

S. Joshua Swamidass

<http://swami.wustl.edu/>

Research Interests

Statistical machine learning and decision theory applied to chemical biology and medicine, with special interest in drug metabolism, drug discovery, and translational research.

Appointments

2011- <i>now</i>	Assistant Professor	Washington University Department of Immunology and Pathology Division of Laboratory and Genomic Medicine
2017- <i>now</i>	Faculty Lead	Washington University Translational Bioinformatics Institute for Informatics
2011-2012	Visiting Scientist	Pfizer Incorporated Indications Discovery Unit
2010-2011	Instructor	Washington University Department of Immunology and Pathology Division of Laboratory and Genomic Medicine
2009-2012	Visiting Scientist	Broad Institute of Harvard and MIT Chemical Biology Program
2009-2010	Laboratory and Genomic Medicine Resident	Washington University Department of Immunology and Pathology Division of Laboratory and Genomic Medicine

Education

2000-2009	Medical Doctorate (M.D.) Graduation with Distinction in Research	University of California, Irvine
2002-2007	Philosophical Doctorate (Ph.D.) Information and Computer Sciences Specialization in Informatics in Biology and Medicine	University of California, Irvine
2002-2006	Master of Sciences (M.S.) Information and Computer Sciences	University of California, Irvine
1996-2000	Bachelor of Sciences (B.S.) Major in Biological Sciences, <i>Cum Laude</i> Minor in Information and Computer Sciences	University of California, Irvine

Funding

2016-2020	Computationally Modeling the Impact of Ontogeny on Drug Metabolic Fate	Library of Medicine R01 LM012482-01
2016-2020	Data and Tools for Modeling Metabolism and Reactivity	Library of Medicine R01 LM012222-01
2011-2014	Computational Biology	Glaxo-Smith-Kline
2011-2012	Indications Discovery Unit	Pfizer Incorporated
2009-2010	Physician Scientist Training Program	Washington University Department of Pathology and Immunology
2000-2009	M.D. Ph.D. Program	University of California, Irvine Funded by National Institute of Health

2005-2007	Biomedical Informatics Training Program	University of California, Irvine Funded by National Institute of Health
2001-2004	Harvey Fellowship	The Mustard Seed Foundation
1996-2000	Regents Scholarship	University of California Regents

Service

2017- <i>now</i>	NIH Review Committee Member (BLIRC)	National Library of Medicine
2016- <i>now</i>	Associate Editor	BMC Medical Informatics and Decision Making
2016-2017	NIH Review Committee Member (XNDA)	National Institute of General Medical Sciences
2015-2016	Science Advisor, Concordia Seminary	American Association for the Advancement of Science Science for Seminaries Program
2015-2016	NIH Review Committee Member (ZLM1 ZH-C)	National Library of Medicine

Publications

Swamidass, S. J. *Chemical Databases, Datastructures, and Algorithms*. Ph.D. dissertation, University of California, Irvine, 2007.

Refereed Journal Articles

48. Ruan, S., **S. J. Swamidass**, and G. D. Stormo. BEESEM: Estimation of binding energy models using HT-SELEX data. *Bioinformatics*, 2017.
47. Dang, N. L., T. B. Hughes, G. P. Miller, and **S. J. Swamidass**. Computational approach to structural alerts: Furans, phenols, nitroaromatics, and thiophenes. *Chemical Research in Toxicology*, 2017.
46. Hughes, T. B. and **S. J. Swamidass**. Deep learning to predict the formation of quinone species in drug metabolism. *Chemical Research in Toxicology*, 30(2):642–656, 2017.
45. Kumar, R. D., **S. J. Swamidass**, and R. Bose. Unsupervised cancer driver detection with parsimony-guided learning. *Nature Genetics*, 48:1288–1294, 2016.
44. Pan, Y., T. Daito, Y. Sasaki, Y. H. Chung, X. Xing, S. Pondugula, **S. J. Swamidass**, T. Wang, A. H. Kim, and H. Yano. Inhibition of dna methyltransferases blocks mutant huntingtin-induced neurotoxicity. *Scientific Reports*, 6:31022, 2016.
43. Hughes, T. B., N. L. Dang, G. P. Miller, and **S. J. Swamidass**. Modeling reactivity to biological macromolecules with a deep multitask network. *ACS Central Science*, 2(8):529–537, 2016.
42. Van Voorhis, W. C., ..., N. L. Dang, ..., T. B. Hughes, ..., **S. J. Swamidass**, ..., and P. Willis. Open source drug discovery with the malaria box compound collection for neglected diseases and beyond. *PLoS Pathogens*, 12(7):e1005763, 2016.
41. Le Dang, N., T. B. Hughes, V. Krishnamurthy, and **S. J. Swamidass**. A simple model predicts UGT-mediated metabolism. *Bioinformatics*, 2016.
40. Lusci, A., M. Browning, D. Fooshee, **S. J. Swamidass**, and P. Baldi. Accurate and efficient target prediction using a potency-sensitive influence-relevance voter. *Journal of Cheminformatics*, 7(1):63, 2015.
39. Kumar, R. D., A. C. Searleman, **S. J. Swamidass**, O. L. Griffith, and R. Bose. Statistically identifying tumor suppressors and oncogenes from pan-cancer genome-sequencing data. *Bioinformatics*, 31(22):3561–3568, 2015.
38. Hughes, T. B., G. P. Miller, and **S. J. Swamidass**. Modeling epoxidation of drug-like molecules with a deep machine learning network. *ACS Central Science*, 1(4):168–180, 2015.
37. Zaretski, J., K. M. Boehm, and **S. J. Swamidass**. Improved prediction of cyp-mediated metabolism with chemical fingerprints. *Journal of Chemical Information and Modeling*, 55(5):972–982, 2015.
36. Li, J., S. Zheng, B. Chen, A. J. Butte, **S. J. Swamidass**, and Z. Lu. A survey of current trends in computational drug repositioning. *Briefings in Bioinformatics*, 1:11, 2015.
35. Hughes, T. B., G. P. Miller, and **S. J. Swamidass**. Site of reactivity models predict molecular reactivity of diverse chemicals with glutathione. *Chemical Research in Toxicology*, 28(4):797–809, 2015.

* Co-corresponding author.

34. Zaretski, J. M., M. R. Browning, T. B. Hughes, and **S. J. Swamidass**. Extending p450 site-of-metabolism models with region-resolution data. *Bioinformatics*, 31(12):1966–1973, 2015.
33. Diamond, M. I., S. Cai, A. Boudreau, C. J. Carey, N. Lyle, R. V. Pappu, **S. J. Swamidass**, M. Bissell, H. Piwnicka-Worms, and J. Shao. Subcellular localization and ser-137 phosphorylation regulate tumor-suppressive activity of profilin-1. *Journal of Biological Chemistry*, 290(14):9075–9086, 2015.
32. **Swamidass, S. J.**, M. Matlock, and L. Rozenblit. Securely measuring the overlap between private datasets with cryptosets. *PLOS One*, 10(2):e0117898, 2015.
31. Matlock, M. K., T. B. Hughes, and **S. J. Swamidass**. Xenosite server: a web-available site of metabolism prediction tool. *Bioinformatics*, 31(7):1136–1137, 2015.
30. Ekins, S., A. M. Clark, **S. J. Swamidass**, N. Litterman, and A. J. Williams. Bigger data, collaborative tools and the future of predictive drug discovery. *Journal of Computer-Aided Molecular Design*, pages 1–12, 2014.
29. **Swamidass, S. J.**, C. N. Schillebeeckx, M. Matlock, M. R. Hurle, and P. Agarwal. Combined analysis of phenotypic and target-based screening in assay networks. *Journal of Biomolecular Screening*, 19(5):782–790, 2014.
28. Matlock, M. and **S. J. Swamidass**. Sharing chemical relationships does not reveal structures. *Journal of Chemical Information and Modeling*, 54(1):37–48, 2013.
27. **Swamidass, S. J.**, Z. Lu, P. Agarwal, and A. J. Butte. Computational approaches to drug repurposing and pharmacology-session introduction. In *Pacific Symposium on Biocomputing*. *Pacific Symposium on Biocomputing*, volume 19, pages 110–113, 2013.
26. Zaretski, J., M. Matlock, and **S. J. Swamidass**. XenoSite: Accurately predicting CYP-mediated sites of metabolism with neural networks. *Journal of Chemical Information and Modeling*, 53(12):3373–3383, 2013.
25. **Swamidass, S. J.** Using economic optimization to design high-throughput screens. *Future Medicinal Chemistry*, 5(1):9–11, 2013.
24. Matlock, M. K., J. M. Zaretski, and **S. J. Swamidass**. Scaffold network generator: a tool for mining molecular structures. *Bioinformatics*, 29(20):2655–2656, 2013.
23. Browning, M. R., B. T. Calhoun, and **S. J. Swamidass**. Managing missing measurements in small-molecule screens. *Journal of Computer-Aided Molecular Design*, pages 1–10, 2013.
22. Zaretski, J., C. Bergeron, T.-w. Huang, P. Rydberg, **S. J. Swamidass**, and C. M. Breneman. Rs-webpredictor: a server for predicting cyp-mediated sites of metabolism on drug-like molecules. *Bioinformatics*, 29(4):497–498, 2013.
21. Sloutsky, R., N. Jimenez, **S. J. Swamidass**, and K. M. Naegle. Accounting for noise when clustering biological data. *Briefings in bioinformatics*, 14(4):423–436, 2013.
20. Calhoun, B., M. Browning, B. Chen, J. Bittker, and **S. J. Swamidass**. Automatically detecting workflows in PubChem. *Journal of Biomolecular Screening*, 17(8):1071–1079, 2012.
19. **Swamidass, S. J.**, B. T. Calhoun, J. A. Bittker, N. E. Bodycombe, and P. A. Clemons. Utility-aware screening with clique-oriented prioritization. *Journal of Chemical Information and Modeling*, 52(1):29–37, 2012.
18. Ranu, S., B. Calhoun, A. Singh, and **S. J. Swamidass**. Probabilistic substructure mining from small-molecule screens. *Molecular Informatics*, 30(9):809–815, 2011.
17. **Swamidass, S. J.** Mining small-molecule screens to repurpose drugs. *Briefings in Bioinformatics*, 12(4):327–335, 2011.
16. [†]**Swamidass, S. J.**, B. T. Calhoun, J. A. Bittker, N. E. Bodycombe, and P. A. Clemons. Enhancing the rate of scaffold discovery with diversity-oriented prioritization. *Bioinformatics*, 27(16):2271–2278, 2011.
15. [‡]**Swamidass, S. J.**, J. A. Bittker, N. E. Bodycombe, S. P. Ryder, and P. A. Clemons. An Economic Framework to Prioritize Confirmatory Tests after a High-Throughput Screen. *Journal of Biomolecular Screening*, 15(6):680, 2010.
14. **Swamidass, S. J.**, C. A. Azencott, K. Daily, and P. Baldi. A CROC stronger than ROC: measuring, visualizing and optimizing early retrieval. *Bioinformatics*, 26:1348–1356, 2010.
13. Nasr, R., **S. J. Swamidass**, and P. F. Baldi. Large scale study of multiple-molecule queries. *Journal of Cheminformatics*, 1(1):7, 2009.
12. **Swamidass, S. J.**, C. A. Azencott, T. W. Lin, H. Gramajo, S. C. Tsai, and P. Baldi. Influence Relevance Voting: an accurate and interpretable virtual high throughput screening method. *Journal of Chemical Information and Modeling*, 49:756–766, 2009.

* Corresponding author. [†] Corresponding author. [‡] Corresponding author. [§] Co-first author.

11. Benz, R. W., **S. J. Swamidass**, and P. Baldi. Discovery of power-laws in chemical space. *Journal of Chemical Information and Modeling*, 48:1138–1151, 2008.
10. Baldi, P., R. W. Benz, D. S. Hirschberg, and **S. J. Swamidass**. Lossless compression of chemical fingerprints using integer entropy codes improves storage and retrieval. *Journal of Chemical Information and Modeling*, 47:2098–2109, 2007.
9. Chen, J. H., E. Linstead, **S. J. Swamidass**, D. Wang, and P. Baldi. ChemDB update—full-text search and virtual chemical space. *Bioinformatics*, 23:2348–2351, 2007.
8. **Swamidass, S. J.** and P. Baldi. Mathematical correction for fingerprint similarity measures to improve chemical retrieval. *Journal of Chemical Information and Modeling*, 47:952–964, 2007.
7. Azencott, C. A., A. Ksikes, **S. J. Swamidass**, J. H. 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*, 47:965–974, 2007.
6. **Swamidass, S. J.** and P. Baldi. Bounds and algorithms for fast exact searches of chemical fingerprints in linear and sublinear time. *Journal of Chemical Information and Modeling*, 47:302–317, 2007.
5. Danziger, S. A., **S. J. Swamidass**, J. Zeng, L. R. Dearth, Q. Lu, J. H. Chen, J. Cheng, V. P. Hoang, H. Saigo, R. Luo, P. Baldi, R. K. Brachmann, and R. H. Lathrop. Functional census of mutation sequence spaces: the example of p53 cancer rescue mutants. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 3:114–125, 2006.
4. Lin, T. W., M. M. Melgar, D. Kurth, **S. J. Swamidass**, J. Purdon, T. Tseng, G. Gago, P. Baldi, H. Gramajo, and S. C. Tsai. Structure-based inhibitor design of AccD5, an essential acyl-CoA carboxylase carboxyltransferase domain of Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, 103:3072–3077, 2006.
3. Chen, J., **S. J. Swamidass**, Y. Dou, J. Bruand, and P. F. Baldi. ChemDB: a public database of small molecules and related cheminformatics resources. *Bioinformatics*, 21:4133–4139, 2005.
2. Ralaivola, L., **S. J. Swamidass**, H. Saigo, and P. F. Baldi. Graph kernels for chemical informatics. *Neural Networks*, 18:1093–1110, 2005.
1. **Swamidass, S. J.**, J. Chen, J. Bruand, P. Phung, L. Ralaivola, and P. F. Baldi. Kernels for small molecules and the prediction of mutagenicity, toxicity and anti-cancer activity. *Bioinformatics*, 21 Suppl 1:i359–368, 2005.

Editorials and Interviews

6. **Swamidass, S. J.** Education: Initiatives to bridge faith and science. *Nature*, 523(7562):531, 2015.
5. Kalgutkar, A. S. Tracking where the o's go. *ACS Central Science*, 1(4):163–165, 2015.
4. Zeliadt, N. Cryptographic methods enable analyses without privacy breaches. *Nature medicine*, 20(6):563–563, 2014.
3. **Swamidass, S. J.** Rubio and the age-of-the-earth question. *Wall Street Journal*, 2012.
2. **Swamidass, S. J.** Interview: Academia's role in drug repositioning multi-stakeholder collaboration. <http://www.terrapiinn.com/template/live/documents.aspx?e=5612&d=8864>.
1. Boldrin, M. and **S. J. Swamidass**. A new bargain for drug approvals. *Wall Street Journal*, 2011.

Conference Talks

26. Hughes, T. B. and **S. J. Swamidass**. An integrated model of metabolism and reactivity. In *The 21st International Symposium on Microsomes and Drug Oxidations*, 2016.
25. Hughes, T. B., N. L. Dang, and **S. J. Swamidass**. Modeling reactive metabolite formation of drug-like molecules with machine learning. In *the 20th North American ISSX Meeting*. International Society for the Study of Xenobiotics, 2015.
24. **Swamidass, S. J.**, N. L. Dang, and T. Hughes. Enhancing structural alerts for toxicity with mechanism-based metabolism and reactivity models. In *the 250th American Chemical Society National Meeting*. American Chemical Society, 2015.
23. **Swamidass, S. J.**, G. P. Miller, and N. L. Dang. Very high accuracy prediction of UDP-glucuronosyltransferase sites of metabolism. In *Experimental Biology*. The American Society for Pharmacology and Experimental Therapeutics, 2015.
22. **Swamidass, S. J.**, G. P. Miller, N. L. Dang, J. Hartman, and S. Cothren. Do chiral fingerprints and descriptors work? In *the 249th American Chemical Society National Meeting*. American Chemical Society, 2015.
21. Hughes, T., G. P. Miller, and **S. J. Swamidass**. Evaluating structural toxicity alerts with metabolism and reactivity models. In *the 249th American Chemical Society National Meeting*. American Chemical Society, 2015.

20. **Swamidass, S. J.** Modeling the reactivity of drug metabolites. In *the 14th CUP Meeting*. OpenEye, 2014.
19. **Swamidass, S. J.** Modeling the reactivity of drug metabolites. In *the 19th North American ISSX Meeting*. International Society for the Study of Xenobiotics, 2014.
18. **Swamidass, S. J.**, J. M. Zaretski, and M. Matlock. Training cytochrome p450 site of metabolism models with region-level data. In *the 247th American Chemical Society National Meeting*. American Chemical Society, 2014.
17. **Swamidass, S. J.** and M. Matlock. Sharing chemical relationships does not reveal structures. In *the 247th American Chemical Society National Meeting*. American Chemical Society, 2014.
16. Zaretski, J. M., M. K. Matlock, and **S. J. Swamidass**. P450 site of metabolism predictions using specialized atomic and molecular descriptors with neural network modeling. In *the 246th American Chemical Society National Meeting*. American Chemical Society, 2013.
15. **Swamidass, S. J.**, M. Matlock, and D. Agrafiotis. Sharing chemical information from screens without revealing structures. In *the 245st American Chemical Society National Meeting*. American Chemical Society, 2013.
14. **Swamidass, S. J.** Phenotypic screening data to suggest repurposing opportunities at an early stage. In *World Drug Repositioning Congress*, Washington DC, 2012.
13. **Swamidass, S. J.** Globally inferring targets from phenotypic small-molecule screens. In *Translational Bioinformatics Conference*, Korea, 2012.
12. **Swamidass, S. J.** Mining small-molecule screens to repurpose drugs. In *the 244st American Chemical Society National Meeting*. American Chemical Society, 2012.
11. **Swamidass, S. J.** and B. T. Calhoun. High accuracy polypharmacology models for large datasets. In *the 243st American Chemical Society National Meeting*. American Chemical Society, 2012.
10. **Swamidass, S. J.** and B. T. Calhoun. Using economic optimization to design high throughput screens. In *the 243st American Chemical Society National Meeting*. American Chemical Society, 2012.
9. **Swamidass, S. J.**, R. Sayan, B. T. Calhoun, and A. Singh. Probabilistic substructure mining from small-molecule screens. In *the 242st American Chemical Society National Meeting*. American Chemical Society, 2011.
8. **Swamidass, S. J.** and B. Calhoun. The anatomy of a pubchem screen. In *the 241st American Chemical Society National Meeting*. American Chemical Society, 2011.
7. **Swamidass, S. J.** and B. Calhoun. Utility-aware virtual high-throughput screening. In *the 241st American Chemical Society National Meeting*. American Chemical Society, 2011.
6. Baldi, P., C. Azencott, and **S. J. Swamidass**. Bridging the gap between neural network and kernel methods: Applications to drug discovery. In *Proceeding of the 2011 conference on Neural Nets WIRN10: Proceedings of the 20th Italian Workshop on Neural Nets*, pages 3–13, Amsterdam, The Netherlands, The Netherlands, 2011. IOS Press.
5. [¶]**Swamidass, S. J.**, J. A. Bittker, N. E. Bodycombe, S. P. Ryder, and P. A. Clemons. Marginal cost of discovery: an economic alternative to false discovery rate. In *the Proceedings of the 2010 Conference on Intelligent Systems for Molecular Biology*, 2010.
4. Benz, R. W., **S. J. Swamidass**, and P. Baldi. Discovery of power-laws in chemical space. In *the 235th American Chemical Society National Meeting*. American Chemical Society, 2008.
3. Chen, J. H., E. Linstead, **S. J. Swamidass**, D. Wang, Y. Dou, and P. F. Baldi. ChemDB: A public database of small molecules and related cheminformatics resources. In *the 235th American Chemical Society National Meeting*. American Chemical Society, 2007.
2. Danziger, S. A., **S. J. Swamidass**, J. Zeng, L. R. Dearth, Q. Lu, J. Chen, J. Cheng, V. Hoang, H. Saigo, R. Luo, P. F. Baldi, R. K. Brachmann, and R. H. Lathrop. Functional census of mutation sequence spaces: the example of p53 cancer rescue mutants. In *the Proceedings of the IEEE/ACM*, 2006.
1. **Swamidass, S. J.**, J. Chen, J. Bruand, P. Phung, L. Ralaivola, and P. F. Baldi. Kernels for small molecules and the prediction of mutagenicity, toxicity and anti-cancer activity. In *the Proceedings of the 2005 Conference on Intelligent Systems for Molecular Biology*, 2005.

Patents

2. **Swamidass, S. J.** and P. F. Baldi. Method for rapidly and accurately searching chemical compounds. U.S. Patent Application 60/970,213, 2007.
1. Baldi, P. F. and S. Tsai. Method of treatment for Mycobacterium tuberculosis. W.I P.O. Patent WO/2007/092930, 2007.^{||}

[¶] Corresponding author. ^{||} Work noted in this *curriculum vitae* led to this patent.

Book Chapter

1. **Swamidass, S. J.** and P. F. Baldi. Statistical distribution of chemical fingerprints, volume 3849. Lecture Notes in Computer Science, 2006.

Invited Talks

Only recent examples included below...

- Mar, 2017 Department of Chemistry, Stanford, Palo Alto CA
- Mar, 2017 Gladstone Institutes, UC San Francisco, San Francisco CA
- Sep, 2016 Department of Biomedical Informatics, University of Pittsburg, Pittsburg PA
- Jul, 2016 Science for Seminaries Retreat, AAAS, Portland OR
- May, 2016 Department of Pharmacology, University of Washington, Seattle WA
- May, 2016 CADD Forum, Janssen Research & Development, Newark NJ
- May, 2016 Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Science, Little Rock AR
- May, 2016 National Center for Toxicological Research, FDA, Little Rock AR
- Mar, 2016 Division of Gastroenterology, Cedars-Sinai Medical Center, Los Angeles CA
- Oct, 2015 Science Seminar Series, Wheaton University, Wheaton IL
- Oct, 2015 Medical School, University of Chicago, Chicago IL
- Mar, 2015 Department of Biomedical Engineering, Oregon Health Science University, Portland OR

Teaching

2016- <i>now</i>	Lecturer	The Digital Society
2011-2015	Lecturer Genomic Statistics	Laboratory and Genomic Medicine Genomics Rotation
2011	Lecturer Classification	Biomedical Engineering 572 Biological Neural Computation
2010	Lecturer Kernel Methods	Biology 5495 Computational Molecular Biology
2003-2005	Lecturer Sequence Analysis and Chemical Informatics	Developmental Biology 203A Biology Basic Training
2005	Lecturer Biologic and Artificial Neural Networks	Information and Computer Sciences 173 Introduction to Artificial Intelligence
2002	Lecturer and Tutor	Medical Biochemistry
1999	Teaching Assistant	Undergraduate General Chemistry

April 23, 2017