 (2019). https://doi.org/10.1038/s41467-019-08475-9, (2019).
- J254. **P. Baldi** and P. Sadowski. Learning in the Machine: Recirculation is Random Backpropagation. *Neural Networks*, 108, 479-494, (2018).
- J253. S. Shao, S. McAleer, R. Yan, and **P. Baldi**. Highly-Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Transactions on Industrial Informatics (TII*), Special Issue on Big Data Analytics in Intelligent Manufacturing, 15,4,2446--2455, (2018).
- J252. J. L. Kwapis, Y. Alaghband, E. A. Kramár, A. J. López, A. Vogel Cierna, A. O. White, G. Shu, D. Rhee, C. M. Michael, E. Montellier, Y. Liu, C. Magnan, S. Chen, P. Sassone-Corsi, **P. Baldi**, D. P. Matheos, and M. A. Wood. Epigenetic Regulation of the Circadian Gene Per1 in the Hippocampus Contributes to Age-Related Changes in Memory and Synaptic Plasticity. *Nature Communications*, 9, 1-14, Article number: 3323, Published 20 August, (2018).
- J251. Sijia Liu, Haiming Chen, Scott Ronquist, Laura Seaman, Nicholas Ceglia, Walter Meixner, Lindsey A. Muir, Pin-Yu Chen, Gerald Higgins, **Pierre Baldi**, Steve Smale, Alfred Hero, Indika Rajapakse. Genome Architecture Mediates Transcriptional Control of Human Myogenic Reprogramming, ISCIENCE, 6, 232–246, Published online: August 7, (2018).
- J250. Kenneth A. Dyar, Dominik Lutter, Anna Artati, Nicholas J. Ceglia, Yu Liu, Danny Armenta, Martin Jastroch, Sandra Schneider, Sara de Mateo, Serena Abbondante, Paola Tognini, Ricardo Orozco-Solis, Saurabh Sahar, Christina Wang, Ronald S. Swerdloff, Seba Nadeef, Valerio Orlando, N. Henriette Uhlenhaut, Selma Masri, Pierre Magistretti , Emiliana Borrelli, **Pierre Baldi**, Jerzy Adamski, Matthias Tschöp, Kristin Eckel-Mahan, Paolo Sassone-Corsi. A circadian metabolomics atlas reveals systems-wide coordination and communication between clocks. *Cell,* 174, 6, P1571-1585.E11, SEPTEMBER 06, 2018, DOI:https://doi.org/10.1016/j.cell.2018.08.042 (2018).
- J249. Gregor Urban, Priyam Tripathi, Talal Alkayali, Mohit Mittal, Farid Jalali, William Karnes, and **Pierre Baldi**. Deep Learning Localizes and Identifies Polyps in Real Time with 96% Accuracy in Screening Colonoscopy. *Gastroenterology*, Volume 155, Issue 4, Pages 1069–1078, (2018).
- J248. G. Urban, K. Bache, D. T.T. Phan, A. Sobrino, A. K. Shmakov, S. J. Hachey, C.W. Hughes, and **P. Baldi**. Deep Learning for Drug Discovery and Cancer Research: Automated Analysis of Vascularization Images. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 16, 3, 1029—1035, (2018).
- J247. Cuong Nguyen, William J. Thrift, Arunima Bhattacharjee, Saba Ranjbar, Tara Gallagher, Mahsa Darvishzadeh-Varcheie, Robert Sanderson, Filippo Capolino, Katrine Whiteson, **Pierre Baldi**, Allon Hochbaum, and Regina Ragan. Longitudinal Monitoring of Biofilm Formation via Robust SERS Quantification of Pseudomonas aeruginosa-Produced Metabolite. *ACS Applied Materials and Interfaces*, 10, 15, 12364—12373, (2018).
- J246. **P. Baldi**, P. Sadowski, and Z. Lu. Learning in the Machine: Random Backpropagation and the Deep Learning Channel. *Artificial Intelligence*, 260, 1-35, (2018).

- J245 P. Chang, J. Grinband, B.D. Weinberg, M. Bardis, M. Khy, G. Cadena, M.-Y. Su, S. Cha, C.G. Filippi, D. Bota, **P. Baldi**, L.M. Poisson, R. Jain and D. Chow. Deep-Learning Convolutional Neural Networks Accurately Classify Genetic Mutations in Gliomas. *American Journal of Neuroradiology* May 2018, DOI: https://doi.org/10.3174/ajnr.A5667. Selected as the best original research paper accepted by the American Journal of Neuroradiology in 2018.
- J244. N. Ceglia, Y. Liu, S. Chen, F. Agostinelli, K. Eckel-Mahan, P. Sassone-Corsi, and **P. Baldi**. CircadiOmics: Circadian Omic Data Web Portal. *Nuclei Acids Research*, Web Server Issue, 46, Issue W1, 2 July 2018, Pages W157–W162, https://doi.org/10.1093/nar/gky441, (2018).
- J243. G. Urban, N. Subrahmanya , and **P. Baldi**. Inner and Outer Recursive Neural Networks for Chemoinformatics Applications. *Journal of Chemical Information and Modeling*, 58 2, 207–211, Publication Date (Web): January 10, DOI: 10.1021/acs.jcim.7b00384, (2018).
- J242. C. K. Lee, I. Hofer, E. Gabel, **P. Baldi**, and M. Cannesson. Development and Validation of a Deep Neural Network Model for Prediction of Postoperative In-Hospital Mortality. *Anesthesiology*, Published on April 16, 2018. doi:10.1097/ALN.0000000000002186, 129, 4, 649--662, (2018).
- J241. David Fooshee, Aaron Mood, Eugene Gutman, Amin Tavakoli, Gregor Urban, Frances Liu, Nancy Huynh, David Van Vranken, and **Pierre Baldi**. Deep Learning for Chemical Reaction Prediction. *Molecular Systems Design & Engineering*, Royal Society of Chemistry, 3, 442 – 452, , DOI: 10.1039/c7me00107j, (2018).
- J240. **P. Baldi**. Deep Learning in Biomedical Data Science. *Annual Review of Biomedical Data Science*, Volume 1, 181-205 (Volume publication date July 2018) , First published as a Review in Advance on April 25, 2018, https://doi.org/10.1146/annurev-biodatasci-080917-013343, (2018).
- J239. C. H. Eng, T. W.H. Backman, C. B. Bailey, C. Magnan, H. G. Marin, L. Katz, **P. Baldi**, and J. D. Keasling. ClusterCAD: a Computational Platform for Type I Modular Polyketide Synthase Design. Nucleic Acids Research, 46, Issue D1, 4 January 2018, Pages D509–D515, https://doi.org/10.1093/nar/gkx893 (2018).
- J238. Chase Shimmin, Peter Sadowski, **Pierre Baldi**, Edison Weik, Daniel Whiteson, Edward Goul, and Andreas Søgaard. Decorrelated Jet Substructure Tagging Using Adversarial Neural Networks. *Physical Review D*, 96, 074034 (2017) - Published 30 October 2017.
- J237. **P. Baldi**, P. Sadowski, and Z. Lu. Learning in the Machine: the Symmetries of the Deep Learning Channel. *Neural Networks*, 95, November 2017, 110-133, (2017).
- J236. M. E. Hughes, **P. Baldi**, J. B Hogenesch et al. Guidelines for genome-scale analysis of Biological rhythms. *Journal of Biological Rhythms*, 32, 5, 380-393, (2017). First Published November 3, (2017).
- J235. P. Sadowski, B. Radics, A. Ananya, Y. Yamazaki, and **P. Baldi**. Efficient Antihydrogen Detection in Antimatter Physics by Deep Learning. *Journal of Physics Communications*, 1, 2, Published 6 September 2017, (2017).
- J234. P. Tognini, M.Murakami, Y. Liu, K. L.Eckel-Mahan, J.C.Newmand, E.Verdind, **P.Baldi**, and P.Sassone-Corsi. Distinct Circadian Signatures in Liver and Gut Clocks Revealed by Ketogenic Diet . *Cell Metabolism*, 2017 Sep 5;26, 3, 523-538.e5. doi: 10.1016/j.cmet.2017.08.015, (2017).
- J233. **P. Baldi**. The Inner and Outer Approaches for the Design of Recursive Neural Networks Architectures. *Data Mining and Knowledge Discovery*, 1-13, (2017). DOI: 10.1007/s10618-017-0531-0. Available at: http://link.springer.com/article/10.1007/s10618-017-0531-0.

- J232. R. Mazziotti, L. Baroncelli, N. Ceglia, G. Chelini, G. Della Sala, C. Magnan, D. Napoli, E. Putignano, D. Silingardi, J. Tola, P. Tognini, **P. Baldi**, T. Pizzorusso. Mir-132/212 is Required for Maturation of Binocular Matching of Orientation Preference and Depth Perception. *Nature Communications*, published 23 May 2017, DOI: 10.1038/ncomms15488, (2017).
- J231. J. Wang, Z. Fang, N. Lang, H. Yuan, M. Su., and **P. Baldi**. A Multi-Resolution Approach for Spinal Metastasis Detection using Deep Siamese Neural Networks. *Computers in Biology and Medicine*, Volume 84, 1 May 2017, Pages 137-146, DOI information: 10.1016/j.compbiomed.2017.03.024, (2017).
- J230. Annie Vogel Ciernia1, Enikö A Kramár, Dina P Matheos, Robbert Havekes, Thekla J Hemstedt, Christophe N. Magnan, Keith Sakata, Ashley Tran, Soraya Azzawi, Alberto Lopez, Richard Dang, Weisheng Wang, Brian Trieu, Joyce Tong, Ruth M Barrett, Rebecca J Post, **Pierre Baldi**, Ted Abel, Gary Lynch & Marcelo A. Wood. Mutation of neuron-specific chromatin remodeling subunit BAF53b: Rescue of plasticity and memory by manipulating actin remodeling. *Learning and Memory*, 24, 199–209, (2017).
- J229. Y. Liu, S. Sun, T. Bredy, M. Wood. R. C. Spitale, and **P. Baldi**. MotifMap-RNA: A Genome-Wide Map of RBP Binding Sites. *Bioinformatics*, Volume 33, Issue 13, 1 July 2017, Pages 2029–2031, https://doi.org/10.1093/bioinformatics/btx087 (2017).
- J228. Rafael M. Ioris,Mirco Galié, Giorgio Ramadori, Jason G. Anderson, Anne Charollais, Georgia Konstantinidou, Xavier Brenachot, Ebru Aras, Algera Goga, Nicholas Ceglia, Carlos Sebastián, Denis Martinvalet, Raul Mostoslavsky, **Pierre Baldi**, and Roberto Coppari. SIRT6 Suppresses Cancer Stem-like Capacity in Tumors with PI3K Activation Independently of Its Deacetylase Activity. *Cell Reports*, 18, 8, 1858–1868, (2017).
- J227. J. Wang, H. Ding, F. Azamian, B. Zhou, C. Iribarren, S. Molloi, and **P. Baldi**. Detecting Cardiovascular Disease from Mammograms with Deep Learning. IEEE Transactions on Biomedical Imaging, 36, 5, 1172-1181, (2017). Print ISSN: 0278-0062, Online ISSN: 1558-254X, Digital Object Identifier: 10.1109/TMI.2017.2655486.
- J226. D. Guest, J. Collado, **P. Baldi**, S. Hsu, G. Urban, D. Whiteson. Jet Flavor Classification in High-Energy Physics with Deep Neural Networks. *Physical Review D*, 94, 112002 (2016) - Published 2 December 2016, (2016).
- J225. P. Sadowski, D. Fooshee, N. Subrahmanya, and **P. Baldi**. On the Synergies Between Quantum Mechanics and Machine Learning in Reaction Prediction. *Journal of Chemical Information and Modeling*, 56 (11), 2125–2128. Publication Date (Web): October 17, 2016 (Letter). DOI: 10.1021/acs.jcim.6b00351, (2016).
- J224. C. Magnan, J. Yu, I. Chang, E. Jahn, Y. Kanomata, J. Wu, M. Zeller, M. Oakes, **P. Baldi**, and S. Sandmeyer. Sequence Assembly of *Yarrowia lipolytica* Strain W29/CLIB89 Shows Transposonable Element Diversity. *PLoS ONE*, 11, 9: e0162363. doi:10.1371/journal.pone.0162363, (2016).
- J223. **P. Baldi** and P. Sadowski. A Theory of Local Learning, the Learning Channel, and the Optimality of Backpropagation. Neural Networks, 83, 51–74, (2016). Available online: 5-AUG-2016, DOI information: 10.1016/j.neunet.2016.07.006, (2016).
- J222. M. Murakami, P. Tognini, Y. Liu, K.Eckel-Mahan, **P. Baldi**, and P. Sassone-Corsi. Gut microbiota directs PPARγ driven reprogramming of the liver circadian clock by nutritional challenge. *EMBO Reports*, DOI 10.15252/embr.201642463| Published online 14.07.2016, (2016).
- J221. F. Agostinelli, N. Ceglia, B. Shahbaba, P. Sassone-Corsi, and P. Baldi. What Time is it? Deep Learning Approaches for Circadian Rhythms. Bioinformatics, 32 (12): i8-i17, (2016) . doi:10.1093/bioinformatics/btw243.

- J220. **P. Baldi**, K. Bauer, C. Eng, P. Sadowski, and D. Whiteson. Jet Substructure Classification in High-Energy Physics with Deep Neural Networks. *Physical Review D*, 93, 094034 (2016) - Published 27 May 2016, (2016).
- J219. S. Masri, T. Papagiannakopoulos, K. Kinouchi, Y. Liu, M. Cervantes, **P. Baldi**, T. Jacks, and P. Sassone-Corsi. Lung Adenocarcinoma Distally Rewires Hepatic Circadian Homeostasis. *Cell*, 165, 4, 896—909, (2016).
- J218. S. L. Borrego, J. Fahrmann, R. Datta, C. Stringari, D. Grapov, M. Zeller, Y. Chen, P. Wang, P. Baldi, E. Gratton, O. Fiehn, and P. Kaiser. Metabolic Changes Associated with Methionine Stress Sensitivity in MDA-MB-468 Breast Cancer Cells. Cancer and Metabolism, 4:9, Published on: 2 May 2016, DOI 10.1186/s40170-016-0148-6, (2016).
- J217. P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, D. Whiteson. Parameterized Neural Networks for High-Energy Physics. *The European Physical Journal C*, 76, 235, 1-7, (2016). DOI 10.1140/epjc/s10052-016-4099-4.
- J216. C. Galiez, C. Magnan, F. Coste, and **P. Baldi**. VIRALpro: a tool to identify viral capsid and tail sequences. *Bioinformatics,* 32 (9): 1405-1407. First published online January 5, 2016, doi:10.1093/bioinformatics/btv727, (2016).
- J215. A. Lusci, M. Browning, D.Fooshee, S. Joshua Swamidass, and **P. Baldi**. Accurate and Efficient Target Prediction Using a Potency-Sensitive Influence-Relevance Voter. *Journal of Cheminformatics*, 7, 63-79, DOI: 10.1186/s13321-015-0110-6, http://www.jcheminf.com/content/7/1/63, (2015).
- J214. S. Abbondante, K. Eckel Mahan, N. Ceglia, **P. Baldi**, and P. Sassone-Corsi. Comparative Circadian Metabolomics Reveal Differential Effects of Nutritional Challenge in the Serum and Liver. *Journal of Biological Chemistry*, 291, 6, 2812-.-2828, doi:10.1074/jbc.M115.681130, (2016).
- J213. D. R. Fooshee, P. K. Aiona, A. Laskin, J. Laskin, S. A. Nizkorodov, and **P. Baldi**. Atmospheric Oxidation of Squalene: Molecular Study Using COBRA Modeling and High-Resolution Mass Spectrometry. *Environmental Sciences and Technology*, 49 (22), pp 13304–13313, (2015). Publication Date (Web): October 22, 2015 le). DOI: 10.1021/acs.est.5b03552.
- J212. S. Lane, P. Di Lena, K. Tormanen, **P. Baldi**,and H. Liu. Function and regulation of Cph2 in Candida albicans. *Eukaryotic Cell*, 14, 11, 1114-1126, (2015).
- J211. E. A. Mayer, J. Labus, K. Tillisch, S. Cole, and **P. Baldi**. Towards a Systems View of Irritable Bowel Syndrome. *Nature Reviews Gastroenterolog and Hepatology,* 12, 592–605, (2015).
- J210. V. Patel, N. Ceglia, M. Zeller, K. Eckel-Mahan, P. Sassone-Corsi, and **P. Baldi**. The Pervasiveness and Plasticity of Circadian Oscillations: The Coupled Circadian-Oscillators Framework. *Bioinformatics*, 31, (19), 3181-3188, [first published online June 6, 2015 doi:10.1093/bioinformatics/btv353, (2015).
- J209. A. Sequeira, B. Rollins, C. Magnan, M. van Oven, **P. Baldi**, R. M. Myers, J. David Barchas, A. F. Schatzberg, S. J.Watson, H. Akil, W. E. Bunney, M. P. Vawter. Mitochondrial Mutations in Subjects with Psychiatric Disorders. *PLOS ONE*, Published: May 26, 2015, https://doi.org/10.1371/journal.pone.012728, (2015). PMCID: PMC4444211.
- J208. **P. Baldi**, P. Sadowski, and D. Whiteson. Enhanced Higgs to $\tau^+\tau^-$ Search with Deep Learning. *Physical Review Letters*, 114, 111801-1—111801-5, (2015).
- J207. P. Sadowski, J. Collado, D. Whiteson, and **P. Baldi**. Deep Learning, Dark Knowledge, and Dark Matter. Journal of Machine Learning Research, Workshop and Conference Proceedings, Proceedings of NIPS Workshop on High-Energy Physics, Volume 42, 81-97, (2015).
- J206. N. P. Hoverter, M. Zeller, M. McQuade, A. Garibaldi, A. Busch, K. Hertel, **P. Baldi**, and M. L. Waterman. The TCF C-clamp DNA Binding Domain Expands the Wnt

Transcriptome via Alternative Target Recognition. *Nucleic Acids Research*, 42, 22, 13615-13632, (2014). PMCID: PMC4267635.

- J205. W. Gordon, M. Zeller, R. Herndon, W. Swindell, H. Ho, F. Espetia, J.E. Gudjonsson, **P. Baldi**, and B. Andersen. A GRHL3-regulated repair pathway suppresses immune-mediated epidermal hyperplasia. *Journal of Clinical Inverstigation*, 124, 12, 5205-5218, (2014). PMCID: PMC4348962.
- J204. M. Zeller, C. Magnan, V. R. Patel, P. Rigor. L. Sender, and **P. Baldi**. A Genomic Analysis Pipeline and Its Application to Pediatric Cancers. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 11, 5, 826-839, (2014).
- J203. C. Magnan and **P. Baldi**. SSpro/ACCpro 5.0: Almost Perfect Prediction of Protein Secondary Structure and Relative Solvent Accessibility. Problem Solved? *Bioinformatics*, 30, 18,: 2592—2597, (2014). First published online May 24, 2014 doi:10.1093/bioinformatics/btu352. PMCID: PMC4215083.
- J202. K. Nagata, A. Randall, and **P. Baldi**. Incorporating Post-Translational Modifications and Unnatural Amino Acids into High-Throughput Modeling of Protein Structures. B*ioinformatics*, 30, 12, 1681-1689 (first published online February 25, 2014) (2014). PMCID: PMC4058938.
- J201. S. Sahar, S. Masubuchi, K. Eckel-Mahan, S. Vollmer, L. Galla, N. Ceglia, S. Selma Masri, T. K. Barth, B. Grimaldi, O. Oluyemi, G. Astarita, W. C. Hallows, D. Piomelli, A. Imhof, **P. Baldi**, J. M. Denu, and P. Sassone-Corsi. Circadian Control of Fatty Acid Elongation by SIRT1-mediated Deacetylation of Acetyl-CoA Synthetase 1. *Journal of Biological Chemistry*, 289, 6091-6097, (2014). PMCID: PMC3937675.
- J200. **P. Baldi**, P. Sadowski, and D. Whiteson. Deep Learning in High-Energy Physics: Improving the Search for Exotic Particles. *Nature Communications*, 5, article 4308, (2014).
- J199. S. Masri, P. Rigor, M. Cervantes, N. Ceglia, C. Sebastian, C. Xiao, M. Roqueta-Rivera, C. Deng, T. F. Osborne, R. Mostoslavsky, **P. Baldi**, and P. Sassone-Corsi. Partitioning Circadian Transcription by SIRT6 Leads to Segregated Control of Cellular Metabolism. *Cell*, 158, 3, 659—672, (2014).
- J198. V. R. Patel, K. Eckel-Mahan, P. Sassone-Corsi, and **P. Baldi**. How Pervasive Are Circadian Oscillations? Trends in Cell Biology, 24, 329-331, (2014).
- J197. **P. Baldi** and P. Sadowski. The Dropout Learning Algorithm. *Artificial Intelligence*, 210C, 78-122, (2014). PMCID: PMC3996711
- J196. K. A. Dyar, S. Ciciliot, L. E. Wright, R. S. Biensø, G. Malagoli Tagliazucchi, V. R. Patel, M. Forcato, M. I. Peña Paz, A. Gudiksen, F. Solagna, M. Albiero, I. Moretti, K. L. Eckel-Mahan, **P. Baldi**, P. Sassone-Corsi, R. Rizzuto, S. Bicciato, H. Pilegaard, B. Blaauw, and S. Schiaffino. Muscle Insulin Sensitivity and Glucose Metabolism are Controlled by the Intrinsic Muscle Clock. *Molecular Metabolism,* 3, 1, 29-41, (2014). PMCID: PMC4264038.
- J195. P. Sadowski and **P. Baldi**. Small-molecule 3D Structure Prediction Using Open Crystallography Data. *Journal of Chemical Information and Modeling*, 53, 12, 3127–3130, (2013). PMCID: PMC3890693.
- J194. K. L. Eckel-Mahan, V. R. Patel, S. de Mateo, N. J. Ceglia, S. Sahar, S. Dilag, K. A. Dyar, R. Orozco-Solis, **P. Baldi**, and Paolo Sassone-Corsi. Reprogramming of the Circadian Clock by Nutritional Challenge. *Cell*, 155, 7, 1464-1478, (2013).
- J193. D. Fooshee, A. Andronico, and **P. Baldi**. ReactionMap: An Efficient Atom-Mapping Algorithm for Chemical Reactions. *Journal of Chemical Information and Modeling*, 53 (11), 2812–2819, (2013). PMID:24160861.
- J192. L. Aguilar-Arnal, O. Hakim, V. R. Patel, **P. Baldi**, G. L. Hager, and P. Sassone-Corsi. Cycles in Spatial and Temporal Chromosomal Organization Driven by the Circadian Clock. *Nature Structural and Molecular Biology*, 20, 1206-1213, (2013).

- J191. T. Fujikawa, E. D. Berglund, V. R. Patel, G. Ramadori, C. R. Vianna, L. Vong, F. Thorel, S. Chera, P. L. Herrera, B. B. Lowell, J. K. Elmquist, **P. Baldi**, and R. Coppari. Leptin Engages a Hypothalamic Neurocircuitry to Permit Survival in the Absence of Insulin. *Cell Metabolism*, 18, 3, 431–444, (2013).
- J190. A. Lusci, G. Pollastri, and **P. Baldi**. Deep Architectures and Deep Learning in Chemoinformatics: the Prediction of Aqueous Solubility for Drug-Like Molecules. *Journal of Chemical Information and Modeling*, 53, 7, 1563–1575, (2013). PMID: 23795551 PMCID: PMC3739985.
- J189. V. A. Feher, A. Randall, **P. Baldi**, R. M. Bush, L. M de la Maza, and R. E. Amaro. A 3-dimensional trimeric β-barrel model for Chlamydia MOMP contains conserved and novel elements of gram-negative bacterial porins. *PLOS One*, published 25 Jul 2013, DOI: 10.1371/journal.pone.0068934, (2013). PMCID: PMC3723809.
- J188. E. A. Gordon, T. Whisenant, M. Zeller, R. Kaake, W. Gordon, P. Krotee, V. Patel, L. Huang, **P. Baldi**, and L. Bardwell. Combining Docking Site and Phosphosite Predictions to Find New Substrates: Identification of Smoothelin-like-2 (SMTNL2) as a c-Jun N-terminal Kinase (JNK) Substrate. *Cellular Signalling*, 25, 12, 2518-2529, (2013). PMCID: PMC4132694.
- J187. M. M. Bellet, E. Deriu, J. Liu, B. Grimaldi, C. Blaschitz, M. Zeller, R. A. Edwards, S. Sahar, S.Dandekar, **P. Baldi**, M. D.George, M. Raffatellu, and P. Sassone-Corsi. The Circadian Clock Regulates the Host Response to *Salmonella*. *Proceedings of the National Academy of Sciences USA*, 110, 24, 9897-9902, (2013). PMID: 23716692. PMCID: PMC3683799.
- J186. I. Chang and **P. Baldi**. A Unifying Kinetic Framework for Modeling Oxidoreductase Catalyzed Reactions. *Bioinformatic*s, 29, 10, 1299-1307, (2013). PMID: 23613486. PMCID: PMC3732027.
- J185. A. Vogel-Ciernia, D. P. Matheos, R. M. Barrett, E. Kramár, Y. Chen, C. Magnan, M. Zeller, A. Sylvain, S. Azzawi, J. Haettig, Y. Jia, A. Tran, R. J. Post, G. R. Crabtree, T. Z. Baram, **P. Baldi**, G. Lynch, and M. A. Wood. Neuron-Specific Nucleosome Remodeling Component BAF53b is Necessary for Synaptic Plasticity and Memory. *Nature Neuroscience*, 16, 552-561, (2013). PMCID: PMC3777648.
- J184. S. Masri, V. R. Patel, K. Eckel-Mahan, S. Peleg, I. Forne, A. Ladurner, **P. Baldi**, A. Imhof, and P. Sassone-Corsi. The Circadian Acetylome Reveals Regulation of Mitochondrial Metabolic Pathways. *Proceedings of the National Academy of Sciences USA*, 110, 9, 3339-3344, (2013). doi:10.1073/pnas.1217632110 . PMCID: PMC3587221.
- J183. S. Lee, H. Yu, A. Randall, Z. Zeng, **P. Baldi**, R. Doong, and K. J. Shea. Engineered Synthetic Polymer Nanoparticles as IgG Affinity Ligands. *Journal of the Americal Chemical Society*, 134, 38, 15765–15772 , (2012). PMID:22924890. PMCID: PMC3482410.
- J182. M. Kayala and **P. Baldi**. ReactionPredictor: Prediction of Complex Chemical Reactions at the Mechanistic Level Using Machine Learning. *Journal of Chemical Information and Modeling*, 52, 10, 2526–2540, (2012). PMID: 22978639.
- J181. P. Di Lena, K. Nagata, and **P. Baldi**. Deep Architectures for Protein Contact Map Prediction. *Bioinformatics*, 28, 2449-2457, (2012). PMID: 22847931. PMCID: PMC3463120.
- J180. N. Hoverter, J. H. Ting, S. Sundaresh, **P. Baldi**, and M. Waterman. A WNT/p21 circuit directed by the C-clamp, a sequence-specific DNA binding domain in TCFs. *Molecular and Cellular Biology*, 32, 3648-3662, (2012). PMID: 22778133. PMCID: PMC34301983.
- J179. V. Patel, K. Eckel Mahan, P. Sassone-Corsi, and **P. Baldi**. CircadiOmics: Integrating Circadian Genomics, Transcriptomics, Proteomics, and Metabolomics. *Nature Methods*, 9, 8, 772-773, (2012).

- J178. **P. Baldi**. Boolean Autoencoders and Hypercube Clustering Complexity. *Designs, Codes and Cryptography*, 65, 3, 383-403, (2012).
- J177 A. S. Hopkin, W. Gordon, R. H. Klein, F. Espitia, K. Daily, M. Zeller, **P. Baldi**, and B. Andersen. GRHL3/GET1 and Trithorax Group Members Collaborate to Activate the Epidermal Progenitor Differentiation Program. *PLoS Genetics*, published 19 Jul 2012, 10.1371/journal.pgen.1002829. PMID: 22829784. PMCID: PMC3400561.
- J176. A. Sequeira, M. V. Martin, B. Rollins, E. A. Moon, W. E. Bunney, F. Macciardi, S. Lupoli, E. Smith, J. Kelsoe, C. Magnan, M. van Oven, **P. Baldi**, D.C. Wallace, M. P. Vawter. Mitochondrial Mutations and Polymorphisms in Psychiatric Disorders. *Frontiers in Behavioral and Psychiatric Genetics*, 20 June 2012, doi: 10.3389/fgene.2012.00103, (2012). PMCID: PMC3379031.
- J175. D. R. Fooshee, T. B. Nguyen, S. A. Nizkorodov, J. Laskin, A. Laskin, and **P. Baldi**. COBRA: A Computational Brewing Application for Predicting the Molecular Composition of Organic Aerosols. *Environmental Science & Technology*, 46, 11, 6048–6055, (2012). PMCID: PMC3385869.
- J174. A. Kao, A. Randall, Y. Yang, V. Patel, W. Kandur, S. Guan, S. Rychnovsky, **P. Baldi**, and L.Huang. Mapping the Structural Topology of the Yeast 19S Proteasomal Regulatory Particle Using Chemical Cross-linking and Probabilistic Modeling. *Molecular and Cellular Proteomics*, 11,12 1566-1577. First Published on April 30, 2012, doi:10.1074/mcp.M112.018374, (2012). PMCID: PMC3518129.
- J173. **P. Baldi** and Z. Lu. Complex-Valued Autoencoders. *Neural Networks*, 33, 136-147 (2012). PMID: 22622264. PMCID: PMC3399055.
- J172. M. Kayala, and **P. Baldi**. Cyber-T Webserver: Differential Analysis of High-Throughput Data. *Nucleic Acids Research*, Web Server Issue, first published online May 7, 2012, 40, 1-7,; doi: 10.1093/nar/gks420, (2012). PMCID: PMC3394347.
- J171. L. Fang, R. M. Kaake, V. Patel, **P. Baldi**, and L. Huang. Mapping the Protein Interaction Network of the Human COP9 Signalosome (CSN) Complex Using a Label-Free QTAX Strategy. *Molecular and Cellular Proteomics*. First Published on April 3, 2012, 11, 5, 138-147,; doi:10.1074/mcp.M111.016352, (2012). PMCID: PMC3418852.
- J170. R. Nasr. R. Vernica, C. Li, and **P. Baldi**. Speeding Up Chemical Searches Using the Inverted Index: the Convergence of Chemoinformatics and Text Search Methods. *Journal of Chemical Information and Modeling*, 52, 4, 891—900, (2012). PMCID: PMC3415597.
- J169. **P. Baldi** and C. Lopes. The Universal Campus: An Open 3D Virtual World Infrastructure for Research and Education. *ACM eLearn Magazine*, April 26, (2012).
- J168. K. L. Eckel-Mahan, V. R. Patel, K. S.Vignola, R. P. Mohney, **P. Baldi**, and P. Sassone-Corsi. Coordination of Metabolome and Transcriptome by the Circadian Clock. *PNAS*, 109 (14) 5541-5546; published ahead of print March 19, 2012, doi:10.1073/pnas.1118726109, (2012). PMCID: PMC3325727.
- J167. D. Ren, M. Dalmau, A. Randall, M. M. Shindel, **P. Baldi**, and S. Wang. Biomimetic Design of Protein Nanomaterials for Hydrophobic Molecular Transport. *Advanced Functional Materials*, Article first published online: 23 APR 2012, DOI: 10.1002/adfm.201200052, (2012). PMCID: PMC3603581.
- J166. X. Qi, K. Daily, K. Nguyen, H. Wang, D. Mayhew, P. Rigor, S. Forouzan, M. Johnston, R. D. Mitra, **P. Baldi**, S. B. Sandmeyer. Retrotransposons Profiling of RNA Polymerase III Initiation Sites. *Genome Research*, April 2012 22: 681-692; Published in Advance January 27, 2012, doi:10.1101/gr.131219.111, (2012). PMCID: PMC3317150.
- J165. K. Nagata, A. Randall, and **P. Baldi**. SIDEpro: a Novel Machine Learning Approach for the Fast and Accurate Prediction of Side-Chain Conformations. *Protein: Structure, Function, and Bioinformatics*, 80, 1, 142–153, (2012). PMCID: PMC3240718.

- J164. **P. Baldi**. Autoencoders, Unsupervised Learning, and Deep Architectures. *Journal of Machine Learning Research*, Workshop and Conference Proceedings, Proceedings of ICML Workshop on Unsupervised and Transfer Learning, Volume 27, 37-50, (2012).
- J163. S. Srikrishnan, A. Z. Randall, **P. Baldi**, and N.A. Da Silva. Rationally Selected Single-Site Mutants of the *Thermoascus aurantiacus* Endoglucanase Increase Hydrolytic Activity on Cellulosic Substrates. *Biotechnology & Bioengineering*, Volume 109, Issue 6, June 2012, Pages: 1595–1599. Article first published online: 2 JAN 2012, DOI: 10.1002/bit.24414, (2012).
- J162. K. Daily, V. Patel, P. Rigor, X. Xie, and **P. Baldi**. MotifMap: Integrative Genome-Wide Maps of Candidate Regulatory Motif Sites for Model Species. *BMC Bioinformatics*, 12:495 (30 December 2011), (2011). PMCID: PMC3293935.
- J161. **P. Baldi**. Data-Driven High-Throughput Prediction of the 3-D Structure of Small Molecules: Review and Progress. A Response to the Letter by the Cambridge Crystallographic Data Center. *Journal of Chemical Information and Modeling*, 51, 12, 3029–3029, (2011).
- J160. A. E. Barry, A. Trieu, F. J. I. Fowkes, J. Pablo,M. Kalantari-Dehaghi, A. Jasinskas, X. Tan, M.A. Kayala, L. Tavul, P.M. Siba, K. P. Day, **P. Baldi**, P. L. Felgner, D. L. Doolan. The Stability and Complexity of Antibody Responses to the Major Surface Antigen of *Plasmodium falciparum* are Associated with Age in a Malaria Endemic Area. *Molecular and Cellular Proteomics*, First Published on August 8, 2011, doi:10.1074/mcp.M111.008326, (2011).
- J159. M. Kayala, C. Azencott, J. Chen, and **P. Baldi**. Learning to Predict Chemical Reactions. *Journal of Chemical Information and Modeling*, 51, 9, 2209–2222, (2011). PMCID: PMC3193800.
- J158. R. Nasr, T. Kristensen, and **P. Baldi**. Tree and Hashing Data Structures to Speedup Chemical Searches: Analysis and Experiments. *Molecular Informatics*. Special Issue on Machine Learning Methods in Chemoinformatics/NIPS, 30, 9, 791-800, (2011).
- J157. B. Tang, P. Di Lena, L. Schaffer, S. R. Head, **P. Baldi**, and E. A. Thomas. Genome-wide Identification of Bcl11b Gene Targets Reveals Role in Brain-Derived-Neurotrophic-Factor Signaling. *PLoS ONE*, 6(9): e23691. doi:10.1371/journal.pone.0023691, (2011). PMCID: PMC3164671.
- J156. S. Boehnke, D. Berg, R. Marino, **P. Baldi**, L. Itti, and D. Munoz. Visual Adaptation and Novelty Responses in the Superior Colliculus. *European Journal of Neuroscience*, 34, 5 766-779, (2011). PMCID: PMC3168683.
- J155. A. Trieu, M. Kayala, C. Burk, D. M. Molina, D. A. Freilich, T. L. Richie, **P. Baldi**, P. L. Felgner, D. L. Doolan. Sterile Protective Immunity to Malaria is Associated with a Panel of Novel Plasmodium falciparum Antigens. *Molecular and Cellular Proteomics*. First Published on May 31, 2011, 10,9,; doi:10.1074/mcp.M111.007948, (2011).
- J154. A. Andronico, A. Randall, R. W. Benz, and **P. Baldi**. Data-Driven High-Throughput Prediction of the 3D Structure of Small Molecules: Review and Progress. *Journal of Chemical Information and Modeling*, 51, 4, 760–776, (2011).
- J153. K. Christiansen, L. Larsen, M. Zhang, Y. Kuznetsov, V. Bilanchone, N.Beliakova-Bethell1, A. Randall, R. DaSilva, K. Nagashima, A. McPherson, **P. Baldi**, and S. Sandmeyer. The Ty3 GAG3 Spacer Controls Intracelluar Condenstation and Uncoating. *Journal of Virology*, published ahead of print on 26 Jan 2011, 85, 7, 3055-3066,; doi: doi:10.1128/JVI.01055-10, (2011).
- J152. M. I. Cruz-Fisher, C. Cheng, G. Sun, S. Pal, A. Teng, D. M. Molina, M. A. Kayala, **P. Baldi**, P. L. Felgner, X. Liang, L. M. de la Maza. Identification of Immunodominant Antigens by Probing a whole *Chlamydia trachomatis* Open Reading Frame Proteome

Microarray Using Sera from Immunized Mice, *Infection and Immunity*, 79, 1, 246—257, (2011).

- J151. K. Daily, P. Rigor, S. Christley, X. Xie, and **P. Baldi**. Data Structures and Compression Algorithms for High- Throughput Sequencing Technologies. *BMC Bioinformatics*, 11, 514-526, (2010). PMCID: PMC2964686.
- J150. A. Kao, C. Chiu, D. Vellucci, Y. Yang, V. R. Patel, S. Guan, A. Randall, **P. Baldi**, S. Rychnowski, L. Huang. Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Molecular and Cellular Proteomics*. First Published on August 24, 2010, 10, 1; doi:10.1074/mcp.M110.002212 (2011).
- J149. I. Chang, M. Heiske, T. Letellier, D. C. Wallace, and **P. Baldi**. Modeling of Mitochondrial Bioenergetics Based on a Composable Chemiosmotic Energy Transduction Rate Law: Theory and Experimental Validation. *PLoS ONE*, 6(9): e14820. doi:10.1371/journal.pone.0014820, (2011).
- J148. C. N. Magnan, M. Zeller, M. A. Kayala, A. Vigil, A. Randall, P. L. Felgner, and **P. Baldi**. High-throughput Prediction of Protein Antigenicity Using Protein Array Data. *Bioinformatics*, 26, 23, 2936-2943, (2010). PMCID: PMC2982151.
- J147. T. C. Whisenant, D. T. Ho, R. Benz, J. Rogers, R. Kaake, E. A. Gordon, L. Huang, **P. Baldi**, and L. Bardwell. Computational Prediction and Experimental Verification of New MAP Kinase Docking Sites and Substrates Including Gli transcription Factors. *PLoS Computational Biology*, published 26 August, (2010), 6, 8.
- J146. R. Nasr, D. Hirschberg, and **P. Baldi**. Hashing Algorithms and Data Structures for Rapid Searches of Fingerprint Vectors. *Journal of Chemical Information and Modeling*, 50, 8, 1358—1368, (2010). PMCID: PMC2926297.
- J145. **P. Baldi** and R. Nasr. When is Chemical Similarity Significant? The Statistical Distribution of Chemical Similarity Scores and Its Extreme Values. *Journal of Chemical Information and Modeling*, 50, 7, 1205—1222, (2010). PMCID: PMC2914517.
- J144. L. Liang, D. Leng, C. Burk, R. Nakajima-Sasaki, M. Kayala, V. L. Atluri, J. Pablo, B. Unal, T. A. Ficht, E. Gotuzzo, M. Saito, J. Morrow, X. Liang, **P. Baldi**, J. M. Vinetz, P. L. Felgner, and R. M. Tsolis. Large Scale Immune Profiling of Infected Humans and Goats Reveals Differential Recognition of *Brucella melitensis* Antigens. *PLoS Neglected Tropical Diseases*, 4, 5, e673, 1—11, (2010).
- J143. P. Crompton, M. Kayala, B. Traore, K. Kayentao, A. Ongoiba, G. Weiss, D. Molina, C. Burk, M. Waisberg, A. Jasinskas, X. Tan, S. Doumbo, D. Doumtabe, Y. Kone, D. Narum, X. Liang, O. Doumbo, L. Miller, D. Doolan, **P. Baldi**, P. Felgner, S. Pierce. A Prospective Analysis of the Antibody Response to *Plasmodium falciparum* Before and After a Malaria Season by Protein Microarray. *PNAS*, 107, 15, 6958—6963, (2010). PMCID: PMC2872454.
- J142. S. J. Swamidass, C. Azencott, K. Daily, and **P. Baldi**. A CROC Stronger than ROC: Measuring, Visualizing, and Optimizing Early Retrieval. *Bioinformatics*, 26, 10, 1348-1356, (2010). PMCID: PMC2865862.
- J141. S. Javanmardi, C. Lopes, and **P. Baldi**. Modeling User Reputation in Wikis. *Statistical Analysis and Data Mining*, 3, 2, 126-139, (2010).
- J140. A. B. Mochon1, J. Ye, M. A. Kayala, J. R. Wingard, C. J. Clancy, M. H. Nguyen, P. Felgner, **P. Baldi**, and H. Liu. Serological Profiling of a *Candida albicans* Protein Microarray Reveals Permanent Host-Pathogen Interplay & Stage-Specific Responses during Candidemia. *PLoS Pathogens*, 6, e1000827, (2010).
- J139. **P. Baldi** and L. Itti. Of Bits and Wows: A Bayesian Theory of Surprise with Applications to Attention. *Neural Networks*, 23,5, 649-666, (2010). PMCID: PMC2860069.

- J138. D. M. Molina, S. Pal, A. Teng, M. Kayala, **P. Baldi**, P. Felgner, X. Liang, L. M. de la Maza. Identification of Immunodominant Antigens of Chlamydia Trachomatis Using Proteome Microarrays. *Vaccine*, 28, 17, 3014-3024, (2010).
- J137. D. Guzman, A. Randall, **P. Baldi**, and Z. Guan. Computational and single-molecule force studies of a macro domain protein reveal a key molecular determinant for mechanical stability. *Proceedings of the National Academy of Sciences USA*, 107,5, 1989—1994, (2010).
- J136. J. Chen and **P. Baldi**. No Electron Left-Behind: a Rule-Based Expert System to Predict Chemical Reactions and Reaction Mechanisms. *Journal of Chemical Information and Modeling*, 49, 9, 2034-2043, (2009).
- J135. **P. Baldi** and D. Hirschberg. An Intersection Inequality Sharper than the Tanimoto Triangle Inequality for Efficiently Searching Large Databases. *Journal of Chemical Information and Modeling*, 49, 8, 1866-1870, (2009).
- J134. C. N. Magnan, A. Randall, and **P. Baldi**. SOLpro: Accurate Sequence-Based Prediction of Protein Solubility. *Bioinformatics*, 25, 17, 2200-2207, (2009).
- J133. R. J. Nasr, S. J. Swamidass, and **P. Baldi**. Large Scale Study of Multiple-Molecule Queries. *Journal of Chemoinformatics*, 1:7, 2009.
- J132. M. C. Brandon, D. C. Wallace, and **P. Baldi**. Data Structures and Compression Algorithms for Genomic Sequence Data. *Bioinformatics*, 25,14, 1731-1738, (2009).
- J131. P. Felgner, M. Kayala, A. Vigil, C. Burk, R. Nakajima-Sasaki, J. Pablo, D. Molina, S. Hirst, J. Chew, D. Wang, G. Tan, M. Duffield, R. Yang, J. Neel, N. Chantratita, G. Bancroft, G. Lertmemongkolchai, D. Davies, **P. Baldi**, S. Peacock, and R. Titball. A Burkholderia pseudomallei protein array reveals serodiagnostic and cross-reative antigens. *PNAS*, 106, 32, 13499-13504, (2009).
- J130 S. J. Swamidass, C. Azencott, H. Gramajo, S. Tsai, and **P. Baldi**. The Influence Relevance Voter: An Accurate and Interpretable Virtual High Throughput Screening Method. *Journal of Chemical Information and Modeling*, 49, 4, 756--766, (2009).
- J129 X. Xie, P. Rigor, and **P. Baldi**. MotifMap: a human genome-wide map of candidate regulatory motif sites. *Bioinformatics*, 25, 2, 167-174, (2009).
- J128 M. J. Sweredoski and **P. Baldi**. COBEpro: A Novel System for Predicting Continuous B-cell Epitopes. *Protein Engineering Design and Selection*, 22, 3, 113-120, (2009).
- J127 A. Randall and **P. Baldi**. SELECTpro: Effective Protein Model Selection Using a Structure-Based Energy Function Resistant to Blunders. BMC Structural Biology, 8:52, (2008).
- J126. L. Itti and **P. Baldi**. Bayesian Surprise Attracts Human Attention. *Vision Research*, 49, 10, 1295-1306, (2009).
- J125. E. Linstead, S. Bajracharya, T. Ngo, P. Rigor, C. Lopes, **P. Baldi.** Sourcerer: Mining and Searching Internet-Scale Software Repositories. *Journal of Datamining and Knowledge Discovery*, 18, 2, 300-336, (2009).
- J124. M. Brandon,E. Ruiz-pesini, D. Mishmar, V. Procaccio, M.T. Lott, K. C. Nguyen, S. Spolim, U. Patil, **P. Baldi**, and D. Wallace. MITOMASTER: A Bioinformatics Tool For the Analysis of Mitochondrial DNA Sequences. *Human Mutation*, Database Issue, 30, 1, 1-6, (2009).
- J123 R. Jurdak, **P. Baldi**, and C. Lopes. Software-Driven Sensor Networks for Short-Range Shallow Water Applications. *.Ad Hoc Networks Journal*, 7, 837-848, (2008).
- J122 J. Cheng, A. N. Tegge, and **P. Baldi**. Machine Learning Methods for Protein Structure Prediction. *IEEE Reviews in Biomedical Engineering*, 1, 1, 41—49, (2008).
- J121. J. Chen and **P. Baldi**. Synthesis Explorer: Organic Chemistry Tutorial System for Multi-Step Synthesis Design and Reaction Prediction. *Journal of Chemical Education*, 85, 12, 1699-1703, (2008).

- J120 L. Chang, S. Sundaresh, J. Elliott, P. A. Anton, **P. Baldi**, A. Licudine, M. Mayer, T. Vuong, M. Hirano, B. D. Naliboff, V. Z. Ameen, and E. A. Mayer, Dysregulation of the Hypothalamic-Pituitary-Adrenal (HPA) Axis in Irritable Bowel Syndrome. *Neurogastroenterology and Motility*, 21, 2, 149-159, (2008).
- J119 **P. Baldi** and R. W. Benz. BLASTing Small Molecules—Statistics and Extreme Statistics of Chemical Similarity Scores. *Bioinformatics*, 24(13): i357-i365, (2008).
- J118 M. J. Sweredoski and **P. Baldi**. PEPITO: Improved Discontinuous B-Cell Epitope Prediction Using Multiple Distance Thresholds and Half-Sphere Exposure. *Bioinformatics*, 24, 12, 1459-1460, (2008).
- J 117. D. L. Doolan, Y. Mu, B. Unal, S. Sundaresh, S. Hirst, C. Valdez, A. Randall, D. Molina, X. Liang, J. A. Oloo, P. L. Blair, J. C. Aguiar, **P. Baldi**, D. Huw Davies, and P. L. Felgner. Profiling Humoral Immune Responses to *P. falciparum* Infection with Protein Microarrays, *Proteomics*, 8, 22, 4680-4694, (2008).
- J 116. **P. Baldi**, D. S. Hirschberg and R. J. Nasr. Speeding Up Chemical Database Searches Using a Proximity Filter Based on the Logical Exclusive-OR. *Journal of Chemical Information and Modeling*, 48, 7, 1367-1378, (2008).
- J115. A. G. Barbour, A. Jasinskas, M. Kayala, D. Huw Davies, A. C. Steere, **P. Baldi**, and P. Felgner. A genome-wide proteome array reveals a limited set of immunogens in natural infections of humans and white-footed mice with *Borrelia bugdorferi*. *Infection and Immunity*, 76, 8, 3374-3389, (2008).
- J114. R. W. Benz, S. J. Swamidass, and **P. Baldi**. Discovery of Power-Laws in Chemical Space. *Journal of Chemical Information and Modeling*, 48, 6, 1138-1151, (2008).
- J113. L. Wu and **P. Baldi**. Learning to Play Go Using Recursive Neural Networks. *Neural Networks*, 21, 9, 1392-1400, (2008).
- J112. A. Randall, J. Cheng, M. Sweredoski, and **P. Baldi**. TMBpro: Secondary Structure, Beta-Contact, and Tertiary Structure Prediction of Transmembrane Beta-Barrel Proteins. *Bioinformatics*, 24, 4, 513-520, (2008).
- J111. **P. Baldi**., R. W. Benz, D. S. Hirschberg, and S. Joshua Swamidass. Lossless Compression of Chemical Fingerprints Using Integer Entropy Codes Improves Storage and Retrieval. *Journal of Chemical Information and Modeling*, 47, 6, 2098-2109, (2007).
- J110. S. Sundaresh; A. Randall; B. Unal; J. M. Petersen; J. T. Belisle; M. Gill Hartley; M. Duffield; R. W. Titball; D. Huw Davies; P. L. Felgner; and **P. Baldi**. From protein microarrays to diagnostic antigen discovery: a study of the pathogen *Francisella tularensis. Bioinformatics*, 23, i508-i518, (2007).
- J109. J. Chen, E. Linstead, S. J. Swamidass, D. Wang, and **P. Baldi**. ChemDB Update—Full-Text Search and Virtual Chemical Space. *Bioinformatics*, 23, 17, 2348-2351, (2007).
- J108. M. Tress, J. Cheng, **P. Baldi**, K. Joo, J. Lee, J. Seo, J. Lee, D. Baker, D. Chivian, D. Kim, I. Ezkurdia. Assessment of Predictions Submitted for the CASP7 Domain Prediction Category. *Proteins*, 69, S8, 137-151, (2007).
- J107. M. Sweredoski, K. Donovan, B. Nguyen, A. J. Shaka, and **P. Baldi**. Minimizing the Overlap Problem in Protein NMR: A Computational Framework for Precision Amino Acid Labeling. *Bioinformatics*, 23, 21, 2829-2835, (2007).
- J106. N. Komarova, L. Wu, and **P. Baldi**. The fixed-size Luria-Delbruck Model with a Nonzero Death Rate. *Mathematical Biosciences*, 210, 1, 253-290, (2007).
- J105. J. Cheng and **P. Baldi**. Improved Residue Contact Prediction Using Support Vector Machines and a Large Feature Set. *BMC Bioinformatics*, 8, 113-121, (2007).
- J104. J. E. Eyles, B. Unal, M. Gill Hartley, S. L. Newstead, H. Flick-Smith, J. L. Prior, P. C. F. Oyston, A. Randal, Y. Mu, S. Hirst, D. M. Molina, D. Huw Davies, T. Milne, K. F. Griffin, **P. Baldi**, R. W. Titball and P. L. Felgner. Immunodominant *Francisella tularensis* antigens identified using proteome microarray. *Proteomics*, 7, 13, 2172-2183, (2007).

- J103. W. Einhäuser, T. N. Mundhenk, **P. Baldi**, C. Koch, and L. Itti. A bottom-up model of spatial attention predicts human error patterns in rapid scene recognition. *Journal of Vision*, 7, 10, article 6, 1-13, (2007).
- J102. S. J. Swamidass and **P. Baldi**. A Mathematical Correction for Fingerprint Similarity Measures to Improve Chemical Retrieval. *Journal of Chemical Information and Modeling*, 47, 3, 952-964, (2007).
- J101. S. J. Swamidass and **P. Baldi**. Bounds and Algorithms for Fast Exact Searches of Chemical Fingerprints in Linear and Sub-Linear Time. *Journal of Chemical Information and Modeling*, 47, 2, 302-317, (2007).
- J100. A. Sadovsky, **P. Baldi**, and F. Wan. A Theoretical Study of the In Vivo Mechanical Properties of Angiosperm Roots: Constitutive Theories and Methods of Parameter Estimation. *Journal of Engineering Materials and Technology*, 129, 3, 483-487, (2007).
- J99. R. Jurdak, **P. Baldi**, and C. Lopes**.** Adaptive Low Power Listening for Wireless Sensor Network, *IEEE Transactions on Mobile Computing*, 6, 8, 988-1004, (2007).
- J98. E. Ruiz-Pesini, M. T. Lott, V. Procaccio, J. Poole, M. C. Brandon, D. Mishmar, C. Yi, J. Kreuziger, **P. Baldi**, and D. C. Wallace. An Enhanced MITOMAP with a Global mtDNA Mutational Phylogeny. *Nucleic Acids Research*, 35 (Database Issue), D823-D828, (2007).
- J97. L. Zicker-Larsen, M. Zhang, N. Beliakova-Bethell, V. Bilanchone, A. Lamsa, K. Nagashima, R. Najdi, K. Kosaka, V. Kovacevic, J. Cheng, **P. Baldi**, G. Wesley Hatfield, and S. Sandmeyer. Ty3 Capsid Mutations Reveal Early and Late Functions of the Amino-Terminal Domain. *The Journal of Virology*, 81, 13, 6957-6972, (2007).
- J96. N. Wang, **P. Baldi**, and B. Gaut. Phylogenetic Analysis, Genome Evolution, and the Rate of Gene Gain in the *Herpesviridae*. *Molecular Phylogenetics and Evolution*, 43, 3, 1066-1075, (2007).
- J95. C. Azencott, A. Ksikes, S. Joshua Swamidass, J. Chen, L. Ralaivola, and **P. Baldi**. One- to Four- Dimensional Kernels for Virtual Screening and the Prediction of Physical, Chemical, and Biological Properties. *Journal of Chemical Information and Modeling*, 47, 3, 965-974, (2007).
- J94. Y. Dou, K. Fox-Walsh, **P. Baldi**, and K. Hertel. Genomic Splice-Site Analysis Reveals Frequent Alternative Splicing Close to the Dominant Splice Site, *RNA*, 12, 12, 2047-2056, (2006).
- J93. S. Sundaresh, D. L. Doolan, S. Hirst, Y. Mu, B. Unal, D. H. Davies, P. L. Felgner, and **P. Baldi**. Identification of Humoral Immune Responses in Protein Microarrays using DNA Microarray Data Analysis Techniques. *Bioinformatics*, 22, 14, 1760-1766, (2006).
- J92. M. Brandon, **P. Baldi**, and D. C. Wallace. Mitochondrial Mutations in Cancer. *Oncogene*, 25, 34, 4647-4662, (2006).
- J91. J. Cheng and **P. Baldi**. A Machine Learning Information Retrieval Approach to Protein Fold Recognition. *Bioinformatics*, 22, 12, 1456-1463, (2006).
- J90. C. V. Lopes, A. Haghighat, A. Mandal, **P. Baldi**, and T. Givargis. Localization of Off-the-Shelf Mobile Devices Using Audible Sound: Architectures, Protocols and Performance Assessment. ACM SIGMOBILE Mobile Computing and Communications Review, 10, 2, 38-50, (2006).
- J89. C. Tagwerker, K. Flick1, M. Cui, C. Guerrero, Y. Dou, B. Auer, **P. Baldi**, L. Huang, and Peter Kaiser. A tandem-affinity tag for two-step purification under fully denaturing conditions: Application in ubiquitin profiling and protein complex identification combined with in vivo cross-linking. *Molecular and Cellular Proteomics*, 5, 4, 737-748. First Published on January 23, 2006, doi:10.1074/mcp.M500368-MCP200, (2006).
- J88. T. Lin, M. Melgar, S. J. Swamidass, J. Purdon, T. Tseng, G. Gago, D. Kurth, **P. Baldi**, H. Gramajo, and S. Tsai. Structure-Based Inhibitor Design of AccD5, an Essential acyl-CoA

Carboxylase Carboxyltransferase Domain of *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences USA*, 103, 9, 3072-3077, (2006).

- J87. G. Pollastri, A. Vullo, P. Frasconi, and **P. Baldi**. Modular DAG-RNN Architectures for Assembling Coarse Protein Structures. *Journal of Computational Biology*, 13, 3, 631-650, (2006).
- J86 J. Cheng, M. J. Sweredoski, and **P. Baldi**. DOMpro: Protein Domain Prediction Using Profiles, Secondary Structure, Relative Solvent Accessibility, and Recursive Neural Networks. *Data Mining and Knowledge Discovery*, 13, 1, 1-10, (2006).
- J85. J. Cheng, A. Randall, and **P. Baldi**. Prediction of Protein Stability Changes for Single Site Mutations Using Support Vector Machines. *Proteins*, 62, 4, 1125-1132, (2006).
- J84 E. T. Wang, G. Kodama, **P. Baldi**, and R. K. Moyzis. Global Landscape of Recent Inferred Darwinian Selection for *Homo Sapiens*. *Proceedings of the National Academy of Sciences USA*, 103, 135-140, (2006).
- J83. J. Cheng, H. Saigo, and **P. Baldi**. Large-Scale Prediction of Disulphide Bridges Using Kernel Methods, Two-Dimensional Recursive Neural Networks, and Weighted Graph Matching. *Proteins*, 62, 3, 617-629, (2006).
- J82. S. A. Danziger, S. J. Swamidass, J. Zeng, L. R. Dearth, Q. Lu, J. H. Chen, J. Cheng, V. P. Hoang, H. Saigo, R. Luo, **P. Baldi**, Rainer K. Brachmann, and Richard H. Lathrop. Functional Census of Mutation Sequence Spaces: The Example of p53 Cancer Rescue Mutants. *IEEE Transactions on Computational Biology and Bioinformatics*, 3, 2, 114-125, (2006).
- J81 K. L. Fox-Walsh, Y. Dou, B. J. Lam, S. Hung, **P. Baldi**, and Klemens J. Hertel. The Architecture of pre-mRNAs Affects Mechanisms of Splice-Site Pairing. *Proceedings of the National Academy of Sciences USA,* 102, 16176-16181, (2005).
- J80. J. Chen, S. J. Swamidass, Y. Dou, J. Bruand, and **P. Baldi**. ChemDB: A Public Database of Small Molecules and Related Chemoinformatics Resources. *Bioinformatics*, 21, 22, 4133-4139, (2005).
- J79. D. M. Chung, Y. Dou, **P. Baldi**, and J. S. Nowick. The Absence of Favorable Aromatic Interactions between Beta-Sheet Peptides. *Journal of the American Chemical Society*,127 (28), 9998-9999, (2005).
- J78. J. Cheng and **P. Baldi**. Three-Stage Prediction of Protein Beta-Sheets by Neural Networks, Alignments, and Graph Algorithms. Proceedings of the 2005 Conference on Intelligent Systems for Molecular Biology, ISMB 05. *Bioinformatics*, 21, Supplement 1, i75-i84, (2005).
- J77. S. J. Swamidass, J. Chen, P. Phung, J. Bruand, L. Ralaivola, and **P. Baldi**. Kernels for Small Molecules and the Prediction of Mutagenicity, Toxicity, and Anti-Cancer Activity. Proceedings of the 2005 Conference on Intelligent Systems for Molecular Biology, ISMB 05. *Bioinformatics*, 21, Supplement 1, i359-i368, (2005).
- J76. J. Cheng, M. J. Sweredoski, and **P. Baldi**. Accurate Prediction of Protein Disordered Regions by Mining Protein Structure Data. *Data Mining and Knowledge Discovery*, 11, 3, 213-222, (2005).
- J75. J. Cheng, A. Z. Randall, M. Sweredoski, and **P. Baldi**. SCRATCH: a Protein Structure and Structural Feature Prediction Server. *Nucleic Acids Research*, 33, 9, Web Server issue, W72-76, (2005).
- J74. **P. Baldi** and M. Rosen-Zvi. On the Relationship Between Deterministic and Probabilistic Directed Graphical Models: from Bayesian Networks to Recursive Neural Networks. *Neural Networks,* special issue on Neural Networks and Kernel Methods for Structured Domains, 18, 8, (2005).

- J73. L. Ralaivola, J. S. Swamidass, H. Saigo, and **P. Baldi**. Graph Kernels for Chemical Informatics. *Neural Networks,* special issue on Neural Networks and Kernel Methods for Structured Domains, 18, 8, 1093-1110, (2005).
- J72. B. Irwin, M. Aye, **P. Baldi**, N. Beliakova-Bethell, H. Cheng, Y. Dou, W. Liou, and S. Sandmeyer. Retroviruses and Yeast Retrotransposons Use Overlapping Sets of Host Genes. *Genome Research*, 15, 641-654, (2005).
- J71. K. A. Salmon, S. Hung, N. R. Steffen, R. Krupp, **P. Baldi**, G. Wesley Hatfield, and R. P. Gunsalus. Global Gene Expression Profiling in *Escherichia coli* K12: The Effects of Oxygen Availability and ArcA. *Journal of Biological Chemistry*, 280, 15, 15084-15096, (2005).
- J70. R. Jurdak, P. Baldi, and C. Videira Lopes U-MAC: A Proactive and Adaptive UWB Medium Access Control Protocol. *Wireless Communications and Mobile Computing Journal*, Special Issue on UWB Communications, 5, 5, 551-566, (2005).
- J69. S. Sundaresh, S. Hung, G. W. Hatfield, and **P. Baldi**. How Noisy and Replicable are DNA Microarray Data? *International Journal of Bioinformatics Research and Applications*, 1, 1, 31-50, (2005).
- J68. H. Davies, X. Liang, J. E. Hernandez, K. M. Romero, T. T. Nguyen, M. Kalantari-Dehaghi, S. Hirst, Y. Mu, A. Randall, **P. Baldi**, L. Villarreal, P. L. Felgner. Profiling the Humoral Response to Infection by Using Proteome Microarrays: High-Throughput Vaccine and Diagnostic Antigen Discovery. *Proceedings of the National Academy of Sciences USA,* 102, 3, 547-552, (2005).
- J67. J. Cheng, L. Scharenbroich, **P. Baldi**, and E. Mjolsness. Sigmoid: Towards a Generative, Scalable Software Infrastructure for Pathway Bioinformatics and Systems Biology. *IEEE Intelligent Systems*, 20, 3, 68-75, (2005).
- J66. M. Brandon, M. Lott, K. Nguyen, S. Spolim, S. Navathe, **P. Baldi**, and D. Wallace. MITOMAP: a human mitochondrial genome database - 2004 update. *Nucleic Acids Research*, 33 (Database issue), D611-613, (2005).
- J65. I. Sokolchik, T. Tanabe, **P. Baldi**, and J. Y. Sze. Polymodal sensory function of the C. elegans OCR-2 channel arises from distinct intrinsic determinants within the protein and is selectively conserved in human TRPV2. *Journal of Neuroscience*, 25, 4, 1015-1023, (2005).
- J64. S. Hampson, B. S. Gaut, and **P. Baldi**. Statistical Detection of Chromosomal Homology Using Shared-Gene Density Alone. *Bioinformatics*, 21, 8, 1339-1348, (2005).
- J63. R. Jurdak, C. Videira Lopes, and **P. Baldi**. A Survey, Classification, and Comparative Analysis of Medium Access Control Protocols for Ad Hoc Networks. *IEEE Communications Surveys and Tutorials*, 6, 1, 2-16, (2004).
- J62. A. Z. Randall, **P. Baldi**, and L. P. Villareal. Structural Proteomics of the Poxvirus Family. *Artificial Intelligence in Medicine,* special issue on "Data mining in Genomics and Proteomics", 31, 2, 105-115, (2004).
- J61. Y. Dou, P. Baisnee, G. Pollastri, Y. Pecout, J. Nowick, and **P. Baldi**. ICBS: A Database of Interactions Between Protein Chains Mediated by Beta-Sheet Formation. *Bioinformatics*, 20, 16, 2767-2777, (2004).
- J60. A. McLysaght, **P. Baldi**, and B. S. Gaut. Extensive Gene Gain Associated with Adaptive Evolution of Poxviruses. *Proceedings of the National Academy of Sciences USA*, 100, 26, 15655-15660, (2003).
- J59. K. Salmon, K. Mekjian, R. P. Gunsalus, S. Hung, **P. Baldi**, and G. Wesley Hatfield. Global Gene Expression Profiling in *Escherichia coli* K12: The Effects of Oxygen Availability and FNR. *Journal of Biological Chemistry*, 278, 32, 29837-29855, (2003).
- J58. S. Hampson, A. McLysaght, B. Gaut, and **P. Baldi**. LineUp: Statistical Detection of Chromosomal Homology with Applications to Plant Comparative Genomics. *Genome Research*, 13, 5, 999-1010, (2003).

- J57. G. A. Weiss, T. A. Roth, **P. Baldi**, and S. S. Sidhu Comprehensive Mutagenesis of the C-terminal Domain of the M13 Gene-3 Minor Coat Protein: The Requirements for Assembly into the Bacteriophage Particle. *Journal of Molecular Biology*, 332, 4, 777-782, (2003).
- J56. **P. Baldi** and G. Pollastri. The Principled Design of Large-Scale Recursive Neural Network Architectures—DAG-RNNs and the Protein Structure Prediction Problem. *Journal of Machine Learning Research*, 4, 575-602, (2003).
- J55. Y. Guermeur, G. Pollastri, A. Elisseeff, D. Zelus, H. Paugam-Moisy, and **P. Baldi.** Combining Protein Secondary Structure Prediction Models with Ensemble Methods of Optimal Complexity. *Neurocomputing*, Vol. 56C, 305-327, (2003) (partially presented at JOBIM 2001 conference).
- J54. **P. Baldi** and W. H. Batchelder. Bounds on Variances of Estimators for Multinomial Processing Tree Models. *Journal of Mathematical Psychology*, 47, 467-470, (2003).
- J53. **P. Baldi**, L. De Nardis, and M. G. Di Benedetto. Modeling and Optimization of UWB Communication Networks Through a Flexible Cost Function. *IEEE Journal on Selected Areas in Communications*, 20, 9, 1733-1744, (2002).
- J52. S. Hung, **P. Baldi**, and G. Wesley Hatfield. Global Gene Expression Profiling in *Escherichia coli* K12: The Effects of Leucine-responsive Regulatory Protein. *Journal of Biological Chemistry*, 277, 43, 40309-40323, (2002).
- J51. P. Baisnee, S. Hampson, and **P. Baldi**. Why Are Complementary DNA Strands Symmetric? *Bioinformatics*, 18, 1021-1033, (2002).
- J50. G. Wesley Hatfield, S. Hung, and **P. Baldi**. Differential Analysis of DNA Microarray Gene Expression Data. *Molecular Microbiology*, 47, 4, 871-877, (2003).
- J49. G. Pollastri and **P. Baldi**. Prediction of Contact Maps by GIOHMMs and Recurrent Neural Networks Using Lateral Propagation from All Four Cardinal Corners. Proceedings of the 2002 Conference on Intelligent Systems for Molecular Biology, ISMB 02. *Bioinformatics*, 18, Supplement 1, S62-S70, (2002). Nominated for Best Paper Award.
- J48. **P. Baldi** and G. Pollastri. A Machine Learning Strategy for Protein Analysis. *IEEE Intelligent Systems*. Special Issue on Intelligent Systems in Biology, 17, 2, 28-35, (2002).
- J47. S. Hampson, D. Kibler, and **P. Baldi.** Distribution Patterns of Over-Represented k-mers in Non-Coding Yeast DNA. *Bioinformatics*, 18, 4, 611-626, (2002).
- J46. G. Pollastri, **P. Baldi**, P. Fariselli, and R. Casadio. Prediction of Coordination Number and Relative Solvent Accessibility in Proteins. *Proteins*, 47, 142-153, (2002).
- J45. G. Pollastri, D. Przybylski, B. Rost, and **P. Baldi**. Improving the Prediction of Protein Secondary Structure in Three and Eight Classes Using Recurrent Neural Networks and Profiles. *Proteins*, 47, 2, 228-235, (2002).
- J44. G. Pollastri, **P. Baldi**, P. Fariselli, R. Casadio. Improved Prediction of the Number of Residue Contacts in Proteins by Recurrent Neural Networks. *Bioinformatics*, 17, S234-S242, (2001).
- J43. A. D. Long, H. J. Mangalam, B. Y. P. Chan, L. Tolleri, G. W. Hatfield, and **P. Baldi**. Improved Statistical Inference from DNA Microarray Data Using Analysis of Variance and a Bayesian Statistical Framework. *Journal of Biological Chemistry*, 276, 23, 19937-19944, (2001).
- J42. **P. Baldi** and A. D. Long. A Bayesian Framework for the Analysis of Microarray Expression Data: Regularized t-Test and Inference of Gene Changes. *Bioinformatic*s, 17, 6, 509-519, (2001).
- J41. P. Baisnée, **P. Baldi**, S. Brunak, and A. Gorm Pedersen. Flexibility of the Genetic Code with Respect to DNA Structure. *Bioinformatic*s, 17: 237-248, (2001).
- J40. **P. Baldi** and P. Baisnée. Sequence Analysis by Additive Scales: DNA Structure for Sequences and Repeats of All Lenghts. *Bioinformatic*s, Vol 16, 10, 865-889, (2000).

- J39. **P. Baldi**, S. Brunak, Y. Chauvin, and H. Nielsen. Assessing the Accuracy of Prediction Algorithms for Classification: An Overview. *Bioinformatics*, Vol. 16, 5, 412-424, (2000).
- J38. **P. Baldi**. On the Convergence of a Clustering Algorithm for Protein-Coding Regions in Microbial Genomes, *Bioinformatics*, Vol. 16, 4, 367-371, (2000).
- J37. **P. Baldi**. Probabilistic Graphical Models in Computational Molecular Biology, *Journal of the Italian Association for Artificial Intelligence (AIIA)*, Vol. 1, 8-12, (2000).
- J36. **P. Baldi**, S. Brunak, Y. Chauvin, and A. Gorm Pedersen. Structural Basis for Triplet Repeat Disorders: A Computational Analysis, *Bioinformatics*, Vol. 15, 11, 918-929, (1999).
- J35. **P. Baldi**, S. Brunak, P. Frasconi, G. Soda, and G. Pollastri. Exploiting the Past and the Future in Protein Secondary Structure Prediction, Bioinformatics, Vol. 15, 11, 937-946, (1999).
- J34. A. Gorm Pedersen, **P. Baldi**, S. Brunak, and Y. Chauvin. The Biology of Eukaryotic Promoter Prediction-a Review. Computers and Chemistry, Special Issue on Genome and Informatics, 23, 191-207, (1999).
- J33. A. Gorm Pedersen, **P. Baldi**, Y. Chauvin, and S. Brunak. DNA Structure in Human RNA Polymerase II Promoters, Journal of Molecular Biology, 281, 663-673, (1998).
- J32. **P. Baldi**, M. C. Vanier, and J. M. Bower. On the Use of Bayesian Methods for Evaluating Compartmental Neural Models, Journal of Computational Neuroscience, 5, 285-314, (1998).
- J31. **P. Baldi**, S. Brunak, Y. Chauvin, and A. Krogh. Naturally Occurring Nucleosome Positioning Signals in Human Exons and Introns, Journal of Molecular Biology, Vol. 263, 5, 503-510, (1996).
- J30. **P. Baldi** and Y. Chauvin. Hybrid Modeling, HMM/NN Architectures, and Protein Applications, Neural Computation, Vol. 8, 7, 1541-1565, (1996).
- J29. **P. Baldi**. Substitution Matrices and Hidden Markov Models, Journal of Computational Biology, Vol. 2, 3, 497-501, (1995).
- J28. **P. Baldi** and K. Hornik. Learning in Linear Networks: A Survey, IEEE Transactions on Neural Networks, Vol. 6, 4, 837-858, (1995)
- J27. **P. Baldi**. Gradient Descent Learning Algorithms Overview: A General Dynamical Systems Perspective, IEEE Transactions on Neural Networks, Vol. 6, 1, 182-195, (1995).
- J26. **P. Baldi** and Y. Chauvin. Hidden Markov Models of the G-Protein Coupled Receptor Family, Journal of Computational Biology, Vol. 1, 4, 311-335, (1994).
- J25. **P. Baldi**, Y. Chauvin, T. Hunkapiller, and M. A. McClure. Hidden Markov Models of Biological Primary Sequence Information, PNAS USA, Vol. 91, 3, 1059-1063, (1994).
- J24. **P. Baldi** and Y. Chauvin. Smooth On-Line Learning Algorithms for Hidden Markov Models, Neural Computation, Vol. 6, 2, 305-316, (1994).
- J23. **P. Baldi** and A. Atiya. How Delays Affect Neural Dynamics and Learning, IEEE Transactions on Neural Networks, Vol. 5, 4, 626-635, (1994).
- J22. **P. Baldi** and Y. Chauvin. Neural Networks for Fingerprint Recognition, Neural Computation, Vol. 5, 3, 402-418, (1993).
- J21. **P. Baldi** and S.S. Venkatesh. Random Interactions in Higher-Orders Neural Networks, IEEE Transactions on Information Theory, Vol. 39, 1, 274-283, (1993).
- J20. **P. Baldi** and F. Pineda. Contrastive Learning and Neural Oscillations, Neural Computation, Vol. 3, No. 4, 526-545, (1991).
- J19. **P. Baldi** and Y. Chauvin. Temporal Evolution of Generalization during Learning in Linear Networks, Neural Computation, Vol. 3, No. 4, 589-603, (1991).
- J18. E. A. Mayer, J. Furness, and **P. Baldi**. Can Regulatory Peptides Be Regarded as Words of a Biological Language?, American Journal of Physiology, Vol. 261(2), G171-G184, (1991).

- J17. S.S. Venkatesh and **P. Baldi**. Programmed Interactions in Higher-Orders Neural Networks: II. The Outer-Product Algorithm, Journal of Complexity, Vol. 7, 443-479, (1991).
- J16. S.S. Venkatesh and **P. Baldi**. Programmed Interactions in Higher-Orders Neural Networks: I. Maximal Capacity, Journal of Complexity, Vol. 7, 316-337, (1991).
- J15. **P. Baldi** and R. Meir. Computing with Arrays of Coupled Oscillators: an Application to Preattentive Texture Discrimination, Neural Computation, Vol. 2, No. 4, 458-471, (1990).
- J14. **P. Baldi**. On a Generalized Family of Colorings, Graphs and Combinatorics, Vol. 6, 95-110, (1990).
- J13. **P. Baldi** and E.C. Posner. Graph Coloring Bounds for Cellular Radio, Computers and Mathematics with Applications, Vol. 19, No. 10, 91-97, (1990).
- J12. **P. Baldi** and A. Atiya. Oscillations and Synchronizations in Neural Networks: an Exploration of the Labeling Hypothesis, International Journal of Neural Systems, Vol. 1, No. 2, 103-124, (1989).
- J11. **P. Baldi** and Y. Rinott. On Normal Approximations of Distributions in Terms of Dependency Graphs, Annals of Probability, Vol. 17, No. 4, 1646-1650, (1989).
- J10. **P. Baldi** and Y. Rinott. Asymptotic Normality of Some Graph Related Statistics, Journal of Applied Probability, Vol. 26, 171-175, (1989).
- J9. **P. Baldi** and K. Hornik. Neural Networks and Principal Component Analysis: Learning from Examples without Local Minima, Neural Networks, Vol. 2, No. 1, 53-58, (1988).
- J8. **P. Baldi** and W. Heiligenberg. How Sensory Maps Could Enhance Resolution Through Ordered Arrangements of Broadly Tuned Receivers, Biological Cybernetics, Vol. 59, No. 4, 5, 313-318, (1988).
- J7. **P. Baldi**. Neural Networks, Acyclic Orientations of the Hypercube and Sets of Orthogonal Vectors, SIAM Journal of Discrete Mathematics, Vol. 1, No. 1, 1-11, (1988).
- J6. **P. Baldi**. Group Actions and Learning for a Family of Automata, Journal of Computer and System Sciences, Vol. 36, No. 1, 1-15, (1988).
- J5. **P. Baldi**. Neural Networks, Orientations of the Hypercube and Algebraic Threshold Functions, IEEE Transactions on Information Theory, Vol. 34, No. 3, 523-530, (1988).
- J4. M. Aschbacher, **P. Baldi**, E. B. Baum, and R.M. Wilson. Embeddings of Ultrametric Spaces in Finite Dimensional Structures, SIAM Journal of Algebraic and Discrete Methods, Vol. 8, No. 4, 564-577, (1987).
- J3. **P. Baldi**. Symmetries and Learning in Neural Network Models, Physical Review Letters, Vol. 59, No. 17, 1976-1978, (1987).
- J2. **P. Baldi** and S.S. Venkatesh. Number of Stable Points for Spin Glasses and Neural Networks of Higher Orders, Physical Review Letters, Vol. 58, No. 9, 913-916, (1987).
- J1. **P. Baldi** and E.B. Baum. Bounds on the Size of Ultrametric Structures, Physical Review Letters, Vol. 56, No. 15, 1598-1600, (1986).

Books

- B5. **P. Baldi**. Deep Learning in Science. Cambridge University Press, (2021).
- B4. **P. Baldi**, P. Frasconi, and P. Smyth. Modeling the Internet and the Web—Probabilistic Methods and Algorithms. Wiley, (2003). Japanese version in 2007.
- B3. **P. Baldi** and G. Wesley Hatfield. DNA Microarrays and Gene Regulation—From Experiments to Data Analysis and Modeling. Cambridge University Press, (2002). English version for China and electronic publishing version through eBrary in 2003 and 2004. Paperback edition in 2011.
- B2. **P. Baldi**. The Shattered Self—the End of Natural Evolution, MIT Press, (2001). Paperback version in 2002.

- B1. **P. Baldi** and S. Brunak. Bioinformatics: The Machine Learning Approach. MIT Press, (1998). Second revised edition (2001). Indian and Chinese versions in 2003.

Refereed Book Chapters

- BC24. M. Samad, F. Agostinelli, and **P. Baldi**. Bioinformatics and Systems Biology of Circadian Rhythms: BIO_CYCLE and CircadiOmics. Methods in Molecular Biology, edition about "Circadian Regulation", Guiomar Solanas Fuster and Patrick-Simon Welz Editors, (2021).
- BC23. **P. Baldi**, P. Sadowski, and D. Whiteson. Deep Learning From Four-Vectors. Artificial Intelligence Applied to Particle Physics (IJMPA: International Journal of Modern Physics A). Editors: Paolo Calafiura, David Rousseau, and Kazuhiro Terao.
- BC22. F. Agostinelli, G. Hocquet, S. Singh, and **P. Baldi**. From Reinforcement Learning to Deep Reinforcement Learning: An Overview. Key Ideas in Learning Theory from Inception to Current State: Emmanuel Braverman's Legacy. Volume in the Springer series: LNCS State-of-the-Art Surveys, Springer, (2018).
- BC21. P. Sadowski and **P. Baldi**. Deep Learning in the Natural Sciences: Applications to Physics. Key Ideas in Learning Theory from Inception to Current State: Emmanuel Braverman's Legacy. Volume in the Springer series: LNCS State-of-the-Art Surveys, Springer, (2018).
- BC20. C. A. Azencott and **P. Baldi**. Virtual High-Throughput Screening with Two-Dimensional Kernels. In: *Hands-On Pattern Recognition. Challenges in Data Representation, Model Selection, and Performance Prediction*, I. Guyon, G. Cawley, G. Droor, and A. Saffari Editors, Lulu Press, (2011).
- BC19. J. Chen, M. A. Kayala, and **P. Baldi**. Reaction Explorer: Towards a Knowledge Map of Organic Chemistry to Support Dynamic Assessment and Personalized Instruction, in Enhancing Learning with Online Resources, Social Networking and Digital Libraries, American Chemical Society Symposium Series Book, R. E. Belford and J. Moore Editors, 191-209, (2010).
- BC18. Randall and **P. Baldi**. Transmembrane Beta-Barrel Protein Structure Prediction. In: Structural Bioinformatics of Membrane Proteins, Editor Dmitrij Frishman, Springer, 83-102, (2010).
- BC17. S. J. Swamidass and **P. Baldi**. Statistical Distribution of Chemical Fingerprints. Proceedings of WILF2005, Springer Verlag. Andrea Tattamanzi Editor.
- BC16. R. Jurdak, C. V. Lopes, and **P. Baldi**. Battery Lifetime Estimation and Optimization for Underwater Sensor Networks. In *Sensor Network Operations*, published by IEEE Press, edited by Shashi Phoha, Thomas F. La Porta and Christopher Griffin **,** 2004 (in press). Edited volume resulting from overflow to special issue on Sensor Networks in IEEE Transactions on Mobile Computing (2005).
- BC15. **P. Baldi**, G. Wesley Hatfield, and Li M. Fu. Clustering Algorithms for Gene Expression Analysis. Handbook of Computational Molecular Biology. S. Aluru Editor, Chapman & Hall/CRC Press, Computer and Information Science Series, (2005).
- BC14. S. Kremer and **P. Baldi**. Hidden Markov Models and Neural Networks. Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics. Wiley, (2004).
- BC13. **P. Baldi**. Surprise: A Shortcut for Attention? In: Neurobiology of Attention. L. Itti, G. Rees, and J. Tsotsos Editors, Elsevier, San Diego, CA, 24-28, (2005).
- BC12. S. P. Hung, G. Wesley Hatfield, S. Sundaresh, and **P. Baldi**. Understanding DNA Microarrays: Sources and Magnitudes of Variances in DNA Microarray Data Sets. In: Genomics, Proteomics, and Vaccines. G. Grandi Editor, John Wiley and Sons, (2003).

- BC11. **P. Baldi**. A Computational Theory of Surprise, in Information, Coding, and Mathematics. M. Blaum Editor, Kluwer, 1-25, (2002).
- BC10. **P. Baldi**, G. Pollastri, P. Frasconi, and A. Vullo. New Machine Learning Methods for the Prediction of Protein Topologies, in Artificial Intelligence and Heuristic Methods in Bioinformatics. P. Frasconi and R. Shamir Editors, IOS Press, 51-73, (2003).
- BC9. **P. Baldi**, S. Brunak, P. Frasconi, G. Pollastri, and G. Soda. Bidirectional Dynamics for Protein Secondary Structure Prediction, in Sequence Learning: Paradigms, Algorithms, and Applications, R. Sun and C. L. Giles Editors, Springer Verlag, 99-120, (2000).
- BC8. **P. Baldi**, S. Brunak, G. Pollastri, and P. Frasconi. Bidirectional IOHMMs and recurrent neural networks for protein secondary structure prediction, in Protein Sequence Analysis in the Genomic Era, Rita Casadio and Lanfranco Masotti Editors, CLUEB, Bologna, Italy, (2000).
- BC7. **P. Baldi**, Y. Chauvin, and V. Mittal Henkle. Software Foundation Libraries for the Design of Intelligent Systems. Neural Nets WIRN Vietri 98. Proceedings of the 10-th Italian Workshop on Neural Nets. M. Marinaro and R. Tagliaferri Editors. Springer Verlag Series on Perspectives on Neural Computing Series, J. Taylor Series Editor, 17-39, (1998).
- BC6. **P. Baldi**. Probabilistic Models of Neuronal Spike Trains, in Adaptive Processing of Temporal Information. C.L. Giles and M. Gori Editors, Lecture Notes in Computer Science, Springer Verlag, New York, 198-228, (1998).
- BC5. **P. Baldi**, Y. Chauvin, and K. Hornik. Back-propagation and unsupervised learning in linear networks, in Back-Propagation: Theory, Architectures and Applications. Y. Chauvin and D.E. Rumelhart Editors, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 389-432, (1995).
- BC4. **P. Baldi**. Gradient Descent Learning Algorithms: A Unified Perspective, in Back-Propagation: Theory, Architectures and Applications. Y. Chauvin and D.E. Rumelhart Editors, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 509-541, (1995).
- BC3. **P. Baldi** and Y. Chauvin. When Neural Networks Play Sherlock Holmes, in Back-Propagation: Theory, Architectures and Applications. Y. Chauvin and D.E. Rumelhart Editors, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 487-507, (1995).
- BC2. **P. Baldi** and Y. Chauvin. Statistical Models of Proteins: An Application to the G-Protein-Coupled Receptor Family, in Modern Approaches in Molecular Bioinformatics. S. Schulze-Kremer Editor, IOS Press, Amsterdam/New York, 53-102 (1994).
- BC1. **P. Baldi**, Y. Rinott, and C. Stein. A Normal Approximation for the Number of Local Maxima of a Random Function on a Graph, in Probability, Statistics and Mathematics: Papers in Honor of Samuel Karlin, T.W. Anderson, K.B. Athreya and D.L. Iglehart, Editors, Academic Press (1989).

Refereed Conference Proceedings (Partial List)

- C91. Stephen McAleer, Kevin Wang, Marc Lanctot, John Lanier, Pierre Baldi, and Roy Fox. Anytime Optimal PSRO for Two-Player Zero-Sum Games. AAAI 2022 Workshop on Reinforcement Learning in Games. (2022).
- C90. Mohammadamin Tavakoli, Alexander Shmakov, Francesco Ceccarelli and Pierre Baldi Rxn Hypergraph: a Hypergraph Attention Model for Chemical Reaction Representation. The Seventh International Workshop on Deep Learning on Graphs: Methods and Applications (AAAI--DLG22), Vancouver, Canada, (2021).
- C89. Keiser A, Dong T, Kramár E, Butler C, Matheos D, Tong L, Berchtold N, Chen S, Samad M, Magnan C, Beardwood J, Shanur S, Rodriguez A, Baldi, P, Cotman C, Wood M. Exercise opens a 'molecular memory window' to facilitate changes in gene expression, synaptic plasticity, and memory. Society for Neuroscience, (2021).

- C88. Farima Farmahinifarahani, Yadong Lu Vaibhav Saini, Pierre Baldi, Cristina Lopes. REX: Static Detection of Relevant Runtime Exceptions with Location Aware Transformer. IEEE SCAM 2021.
- C87. S. McAleer, F. Agostinelli, A. Shmakov, R. Fox, and P.Baldi. XDO: A Double Oracle Algorithm for Extensive-Form Games NeurIPS 2021, in press, (2021).
- C86. Y. Razeghi, S. Kask, Y. Lu, **P. Baldi**, S. Agarwal, R. Dechter. Deep Bucket Elimination. Proceedings of IJCAI-21 (the 30th International Joint Conference on Artificial Intelligence).
- C85. J. Liu, J. Ott, J. Collado, B. Jargowsky, W. Wu, J. Bian, **P. Baldi**, Deep-Learning-Based Kinematic Reconstruction for DUNE, in proceeding of the 34th Conference on Neural Information Processing Systems (NeurIPS), Machine Learning and the Physical Sciences Workshop, (2020).
- C84. Shmakov, A., Lanier, J., McAleer, S., Achar, R., Lopes, C.V. and Baldi, P. (2020). ColosseumRL: A Framework for Multiagent Reinforcement Learning in N-Player Games. In Proceedings of Challenges and Opportunities for Multi-Agent Reinforcement Learning (COMARL AAAI).
- C83. S. McAleer, J. B. Lanier, R. Fox, and **P. Baldi**. Pipeline PSRO: A Scalable Approach for Finding Approximate Nash Equilibria in Large Games. Accepted, NeurIPS, (2020).
- C82. Y. Lu, J. Collado, K. Bauer, D. Whiteson, **P. Baldi**. Sparse Image Generation with Decoupled Generative Models. NeurIPS 2019, Machine Learning and Physical Sciences Workshop.
- C81. L. Li, D. Pluta, B. Shahbaba, **P. Baldi**, N. Fortin, H. Ombao. Modeling Dynamic Functional Connectivity with Latent Factor Gaussian Processes. Proceedings of NIPS 2019.
- C80. Keiser AA1, Butler C2, Kramár EA1, Matheos DP1, Berchtold NC2, Chen S3, Samad M3, Magnan C3, Baldi, P3, Cotman CW2, Wood MA. Effects of exercise on memory performance and hippocampal gene expression. Society for Neuroscience, (2019).
- C79. Lingge Li, Dustin Pluta, Babak Shahbaba and Pierre Baldi. Latent Factor Gaussian Process with Log-Euclidean Metric for Dynamic Covariance Modeling. ICML19, submitted, 2019.
- C78. S. McAleer, F. Agostinelli, A. K Shmakov, and **P. Baldi**. Solving the Rubik's Cube with Approximate Policy Iteration. ICLR 2019, New Orleans.
- C77. **P. Baldi** and R. Vershynin. On Neuronal Capacity. Advances in Neural Information Processing Systems (NIPS) XX (Montreal, Canada), NIPS 2018. (2018). Accepted for oral presentation.
- C76. Vaibhav Saini, Farima Farmahinifarahani, Yadong Lu, Di Yang, Pedro Martins, Hitesh Sajnani, Pierre Baldi and Cristina Lopes. Towards Automating Precision Studies of Clone Detectors. ICSE 2019, in press (acceptance rate of 21%). Best Artifact Award.
- C75. Vaibhav Saini, Farima Farmahini Farahani, Yadong Yu, Pierre Baldi and Cristina Lopes. Oreo: Detection of Clones in the Twilight Zone. ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering ESEC/FSE 2018, in press, (2018). ESEC/FSE 2018 Distinguished Paper Award.
- C74. Christine Lee, Ira Hofer, Pierre Baldi, Maxime Cannesson, Deep Learning for Predicting Postoperative Outcomes: AKI, Reintubation, and In-Hospital Mortality. Abstract for the American Society of Anesthesiologists Conference (October 21-25, 2017).
- C73. C. Lee, I. Hofer, M. Cannesson, and **P. Baldi**. Deep Learning for Predicting in Hospital Mortality. Abstract submitted for Society for Technology in Anesthesia. Accepted for oral presentation, January 17, 2017, San Diego. Best in Show Award.
- C72. P. Sadowski, D. Whiteson, and **P. Baldi**. Searching for Higgs Boson Decay Modes with Deep Learning. NIPS 2014.

- C71. D. Chicco, P. Sadowski, and **P. Baldi**. Deep Autoencoder Neural Networks for Gene Ontology Annotation Predictions. Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, Newport Beach, CA, September 2014. Accepted also for oral presentation.
- C70. **P. Baldi** and P. Sadowski. Understanding Dropout. Advances in Neural Information Processing Systems (NIPS) 26 (Lake Tahoe, NV), C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger Editors, 2814--2822, (2013). Accepted also for oral presentation.
- C69. Francesco Napolitano, Roberto Tagliaferri, and **Pierre Baldi**. An Adaptive Reference Point Approach to Efficiently Search Large Chemical Databases in Recent Advances of Neural Networks Models and Applications, Proceedings of the 23rd Workshop of the Italian Neural Networks Society (SIREN), May 23–25, Vietri sul Mare, Salerno, Italy (2013). Simone Bassis, Anna Esposito and Francesco Carlo Morabito Editors, Smart Innovation, Systems and Technologies Series, Volume 26, PP. 63-74, Springer Cham Heidelberg New York Dordrecht London, ISSN 2190-3018 ISSN 2190-3026 (electronic), ISBN 978-3-319-04128-5 ISBN 978-3-319-04129-2 (eBook), DOI 10.1007/978-3-319-04129-2.
- C68. P. Di Lena, K. Nagata, and **P. Baldi**. Deep Spatio-Temporal Architectures and Learning for Protein Structure Prediction. Advances in Neural Information Processing Systems (NIPS) 25 (Lake Tahoe, NV)}, Editors: F. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger, 512—520, Curran Associates Inc, (2012).
- C67. **P. Baldi**, R. Baronio, E. De Cristofaro, P. Gasti, and G. Tsudik. Efficient and Secure Testing of Fully-Sequenced Human Genomes (Extended Abstract). Oral Presentation at the Provable Privacy (PropPriv) workshop in Vigo, Spain, (2012).
- C66. M. Kayla and **P. Baldi**. A Machine Learning Approach to Predict Chemical Reactions. Advances in Neural Information Processing Systems (NIPS) 24 (Granada, Spain), Editors: J. Shawe-Taylor, R.S. Zemel, P.L. Bartlett, F. Pereira, and K.Q. Weinberger, 747—755, Curran Associates Inc., (2011).
- C65. **P. Baldi**, R. Baronio, E. De Cristofaro, P. Gasti, and G. Tsudik. Countering GATTACA: Efficient and Secure Testing of Fully Sequenced Human Genomes. ACM Computer and Communication Security Conference (CCS 2011).
- C64. **P. Baldi**. Autoencoders, Unsupervised Learning, and Deep Architectures. Neural Information Processing Systems (NIPS) 2010 Conference, Workshop on Deep Learning and Unsupervised Feature Learning, Whistler, Canada, December 2010.
- C63. **P. Baldi**, C. Azencott, and S. J. Swamidass. Bridging the Gap Between Neural Network and Kernel Methods: Applications to Drug Discovery. Proceedings of the 22nd Italian Workshop on Neural Networks, WIRN 2010, IOS Press, (2010).
- C62. Linstead, E., Hughes, L., Lopes, C., **Baldi, P**. "Information-Theoretic Metrics for Project-Level Scattering and Tangling." International Conference on Software Engineering and Knowledge Engineering (SEKE). Redwood City, CA. July 2010.
- C61. F. Napolitano, R. Tagliaferri, and **P. Baldi**. A Scalable Reference-Point Based Algorithm to Efficiently Search Large Chemical Databases. IEEE Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN 2010), (2010).
- C60. E. Linstead, L. Hughes, C. Lopes, and **P. Baldi**. "Information-Theoretic Metrics for Project-Level Scattering and Tangling." International Conference on Software Engineering and Knowledge Engineering (SEKE). Redwood City, CA. July 2010.
- C59. E. Linstead, L. Hughes, C. Lopes, and **P. Baldi**. Software Analysis with Unsupervised Topic Models. Neural Information Processing Systems (NIPS) Conference, Workshop on "Applications for Topic Models", Whistler, Canada, December 2009.

- C58. E. Linstead, L. Hughes, and **P. Baldi**. The Evolution of Concerns, Scattering, and Tangling in Eclipse and ArgoUML. Third International Symposium on Empirical Software Engineering and Measurement (ESEM), Lake Buena Vista, Florida, October, poster, (2009).
- C57. S. Javanmardi, C. Lopes, and **P. Baldi**. User Contribution and Trust in Wikipedia. Proceedings of the 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing, November 11-14, Washington D.C., (2009).
- C56. E. Linstead and **P. Baldi**. Mining the Coherence of GNOME Bug Reports with Statistical Topic Models. MSR '09: Proceedings of the Sixth International Working Conference on Mining Software Repositories, Vancouver, BC, Canada, IEEE Computer Society, in press, (2009).
- C55. E. Linstead, L. Hughes, C. Lopes, and **P. Baldi**. Exploring Java Software Vocabulary: A Search and Mining Perspective. SUITE 2009: Proceedings of the First International Workshop on Search-Driven Development - Users, Tools, and Applications." Vancouver, BC. May 16, 2009.
- C54. E. Linstead, C. Lopes, and **P. Baldi**. An Application of Latent Dirichlet Allocation to Analyzing Software Evolution. International Conference on Machine Learning and Applications (ICMLA), San Diego, (2008).
- C53. **P. Baldi**, C. Lopes, E. Linstead, S. Bajracharya. A Theory of Aspects as Latent Topics. 2008 ACM SIGPLAN Conference on Object-Oriented Programming Systems, Languages, and Applications. OOPSLA '08, October 19-23, Nashville, TN, (2008).
- C52. D. S. Hirschberg and **P. Baldi**. Effective Compression of Monotone and Quasi-Monotone Sequences of Integers. Proceedings of the 2008 Data Compression Conference (DCC 08), 520, Snowbird, Utah, J. A. Storer and M.W. Marcellin Editors, IEEE Computer Society Press, 520, (2008).
- C51. E. Linstead, P. Rigor, S. Bajracharya, C. Lopes, and **P. Baldi**. Mining Internet-Scale Software Repositories. Advances in Neural Information Processing Systems 20, NIPS 2007, Y. Weiss, B. Scholkopf, and J. Platt Editors, 929-936, MIT Press, Cambridge, MA, (2008).
- C50. O. Lemos, S. Bajracharya, J. Ossher, R. Morla, P. Masiero, **P. Baldi**, and C. Lopes. CodeGenie: Using Test-Cases to Search and Reuse Source Code. Proceedings of the 22nd IEEE/ACM International Conference on Automated Software Engineering (ASE'07), Tool Demonstration Session, 525-526, Atlanta, GA, November 2007.
- C49. E. Linstead, P. Rigor, S. Bajracharya, C. Lopes, and **P. Baldi**. Mining Concepts from Code with Probabilistic Topic Models. Proceedings of the 22nd IEEE/ACM International Conference on Automated Software Engineering, (ASE'07), 461-464, Atlanta, GA, November 2007.
- C48. E. Linstead, P. Rigor, S. Bajracharya, C. Lopes, and **P. Baldi**. Mining Eclipse Developer Contributions via Author-Topic Models. Proceedings of the MSR 2007: International Workshop on Mining Software Repositories, 30-33, Minneapolis, MN, (2007). Best Paper Award, MSR "Scale" Challenge.
- C47. R. Jurdak, P.M.Q. Aguiar, **P. Baldi**, and C.V. Lopes. "Software Modems for Underwater Sensor Networks," Proceedings of IEEE/OES Oceans'07, June 2007. Aberdeen, Scotland.
- C46. L. Wu and **P. Baldi**. "A Scalable Machine Learning Approach to Go", Advances in Neural Information Processing Systems 19, NIPS 2006, Y. Weiss, B. Scholkopf, and J. Platt Editors, 1521-1528, MIT Press, Cambridge, MA, (2007).
- C45. R. Jurdak, C.V. Lopes, and **P. Baldi**. "Software Acoustic Modems for Short Range Mote-Based Underwater Sensor Networks," In Proceedings of IEEE Oceans – Asia Pacific, 1-7, Singapore, May, 2006.
- C44. R. Jurdak, C. V. Lopes, P. M. Q. Aguiar, and **P. Baldi**. "A Comparative Analysis and Experimental Study on Wireless Aerial and Underwater Acoustic Communications".

Proceeding of the 2006 International Conference on Digital Telecommunications, 31, IEEE Computer Society, (2006).

- C43. **P. Baldi**. "Chemoinformatics, Drug Design, and Systems Biology", Proceedings of the 16th International Conference on Genome Informatics, (GIW 2005), Yokohama, Japan, December 2005.
- C42. **P. Baldi**. "Exploring Chemical Space with Computers: Informatics Challenges for AI and Machine Learning", BIOMAT V Symposium, Petrópolis, Brazil, December 2005, World Scientific Co.Pte.Ltd.
- C41. **P. Baldi** and L. Itti. "Attention: Bits versus Wows", Proceedings of the International Conference on Neural Networks and Brain (ICNN&B05), Edited by Mingsheng Zhao and Zhongzhi Shi, Vol. 1, PL-56-PL61, Beijing, China, IEEE Press, October 2005.
- C40. L. Itti and **P. Baldi**. "Bayesian Surprise Attracts Human Attention", Advances in Neural Information Processing Systems 18, NIPS 2005, Y. Weiss, B. Scholkopf, and J. Platt Editors, MIT Press, Cambridge, MA, 547--554, (2006).
- C39. R. Jurdak, **P. Baldi**, and C. V. Lopes. "State-Driven Energy Optimization in Wireless Sensor Networks," Proceedings of the IEEE/IEE International Conference on Sensor Networks (SENET), 356-363, Montreal, Canada. August 2005.
- C38. R. Jurdak, **P. Baldi**, C. V. Lopes. "Energy-Aware Adaptive Low Power Listening for Sensor Networks", In Proceedings of the 2nd International Workshop on Networked Sensing Systems (INSS), pp. 24-29, ISBN 0-9743611-2-7, San Diego, CA. June 2005.
- C37. R. Jurdak, C. V. Lopes, and **P. Baldi**. "Programming Model for Supporting Complex Optimizations in Sensor Networks," Proceedings of the Fourth International IEEE/ACM Conference on Information Processing in Sensor Networks (IPSN'05). Los Angeles, CA. April 2005.
- C36. J. Cheng and **P. Baldi**. Three-Stage Prediction of Protein Beta-Sheets by Neural Networks, Alignments, and Graph Algorithms. Proceedings of the 2005 Conference on Intelligent Systems for Molecular Biology, ISMB 05. Bioinformatics, Supplement 1, i75-84, (2005).
- C35. S. J. Swamidass, J. Chen, P. Phung, J. Bruand, L. Ralaivola, and **P. Baldi**. Kernels for Small Molecules and the Prediction of Mutagenicity, Toxicity, and Anti-Cancer Activity. Proceedings of the 2005 Conference on Intelligent Systems for Molecular Biology, ISMB 05. Bioinformatics, 21, Supplement 1, i359-368, (2005).
- C34. R. Jurdak, C. V. Lopes, **P. Baldi**. Adaptive Low Power Listening for Sensor Networks. 11th Annual International Conference on Mobile Computing and Networking.
- C33. L. Itti, **P. Baldi**. A Principled Approach to Detecting Surprising Events in Video. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2005.
- C32. L. Ralaivola, L. Wu, and **P. Baldi**. SVM and Pattern-Enriched Common Fate Graphs for the Game of GO. European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium, April 2005. Accepted for Oral Presentation.
- C31. A. Mandal, C. V. Lopes, T. Givargis, A. Haghighat, R. Jurdak, and **P. Baldi**. Beep: 3D Indoor Positioning Using Audible Sound. Proceedings of the 2004 IEEE Consumer Communications and Networking Conference, 348—353, (2004).
- C30. **P. Baldi**, J. Cheng, and A. Vullo. Large-Scale Prediction of Disulphide Bond Connectivity. Advances in Neural Information Processing Systems 17, NIPS 2004, in press, (2004).
- C29. R. Jurdak, C. Videira Lopes, and **P. Baldi**. An Identification Scheme for Location Systems Using Robust Acoustic Signals. Proceeding of the ACS/IEEE International Conference on Pervasive Services, Beirut, Lebanon. IEEE Computer Society Press, (2004).
- C28. G. Pollastri, **P. Baldi**, A. Vullo, and P. Frasconi. Prediction of Protein Topologies Using Generalized IOHMMs and RNNs. Advances in Neural Information Processing Systems 15,

S. Becker, S. Thrun, and K. Obermayer Editors, Volume 15, MIT Press, Cambridge, MA, (2003).

- C27. G. Pollastri and **P. Baldi**. Prediction of Contact Maps by GIOHMMs and Recurrent Neural Networks Using Lateral Propagation From All Four Cardinal Corners. Proceedings of the 2002 Conference on Intelligent Systems for Molecular Biology, ISMB 02. Bioinformatics, 18, Supplement 1, S62-S70, (2002). Nominated for Best Paper Award.
- C26. L. De Nardis, **P. Baldi**, and M. G. Di Benedetto. UWB Ad-Hoc Networks. Proceedings of the 2002 IEEE Conference on Ultra Wide Band Systems and Technologies (UWBST 2002), Baltimore, MD, 219-224, (2002).
- C25. G. Pollasti, **P. Baldi**, P. Fariselli, R. Casadio. Improved Prediction of the Number of Residue Contacts in Proteins by Recurrent Neural Networks. Proceedings of the 2001 Conference on Intelligent Systems for Molecular Biology, (ISMB01), Copenhagen, Denmark, AAAI Press (2001).
- C24. M. G. Di Benedetto and **P. Baldi**. A Model for Self-Organizing Large Scale Wireless Networks. International Symposium on 3G Infrastructure and Services. Athens, Greece, July 2001.
- C23. **P. Baldi**, G. Pollastri, C. A. F. Andersen, and S. Brunak. Matching Protein β-Sheet Partners by Feedforward and Recurrent Neural Networks. Proceedings of the 2000 Conference on Intelligent Systems for Molecular Biology, (ISMB00), La Jolla, CA, 25-36, AAAI Press (2000).
- C22. S. Hampson, **P. Baldi**, D. Kibler, and S. Sandmeyer. Analysis of Yeast's ORFs Upstream Regions by Parallel Processing, Microarrays, and Computational Methods. Proceedings of the 2000 Conference on Intelligent Systems for Molecular Biology, (ISMB00), La Jolla, CA, 190-201, AAAI Press (2000).
- C21. **P. Baldi**, G. Pollastri, C. A. F. Andersen, and S. Brunak. Protein β-Sheet Partner Prediction by Neural Networks. In: Artificial Neural Networks in Medicine and Biology. Proceedings of the ANNIMAB-1 Conference, Göteborg, Sweden, H. Malmgren, M. Borga and L. Niklasson Editors, Springer Verlag, (2000).
- C20. **P. Baldi**, S. Brunak, Y. Chauvin, and A. Gorm Pedersen. Computational Applications of DNA Structural Scales, Proceedings of the 1998 Conference on Intelligent Systems for Molecular Biology, (ISMB98), Montreal, Canada, 35-42, AAAI Press (1998).
- C19. **P. Baldi**, A. Gorm Pedersen, S. Brunak, and Y. Chauvin. Characterization of Prokaryotic and Eukaryotic Promoters Using Hidden Markov Models, Proceedings of the 1996 Conference on Intelligent Systems for Molecular Biology (ISMB96), Washington University, Saint Louis, 182-191, AAAI Press, (1996).
- C18. **P. Baldi**, S. Brunak, Y. Chauvin, and A. Krogh. Hidden Markov Models for Human Genes: Periodic Patterns in Exon Sequences, Theoretical and Computational Methods in Genome Research, S. Suhai Editor, Plenum, NY, (1996).
- C17. **P. Baldi** and K. Hornik. Universal Approximation and Learning of Trajectories Using Oscillators, Advances in Neural Information Processing Systems 8, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo Editors, MIT Press, Cambridge MA, 451-457, (1996).
- C16. **P. Baldi** and Y. Chauvin. Protein Modeling with Hybrid Hidden Markov Model/Neural Network Architectures, Proceedings of the 1995 Conference on Intelligent Systems for Molecular Biology (ISMB95), Cambridge (UK), 39-47, (1995).
- C15. **P. Baldi**, Y. Chauvin, S. Brunak, J. Engelbrecht, and A. Krogh. Periodic Sequence Patterns in Human Exons, Proceedings of the 1995 Conference on Intelligent Systems for Molecular Biology (ISMB95), Cambridge (UK), 30-38, (1995).
- C14. P. Smyth, U. Fayadd, P. Perona, M. Burl, and **P. Baldi**. Inferring Ground Truth from Subjective Labelling of Venus Images. Advances in Neural Information Processing Systems,

G. Tesauro, D. S. Touretzky and T. K. Leen Editors, MIT Press, Cambridge MA, 1085-1092, (1995).

- C13. **P. Baldi** and Y. Chauvin. Trading Decision Learning: from Theory to Personal Traders, Proceedings of the Second International Conference on Neural Networks in the Capital Markets, Pasadena, CA (1994).
- C12. **P. Baldi**, S. Brunak, Y. Chauvin, J. Engelbrecht, and A. Krogh. Hidden Markov Models of Human Genes, Advances in Neural Information Processing Systems, J. D. Cowan, G. Tesauro and J. Alspector Editors, Morgan Kauffmann Publishers, 761-768, (1994).
- C11. **P. Baldi** and Y. Chauvin. Discrimination of Tyrosine and Serine/Threonine Kinase Sub-Families by Hidden Markov Models, Proceedings of The Third International Conference on Bioinformatics and Genome Research, World Scientific Publishing Co., New Jersey, (1994).
- C10. **P. Baldi** and Y. Chauvin. Modeling Protein Families and Human Genes: Hidden Markov Models and a Little Beyond, Proceedings of the 1994 Fifth Generation Computing Symposium, Workshop on Fusion of Molecular Biology and Knowledge Processing, (Tokyo, Japan), (1994).
- C9. **P. Baldi**, Y. Chauvin, T. Hunkapiller, and M. A. McClure. Hidden Markov Models in Molecular Biology: New Algorithms and Applications, Advances in Neural Information Processing Systems, S. J. Hanson, J. D. Cowan and C. Lee Giles Editors, Morgan Kauffmann Publishers, 747-754, (1993).
- C8. **P. Baldi** and Nikzad Toomarian. Learning Trajectories with a Hierarchy of Oscillatory Modules, Proceedings of the 1993 IEEE International Conference on Neural Networks, San Francisco, CA, Volume III, 1172-1176, and, in similar form, in Proceedings of the 1993 European Symposium on Artificial Neural Networks (ESANN93), Brussels, 183-188.
- C7. **P. Baldi** and Y. Chauvin. Trading Decision Learning, Neural Networks for Computing Conference (abstract), Snowbird, UT, (1992).
- C6. **P. Baldi**. A Modular Hierarchical Approach to Learning, Proceedings of the 2nd International Conference on Fuzzy Logic and Neural Networks, (Iizuka, Japan), Vol. 2, 985-988, (1992).
- C5. **P. Baldi**. Computing with Arrays of Bell-Shaped and Sigmoid Functions. Bernstein Polynomials, the Heat Equation and Universal Approximation Properties, Proceedings of the 1990 Conference on Neural Information Processing Systems, Denver, CO, Morgan Kauffman Publishers, 735-742, (1991).
- C4. **P. Baldi**, Y. Rinott, and C. Stein. On the Distribution of the Number of Local Minima of a Random Function on a Graph, Proceedings of the 1989 Conference on Neural Information Processing Systems, Denver, CO, Morgan Kauffman Publishers, 727-732, (1990).
- C3. **P. Baldi**. Neural Networks and Principal Component Analysis: Landscapes and Algorithms, Proceedings of the 1988 Conference on Neural Information Processing, Denver, CO, Morgan Kauffman Publishers, 65-72, (1989).
- C2. **P. Baldi** and S.S. Venkatesh. On Properties of Networks of Neuron-Like Elements: Complexity and Capacity, Proceedings of the IEEE Conference on Neural Information Processing Systems (Denver 1987), published by the American Institute of Physics, (D.Z. Anderson, Editor).
- C1. **P. Baldi** and E.B. Baum. Caging and Exhibiting Ultrametric Structures, Proceedings of the Conference on Neural Networks for Computing, Snowbird, Utah, April 1986, published by the American Institute of Physics, 35-40, (1986), (John S. Denker, Editor).

Other

- R3. F Agostinelli, M Hoffman, P Sadowski, **P Baldi**. Learning activation functions to improve deep neural networks. arXiv preprint arXiv:1412.6830 (2014). [over 500 citations]
- R2. **P. Baldi** and K. Muller and G. Schneider. Editorial: Charting Chemical Space:

- Challenges and Opportunities for Artificial Intelligence and Machine Learning. Molecular Informatics, 30, 9, 751, (2011).
- R1. **P. Baldi** What Genes are Made Of. Review of edited book entitled: The Genomic Revolution: Unveiling the Unity of Life. Pathways, (2003).
- Several tutorials (e.g. Chemoinformatics Tutorial given at the 2006 ISMB Conference).

Recent Articles in the Press Covering Some of the Publications Above

For AI and the Rubik's Cube:

- https://www.technologyreview.com/s/611281/a-machine-has-figured-out-rubiks-cube-all-by-itself/
- https://gizmodo.com/self-taught-ai-masters-rubik-s-cube-in-just-44-hours-1826918072
- https://www.hpcwire.com/2018/07/25/new-deep-learning-algorithm-solves-rubiks-cube/
- https://hothardware.com/news/university-california-deep-learning-machine-teaches-itself-solve-rubiks-cube
- https://www.latimes.com/local/lanow/la-me-ln-rubiks-cube-20180623-story.html
- https://interestingengineering.com/an-ai-system-taught-itself-how-to-solve-the-rubiks-cube-in-just-44-hours
- https://hub.packtpub.com/deepcube-a-new-deep-reinforcement-learning-approach-solves-the-rubiks-cube-with-no-human-help/
- https://www.yahoo.com/news/machine-learning-solve-rubik-apos-162300548.html
- https://www.popularmechanics.com/culture/a21562414/machine-learning-finally-tackles-the-rubiks-cube/
- https://www.cnet.com/news/machines-can-now-finish-the-rubiks-cube-without-human-help/
- https://www.fudzilla.com/news/ai/46563-ai-defeats-rubik-s-cube-without-human-help
- https://www.digitaltrends.com/computing/ai-rubiks-cube-solution-university-of-california/
- https://techxplore.com/news/2018-06-deepcube-solver-approach-cube.html
- http://www.hardocp.com/news/2018/06/17/machine_has_figured_out_rubikrsquos_cube_all_by_itself
- https://www.techspot.com/news/75140-scientists-have-developed-machine-can-solve-rubik-cube.html

From Erno Rubik, the inventor of the Rubik's Cube:

- https://plus.google.com/+ErnoRubik/posts/VvnVFx5HP8i
- https://plus.google.com/+ErnoRubik/posts/JHcT6S3kzC8

For AI and Drug Discovery:

- http://analytics.dkv.global/data/pdf/AI-for-DD-Q4/Top-100-AI-Leaders.pdf
- https://www.linkedin.com/pulse/top-100-ai-leaders-drug-discovery-advanced-healthcare-colangelo/
- https://www.nature.com/articles/d41586-018-05267-x
- https://ai-pharma.dka.global/ai-leaders/

Databases, Software, and Web Servers

- Please see: www.ics.uci.edu/~pfbaldi and www.igb.uci.edu/servers/servers.html.

Protein Structure Prediction (SCRATCH suite):
SSpro: Protein secondary structure prediction server (3 classes).
SSpro8: Protein secondary structure prediction server (8 classes).
ACCpro: Solvent accessibility prediction server.
CONpro: Residue contact number prediction server.
DIpro: Disulphide bridge prediction server.
BETApro: Beta-residue and beta-strand prediction server.
MUpro: Single amino acid mutation stability prediction server.
DISpro: Disordered region prediction server.
DOMpro: Domain prediction server.
CMAPpro: Contact map prediction server.
SVMcon: Amino acid contact prediction server (using Support Vector Machines).
CCMAPpro: Coarse contact map prediction server.
CMAP23Dpro: 3D reconstruction from contact map.
3Dpro: 3D structure prediction server (ab initio).
TMBpro: Transmembrane beta-barrel features and tertiary structure prediction server.
SELECTpro: Protein model selection server using a structure-based energy function.
SOLpro: Protein solubility prediction server.
SIDE-pro: Protein side-chain conformation prediction server
D-Finder: Kinase docking site prediction server.

Mass Spectrometry suite:
Link Finder: Processes mass spectrometry data and allows users to identify peptides linked by MS cleavable crosslinkers.

Immunology suite:
BEpro: Discontinuous B-cell epitope prediction server.
COBEpro: Continuous B-cell epitope prediction server.
ANTIGENpro: Protein antigenicity prediction server.
.
Sequence Modeling and Analysis:
HMMpro: Hidden Markov Model simulator for biological sequence analysis, with graphical interface.

Comparative Genomics:
LineUp: Comparative genomics server (order + density).
CloseUp: Comparative genomics server (density alone).

DNA Microarray Analysis:
Cyber-T: DNA microarray gene expression analysis server.

Databases and Systems Biology:
ICBS: Inter-chain beta sheet database of protein-protein interactions and web server.
PSPDB: Poxvirus structural proteomics database and webserver.
Sigmoid Database: Database for molecular interactions and pathways (Systems Biology).
Sigmoid Architecture.
GOnet: Yeast database and visualization tool combining gene (SGD), gene ontology (GO), and gene interaction information (GRID).
MotifMap: Database and web server of genome wide Transcription Factor binding sites for all model organisms with alignments and evolutionary conservation scores.
MotifMap-RNA: Database and web server of genome wide RNA-binding protein binding sites for all model organisms with alignments and evolutionary conservation scores.

Circadiomics: Integrated genomic, proteomic, and metabolomic database, software, and web server for the study of circadian rhythms.

Mitochondrial Modeling:
Systems biology mathematical models of mitochondria.

High-Throughput Sequencing (HTS) Pipeline:
Computational pipeline and web server for analyzing all the data produced by the UCI Genomics High-Throughput Facility (GHTF) for the entire campus. The pipeline is used to produce, store, and analyze reads, map them to relevant genomes, and apply relevant quantitative analyses for sequencing, ChIP-seq, and RNA-seq projects.

Chemoinformatics:
ChemDB: Chemoinformatics portal including: (1) a large database (5M) of organic compounds for molecular docking, drug screening, and retrosynthesis applications; (2) multiple Web server machine learning-based predictors of physical, chemical, and biological properties; (3) an organic chemistry expert system (Reaction Explorer) used to power several applications, such as Synthesis Explorer and Mechanism Explorer. Synthesis Explorer and Mechanism Explorer are interactive tutorial systems to learn undergraduate-level organic chemistry. Synthesis Explorer and Mechanism Explorer have been adopted and used in relevant undergraduate organic chemistry classes at UCI. Reaction Explorer has been licensed and is being distributed worldwide by Wiley; (4) an organic chemistry expert system (Reaction Predictor) that uses deep learning to predict the outcome of chemical reactions.

Machine Learning: Modules for deep learning, recursive and recurrent neural networks, dropout.

Software Mining:
Sourcerer: Database and information retrieval system for source code. Search engine for opensource software.

AI:
DeepCube: Deep Reiforcement Learning system to solve the Rubik's cube and other related combinatorial puzzles.

Invention Disclosures and Patent Applications (Available Upon Request)

INVITED TALKS, TUTORIALS, PRESENTATIONS (Partial List)

Sample before 2005:

- University of Florida, Gainesville
- MIT
- Genome Therapeutics
- University of Madrid, Spain
- University of Bologna, Italy
- Bioinformatics School, San Miniato, Italy
- Systems Biology Conference, Caltech
- ISMB Conference
- JOBIM Conference in Toulouse, France
- University of Paris VI, Paris, France
- Ecole Normale Superieure, Paris, France
- Fred Hutchinson Cancer Research Center
- University of Rome, Italy

- Interface Conference
- University of Rio de Janeiro, Brazil
- VII Brazilian Symposium on Neural Networks, Recife, Brazil
- Institute for Mathematics and its Applications (IMA), University of Minnesota
- University of California, San Diego
- RFIA Conference (Reconaissance des Formes et Intelligence Artificielle) , Toulouse, France
- Joint Symposium on Neural Computation
- International Joint Conference on Neural Networks
- University of Louisiana, Lafayette
- University of Algarve, Faro, Portugal
- IEEE CSB (Computational Systems Bioinformatics), Stanford
- University of Arizona, Tucson
- Genentech
- UCSD

Since 2005:

- Invited Tutorial in Bioinformatics, GENSIP 2005, IEEE International Workshop on Genomic Signal Processing and Statistics, Newport, RI, (May 2005)
- Keynote Speaker, RFIA Conference (Reconnaissance des Formes et Intelligence Artificielle) , Nice, France (June 2005)
- Keynote speaker, IJCNN, Montreal, Canada (August 2005)
- Keynote speaker, Computer Vision and Pattern Recognition Conference (CVPR), San Diego, CA (July 2005)
- Invited speaker The Chinese University of Hong Kong (Distinguished Lecture Series)
- Keynote Speaker, International Conference on Neural Networks and Brain, Beijing, Chain (October 2005)
- Invited speaker NIHS Toxicology Division, Tokyo, Japan
- Invited speaker, Kyoto University (December 2005)
- Keynote speaker, Second International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics Crema, ITALY (September 2005)
- Keynote Speaker, Bioinformatics Conference, Atlanta, GA (November 2005)
- Keynote speaker BIOMAT Petropolis, Brazil (December 2005)
- Keynote speaker for the 16th International Conference on Genome Informatics (GIW 2005) December 19-21, Yokohama, Japan
- Invited speaker, Indiana University, Bloomington
- Invited speaker, Mitre Corporation
- Keynote speaker, Gubelkian Institute of Sciences, Portugal (June 2006)
- Invited speaker, University of Rome, Italy (June 2006)
- Invited speaker, University of Lisbon, Portugal (June 2006)
- Chemoinformatics Tutorial ISMB (Intelligent Systems for Molecular Biology) Conference, Fortaleza, Brazil (August 2006)
- Invited speaker, University of Naples, Italy (October 2006)
- Invited speaker, University of Benevento, Italy (October 2006)
- Invited speaker, Iowa State University (March 2007)
- Keynote speaker 2007 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, Hawaii, (April 2007)

- Keynote speaker, Italian Conference on Bioinformatics, Naples, Italy (April 2007)
- Keynote speaker Physical and Chemical Foundations of Bioinformatics Methods, Dresden, Germany (June 2007)
- Invited keynote speaker, Conference on Physical and Chemical Foundations of Bioinformatics Methods, Dresden (June 2007)
- Invited speaker, Iowa State University, 2007
- Invite keynote speaker for the IEEE 2007 Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Honolulu, Hawaii (April 2007).
- Invited keynote speaker Italian Bioinformatics Conference, Naples, Italy (April 2007)
- Invited speaker workshop on Interdisciplinary Strategic Issues in e-Science and Cyber-Infrastructure, Caltech, Pasadena, California (June 2007)
- Invited speaker, University of Naples, Italy (June 2007)
- Invited speaker, University of Benevento, Italy (June 2007)
- Invited keynote speaker, 2007 International Conference on Machine Learning and Applications (ICMLA'07), Cincinnati, Ohio (December 2007).
- Invited distinguished speaker, Shumaker Bioinformatics Seminar of the Informatics Institute at the University of Missouri Columbia, Missouri (October 2007)
- Invited keynote speaker, BioSys Symposium on Human Post-Genomics, Evolution and the Future of Life held at the Royal Library, Copenhagen, Denmark (May 2008)
- Invited speaker, University College, Dublin, Ireland (June 2008)
- Invited speaker, Biogen, Italy (June 2008)
- Invited speaker, University of Salerno, Italy (June 2008)
- Invited keynote speaker, 2008 International Conference on Artificial Neural Networks (ICANN 2008), Prague, Czech Republic (September 2008)
- Invited member of the scientific committee of the 2nd International Workshops on Practical Applications of Computational Biology and Bioinformatics, Salamanca, Spain (October 2008)
- Invited distinguished speaker series, University of California, Riverside (June 2009)
- Invited speaker, Joint Statistical Meeting, Washington DC (August 2009)
- Invited speaker, University of Paris (September 2009)
- Invited speaker, Safra Program Distinguished Lectures Series, Tel Aviv University, Tel Aviv, Israel (March 2010)
- Invited plenary speaker, Clinical Genomic Analysis Workshop, IBM Research, Haifa, Israel (March 2010)
- Invited speaker, Hebrew University, Jerusalem, Israel (March 2010)
- Invited speaker, Bioinformatics Colloquium, UCLA (March 2010)
- Invited speaker, Computational Intelligence Methods for Data Analysis in Oncology Bioinformatics, Vietri sul Mare, Italy (May 2010)
- Director Erice Summer School in Bioinformatics (September 2010)
- Invited speaker, University of Evry and Genopole, Evry, France (September 2010)
- Invited distinguished speaker, University of Alberta, Alberta Ingenuity Centre for Machine Learning, Edmonton. Canada (October 2010)
- Invited speaker, Information Theory and Applications Workshop, UCSD (February 2010)
- Invited speaker, Information Theory and Applications Workshop, UCSD (February 2011)

- Invited speaker, Institute for Pure and Applied Mathematics (IPAM), UCLA, Program on Navigating Chemical Compound Space for Materials and Bio Design (March 2011)
- Invited speaker, Institute for Pure and Applied Mathematics (IPAM), UCLA, Program on Navigating Chemical Compound Space for Materials and Bio Design (April 2011)
- Invited speaker, Center of Genomic Regulation, Barcelona, Spain (June 2011)
- Invited speaker, Invited speaker Bioinformatics Research Centre, Aarhus University, Denmark (June 2011)
- Invited speaker, International Conference on Machine Learning (ICML) Workshop on Unsupervised and Transfer Learning, Seattle, Washington (July 2011)
- Invited keynote speaker, 2011 Pattern Recognition in Bioinformatics Conference (PRIB November 2011), Delft, Netherlands (November 2011)
- Invited speaker, Chemical & Biomolecular Engineering Colloquium, University of California Berkeley, (November 2011)
- Invited keynote speaker, NIPS workshop "From Statistical Genetics to Personalized Medicine", Granada, Spain, December 2011
- Invited speaker, Bioinformatics and Computational Biology Series, Iowa State University (ISU), Ames, Iowa, April 2012
- Invited lecturer, Lipari Summer School in Computational Biology and Bioinformatics on Pharmacogenomics, Lipari, Italy, July 2012
- Invited keynote speaker, JOBIM 2012 Conference, Rennes, France (July 2012).
- Invited keynote speaker, ACM Conference on Bioinformatics and Computational Biology (ACM-BCB 2012), Orlando, Florida, (October 2012)
- Invited keynote speaker, Third Immunoinformatics and Computational Immunology Workshop (ICIW 2012), Orlando, Florida, (October 2012)
- Invited speaker, Amazon (November 2012)
- Invited speaker, Deep Learning, MIT (December 2012)
- Invited speaker, Deep Learning, Toronto University (December 2012)
- Invited speaker, Deep Learning, Cornell University (March 2013)
- Invited speaker, Deep Learning, Caltech (April 2013) [IST Lunch Bunch talk]
- Invited speaker, Deep Learning, UCSD (May 2013)
- Invited speaker, Course on Bayesian Methods and Neural Networks, Univeristy of Padova, Italy, (May 2013)
- Invited speaker, Deep Learning, Carnegie Mellon University (CMU-Pitt PhD program in Computational Biology (CPCB)) August 2013
- Invited plenary speaker, 20-th anniversary of the Center for Biological Sequence Analysis, Technical University of Denmark (DTU), Copenhagen, Denmark (October 2013)
- Invited to the workshop on Learning Data Representation: Hierarchies and Invariance, McGovern Institute, MIT. Sponsored by the new Center for Brain Minds and Machines and the joint IIT-MIT Laboratory for Computational and Statistical Learning. MIT (November 2013)
- Invited speaker 2014 Tarragona International Summer School on Trends in Computing (July 2014)
- Invited speaker 2014 Summer School on RNA at Boston College (July 2014)
- Keynote speaker MLSB 2014 (Eighth International Workshop on Machine Learning in Systems Biology), Strasbourg, France, (September 2014)

- Invited speaker 2014 MLPM Summer School (Machine Learning for Personalized Medicine), Institut Curie, Paris, France, (September 2014)
- Invited speaker, NIPS Workshop on High-Energy Particle Physics, Machine Learning, and the HiggsML Data Challenge, Montreal, Canada, (December 2014)
- Invited speaker, Brigham Young University, UT in March 2015
- Keynote speaker joint session held by ICLR and AISTATS, May 9, 2015, in San Diego, CA
- Invited keynote speaker, Yandex Conference on Machine Learning in Applications, Berlin, Germany, October 2015.
- Invited speaker, Data Science Workshop, CERN, Switzerland, November 9-13, 2015.
- Invited speaker, Department of Bioinformatics, University of North Carolina, Charlotte, December 2015.
- Invited speaker, Chapman University, February 2016.
- Invited Speaker, San Diego State University, February 2016.
- Invited speaker, Distinguished Lecture Series, Computer Science Department, Wayne State University, March 22, 2016.
- Invited speaker, 1st Data Learning and Inference (DALI), Sestri Levante, Italy, March 30-April 1, 2016.
- Invited speaker, Scuola Normale Superiore, Pisa, Italy, April 1, 2016.
- Invited speaker, 99th Canadian Chemistry Conference and Exhibition, Halifax, Nova Scotia, June 7-9, 2016.
- Invited speaker, Computational & Systems Biology Seminar Series, UT Southwestern Medical Center, Dallas, Texas, June 12-13, 2016.
- Keynote speaker, 1st European Conference on Translational Bioinformatics, Copenhagen Denmark, April 25-27, 2016.
- Invited speaker, QBiC Symposium 2016, Senri Life Science Center, Osaka, Japan, September 5-7, 2016.
- Keynote speaker, Jubilee of Mercy for Professors, Symposium on the Technologies for Knowledge, Rome, Italy, September 7-11, 2016.
- Invited talk, University of Catanzaro, Italy, September 12, 2016.
- Keynote speaker, 22nd Internation Conference on Computing in High-Energy and Nuclear Physics, San Francisco, California, October 10-14, (2016).
- Keynote speaker, 15th IEEE International Conference on Machine Learning and Applications (ICMLA 2016), Anaheim, California, December 18-20, 2016.
- Invited lecture, Perelman School of Medicine, Department of Systems Pharmacology and Translational Therapeutics, University of Pennsylvania, Spring 2017.
- Keynote speaker, the 22nd International Conference on Computing in High Energy and Nuclear Physics (CHEP 2016), San Francisco, CA, October 10-14, 2016.
- Keynote speaker, 15th IEEEE International Conference on Machine Learning and Applications (ICMLA 2016), Anaheim, CA, Dec 18-20, 2016.
- Invited talk, Firmeninch, Geneva, January 11, 2017.
- Invited Keynote Speaker, Machine Learning Prague conference, April 2017.
- Invited talk, Director's Colloquium Speaker, Physics Colloquium, Argonne National Laboratory, April 28, 2017.
- Keynote Speaker, "Braverman's Readings in Learning Theory and Related Areas", Boston, MA, April 28-30, 2017.
- Invited lecturer, Deep Learning Summer School, Bilbao, Spain, July 2017.
- Keynote speaker, 40th German AI Conference, Dortmund, Germany, September 2017.

- Keynote speaker, 16th Mexican AI Conference, Ensenada, Mexico, October 2017.
- Invited speaker, Smale Institute, Hong Kong, July 2018.
- Invited lecturer, Deep Learning Summer School, Genova, Italy, July 2018.
- Invited Inaugural Keynote Speaker, 41st National Conference on Biomedical Engineering, Leon, Mexico, October 18-20, 2018.
- Keynote Speaker, 6th Bench to Bedside Symposium, Arnold and Mabel Beckman Center, UCI, March 9, 2019.
- Keynote Panelist, Road to Reinvention 2019: Leadership in the Digital Age, Arnold and Mable Beckman Center, UCI, March 21, 2019 .
- Speaker, Deep Learning Workshop, Copenhagen, Denmark, April 7-11, 2019.
- Invited speaker, AAAS Pacific Division, 100th Annual Meeting, Southern Oregon University, Ashland, Oregon, June 18 - 21, 2019.
- Invited Keynote Speaker, Public Lecture on Artificial Intelligence, DESY, Hamburg, Germany, July 7, 2019.
- Invited speaker, Institute for Mathematical Behavioral Sciences, UCI. AI, Deep Learning, and Virtualization. February 21, 2020.
- Keynote Speaker, 16th IEEE International Conference on Computational Intelligence in Bioinformatics and Computational Biology, Certosa di Pontignano, Italy, July 9-11, 2019.
- Invited Colloquium Spealer, Department of Mathematics, UCLA, March 5, 2020.
- Invited Colloquium Speaker, Department of Mathematics, USC, March 6, 2020.
- Invited Keynote Speaker, AI for Reaction Outcome and Synthetic Route Prediction Conference, March 8-10, 2020, Bristol, UK.
- Invited Keynote Speaker, Fifty-First Southeastern International Conference on Combinatorics, Graph Theory and Computing, Florida Atlantic University, Boca Raton, Florida, March 9-13, 2020.
- Invited Speaker, 90th Birthday Celebration for Stephe Smale, University of Michigan, Ann Arbor, July 12-15, 2020.
- Keynote Speaker, 3rd Advanced Course on Data Science & Machine Learning (ACDL 2020), Certosa di Pontignano (Siena) Tuscany, Italy, July 13-17, 2020.
- Invited Keynote Speaker, 6th International Conference on Machine Learning, Optimization & Data Science (LOD 2020), Certosa di Pontignano (Siena – Tuscany), Italy, July 19-22, 2020
- Invited Keynote Speaker, Workshop on "Biologically Plausible Learning", Satellite Workshop at the 6th International Conference on Machine Learning, Optimization & Data Science (LOD 2020), Certosa di Pontignano, Siena, Italy, July 19, 2020.
- Invited Keynote Speaker, Lipari School on Computational Life Sciences, Lipari, Italy, July 25-August 1, 2020.
- Deep Learning Track Chair, 19th Mexican International Conference on Artificial Intelligence (MICAI 2020) , Mexico City, Mexico, October 12 to 17, 2020.
- Invited Keynote Speaker, National Colloquium on Artificial Intelligence, Mexico, Oct. 28th, 2020.
- Invited speaker, Meeting of the American Academy Society of Spine Radiologists (ASSR 2021), February 20, 2021.
- Invited speaker (Attention Mechanisms for Machine Learning in Physics), CERN Workshop, April 15, 2021.
- Invited speaker, NIH Workshop on Reaction Informatics, May 17-21, 2021.
- Invited Plenary Speaker, the 16th International Work-Conference on Artificial Neural Networks, IWANN 2021, June 16, 2021.

- Invited Keynote Speaker, 4th Advanced Course on Data Science & Machine Learning (ACDL 2021), Certosa di Pontignano (Siena) Tuscany, Italy, July 19--23, 2021.
- Invited Keynote Speaker, Conference on Mathematics of Machine Learning, Center for Interdisciplinary Research (ZIF), Bielefeld, Germany, August 4—7, 2021.
- Co-Chair workshop on "Multi-omics data integration for modeling biological systems", 30th ACM International Conference on Information and Knowledge Management (CIKM2021), Queensland, Austalia, November 1-5, 2021.
- Invited Keynote Speaker, 18th IEEE International Conference in Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Melbourne, Australia, October 13-15, 2021.
- Invited Keynote Speaker, The BIOMAT 2021 International Symposium on Mathematical and Biological Sciences. November 1 -5, 2021.
- Invited Speaker, Bioinformatics Seminar Series, Purdue University, Nov 18, 2021.
- Invited Keynote Speaker, AI and Bio-Medical Imaging, Foro Internacional de Tuberculosis, December 2, 2021.
- Invited Keynote Speaker, 5th Northern Lights Deep Learning Conference, 10-12 January 2022, Tromsø ("North Pole"), Norway.
- Invited Speaker, Institute for Applied Computational Science (IACS) Distinguished Lecture Series, Harvard University, March 4, 2022.
- Invited speaker, 100th Anniversary of the Italian Mathematical Union (UMI) and 800th Anniversary of the University of Padova, Padova, Italy, May 23-27, 2022.
- Invited speaker, Computer Laboratory, Cambridge University, (2022).
- Invited Keynote Speaker, Lipari School on Computational Life Sciences, Lipari, Italy, July 24-30, 2022.
- Invited Keynote Lecturer, Deep Learn Summer School 2022, Las Palmas de Gran Canaria, Spain, July 25-29, 2022

Editorial Activities (Abridged)

- Neural Networks (Member Editorial Board, Associate Editor) (1995-)
- Data Mining and Knowledge Discovery (Associate Editor) (2006-2021)
- International Journal of Bioinformatics Research and Applications (Member Editorial Board) (2003-2008)
- Data Mining and Knowledge Discovery (Member Editorial Board) (2002-)
- Chemistry Central Journal (Member Editorial Board, Founding Editor) (2006-2010)
- Journal of Chemical Information and Modeling (Member Editorial Board) (2008-2011)
- IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) (Associate Editor) (2013-)
- Artificial Intelligence Journal, Associate Editor (2017-).

Reviewer (Abridged)

- Reviewer for NIH, NSF, and other agencies (Portuguese Science Foundation, Italian Science Foundation, etc)
- Standing member of the NIH Biodata Management and Analysis Study Section [BDMA] with a six-year term [7/1/2012 through 8/16/2018]. Chair of the Section starting in July 2014.

- Reviewer for many journals including Science, Science Advances, Nature, Nature Communications, Nature Machine Intelligence, Physical Review Letters, IEEE Transactions on Information Theory, IEEE/ACM Transactions on Computational Biology and Bioinformatics, IEEE Transactions on Neural Networks, Neural Computation, Artificial Intelligence, Data Mining and Knowledge Discovery, Journal of Computational Biology, Journal of Molecular Biology, Bioinformatics, Proteins, Genome Research, Nucleic Acids Research, , CABIOS.

FUNDING

Grants and Gifts (available upon request)

Paid Consultancies and Reviews (Abridged)

Banque de France; TI Capital ; SmithKline-Beecham; Paracel; DE Shaw; Calspan; Ecobalance/Dames and Moore; Colosseum Fund; Banca del Salento; Elitra; Allergan; Genome Canada; San Antonio Life Sciences Institute; Fundacion BBVA, Spain; Italian Ministry for Education University and Research (MIUR); Pennsylvania Department of Health's Interim Performance Review; MIT Press; Cambridge University Press; Wiley; Harvard University; Mitre Corporation; Center for Biological Sequence Analysis, DTU, Denmark; Max Planck Institute for Biological Cybernetics, Germany; Christian Doppler Forschungsgesellschaft, Austria; Danish Agency for Science, Technology and Innovation; Quest Diagnostics; Strategy and Funding Directorate at the Health Research Board, Ireland; Hitchcock Foundation (Dartmouth); Pimco; Trond Mohn Foundation, Norway; Deep Radiology, Anivive Lifesciences.

Entrepreneurial Activities (Abridged)

- Co-founder of Net-ID, Inc. in 1991.
- Founder of IDLAB in 2006 (royalties and consulting).
- Co-founder of Reaction Explorer, LLC in 2010.
- Helped the early development of Verdezyne, Antigen Discovery, DocBot, and Anivive.

Society Memberships

- Member and Elected Fellow of the American Association for the Advancement of Science (AAAS
- Member and Elected Fellow of the Association for the Advancement of Artificial Intelligence (AAAI)
- Member and Elected Fellow of the Institute of Electrical and Electronic Engineers (IEEE)
- Member of the American Chemical Society (ACS)
- Member and Elected Fellow of the Association for Computing Machinery (ACM)
- Member and Elected Fellow of the International Society for Computational Biology (ISCB)

Mentoring

Current PhD Students, Programmers, and Postdoctoral Fellows

- Sherif Abdelkarim
- Siwei Chen
- Ekaterina Deyneka
- Roman Drai
- John B. Lanier
- Junze Liu
- Jordan Ott
- Muntaha Samad
- Alexander Shmakov
- Amin Tavakoli

Graduated PhD Students, Postdoctoral Fellows, and Researchers (Primary Advisor, Past 15 years)

- Forest Agostinelli, University of South Carolina
- Alessio Andronico, University of Paris VI
- Chloe Azencott, Mines ParisTech, Institut Curie, and INSERM
- Kevin Bache, Google
- Pierre-François Baisnée, Institut de Recherche pour le Dévelopment, France
- Ryan Benz, Applied Proteomics
- Martin Brandon, ADNET Systems
- Andrew Brethorst, The Aerospace Corporation
- Nicholas Ceglia, Memorial Sloan Kettering Cancer Center
- Ivan Chang, University of California, Irvine
- Jonathan Chen, Stanford University
- Jianlin Cheng, University of Missouri
- Julian Collado, Blackberry-Cylance
- Kenneth Daily, Amazon
- Pietro Di Lena, University of Bologna and Cesena, Italy
- Yimeng Dou, Verdezyne
- David Fooshee, NeoGenomics
- Steven Hampson, Deceased
- Lars Hertel, Linkedin
- Qian-Nan Hu, Shanghai Institutes for Biological Sciences
- Raja Jurdak, CSIRO, Brisbane, Australia
- Matt Kayala, Google
- Christine Lee, Edwards Life Science
- Lingge Li, Facebook
- Erik Linstead, Chapman University
- Yu Liu, Google
- Yadong Lu, Microsoft
- Alessandro Lusci, Viant Inc.
- Christophe Magnan, NeoGenomics
- Antonio Maratea, University of Naples-Parthenope, Italy

- Stephen McAleer, Carnegie Mellon University
- Ken Nagata, Google
- Ramzi Nasr, comScore
- Vishal Patel, The Retail Equation
- Gianluca Pollastri, University College Dublin, Ireland
- Liva Ralaivola, University of Provence/Aix-Marseille I, France
- Arlo Randall, Antigen Discovery
- Paul Rigor, Osprey Data
- Alex Sadovsky, Defense Logistics Agency
- Peter Sadowski, University of Hawaii
- Suman Sundaresh, Linkedin
- S. Joshua Swamidass, Washington University, Saint Louis
- Mike Sweredoski, Caltech
- Gregor Urban, Google
- Eric Wang, Tandem Diagnostics
- Lin Wu, Kunming University of Science and Technology
- Michael Zeller, Google

UCI Academic Service (available upon request)

Other Activities (Available Upon Request)