David L. Wild
Associate Professor of Computational Biology and Director of Computing
Keck Graduate Institute of Applied Life Sciences
535 Watson Drive
Claremont, California 91711
(909) 607-8566 (work)
(626) 396-9901 (home)
david_wild@kgi.edu

Education

1982	D. Phil. Molecular Biophysics, University of Oxford, UK
1977	M. Phil. Biophysics, Council for National Academic Awards, UK
1974	B.A. (Hons.) Physics, University of York, UK

Professional Experience

1999- present	Founding Faculty, Associate Professor of Computational Biology and Director of Computing, Keck Graduate Institute of Applied Life Sciences, Claremont, CA
1999- present	Plenary Faculty, Claremont Graduate University, Claremont, CA
1997 -1999	Senior Bioinformatics Analyst, Oxford Molecular and Bioinformatics Group, Glaxo Wellcome Medicines Research Centre, Stevenage, UK
1994 -1997	Senior Staff Scientist, Structural Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA
1989-1994	Staff Scientist, Biological Structures and Biocomputing Programme, European Molecular Biology Laboratory, Heidelberg, Germany
1988-1989	Senior Research Scientist, Allelix Biopharmaceuticals Inc., Mississauga, Canada
1987-1988	Research Scientist, Allelix Biopharmaceuticals Inc.
1984-1987	Staff Scientist, European Molecular Biology Laboratory, Grenoble, France
1982-1984	Research Fellow, Department of Applied Biochemistry and Food Science, University of Nottingham, UK
1980-1982	Computing Assistant, Department of Metallurgy and Science Materials, University of Oxford, UK
1974-1976	Research Assistant, Manchester Polytechnic, UK

Visiting Appointments

2004	Senior Fellow, Institute of Pure and Applied Mathematics, UCLA
2002-present	Visitor, Gatsby Computational Neuroscience Unit , University College, London, UK
2001	Visitor, European Bioinformatics Institute, Hinxton, UK

Honors and Awards

1981	Agricultural and Food Research Council, UK - Research Fellowship
1976	Medical Research Council, UK - Postgraduate Studentship

Current Research Support

Awarded

Protein Structure Initiative (Structural Genomics) NIH/NIGMS
P01 GM63208: Tools & Data Resources in Support of Structural Genomics
Project III: Machine Learning Techniques Applied to Protein Fold and Remote Homology Recognition (David Wild, Co-PI) 04/01/02 – 03/31/07, $766,098
The major goal of this project is to develop novel advanced machine learning techniques and apply them to protein fold classification and remote homology recognition, with the aim of improving methods for remote homologue detection to aid the selection of potential targets for structure determination and for classifying newly determined structures in structural genomics projects

Emerging Models and Technologies for Computation, National Science Foundation
CCF-0524331: Reconstructing Metabolic and Transcriptional Networks using Bayesian State Space Models (David Wild, PI) 07/16/05 – 30/06/08, $400,000
The overall goal of this multidisciplinary project is to combine functional genomics and computational modeling into a novel integrative systems approach aimed at identifying key components of the regulatory networks involved in cell physiology. Our proposal aims to develop a computational framework based on a probabilistic modeling technique (Bayesian state space models), to reconstruct transcriptional and metabolic networks representative of the response of E. coli to acid stress.

Sun Academic Equipment Grant
6/15/02-6/30/05 $49,995 V-880 server
Database and Beowulf-type cluster for enhanced genome annotation

Alfred P. Sloan Foundation
Professional Masters Degrees in Bioinformatics/Computational Molecular Biology
01/01/00-12/31/00 $130,000 (Participant) Grant to develop computational biology curriculum at KGI.

Under Review

NIH Bioengineering Research Partnerships
Modeling gene regulatory networks in solar elastosis
Approx $3,000,000 over 5 years
The overall goal of this multidisciplinary project is to combine functional genomics and computational modeling into a novel integrative systems approach aimed at identifying new and functionally relevant gene regulatory networks important to the pathobiology of aging human skin. To understand the complex functional interactions between these intracellular biochemical pathways, altered changes in the deposition of extracellular proteins and the complex phenotype of human dermal aging, we have proposed the development of a novel probabilistic Bayesian state-space modeling approach which we hypothesize will allow us to identify the causal relationships between the regulatory intracellular networks that result in this multifactorial human disorder.

Professional Activities

Editorial Board, *Current Bioinformatics*, 2005
Organizing Committee, Institute for Pure and Applied Mathematics, UCLA, Proteomics Reunion Conference, 2005
Program Committee, Intelligent Systems in Molecular Biology, 2005, Detroit.
NIH Study Section BDMA (Biodata Management and Analysis), 2004.
Organizing Committee, Institute for Pure and Applied Mathematics, UCLA, Program on Proteomics, 2004
NIH Study Section SSSH-90 (Computational Biology), 2003.
Reviewing Committee, 3rd International Conference on Systems Biology, 2002, Stockholm.
Associate Guest Editor IEEE Intelligent Systems. Special Issue, "Intelligent Systems in Biology," November/December 2001; March/April 2002.

Tutor, Microarray Bioinformatics course, Oxford University, Centre for Professional Development, 2001-2003

1989-1994 Collaborative Computational Project (CCP4 – Protein Crystallographic Software), Working Group 2

Referee for *Bioinformatics, Protein Engineering, Phys. Rev. E, Proc. Roy. Soc. B., IEEE Transactions on Computational Biology & Bioinformatics, BMC Bioinformatics, Nucleic Acids Review, FEBS Letters, Intelligent Systems in Molecular Biology, Pacific Symposium on Biocomputing*

Consultant for Celltech-Chiroscience, 2001

Member of International Society for Computational Biology
Member of Molecular Graphics and Modeling Society
Member of the Protein Society

Publications

Submitted and in preparation:

1. Chu, W., Ghahramani, Z. and Wild, D.L. **Bayesian Segmental Models with Multiple Sequence Alignment Profiles for Protein Secondary Structure and Contact Map Prediction**. Submitted to *IEEE/ACM Transactions on Computational Biology & Bioinformatics* (2005).
2. Chu, W., Ghahramani, Z., Krause, R. and Wild, D.L. **Identifying Protein Complexes in High Throughput Protein Interaction Screens Using an Infinite Latent Feature Model.** Submitted to *Pacific Symposium on Biocomputing* (2005).

Papers in refereed journals and proceedings:

3. Podtelezhnikov, A. and Wild, D.L. **Exhaustive Metropolis Monte Carlo sampling and analysis of polyalanine conformations adopted under the influence of hydrogen bonds.** *Proteins: Structure, Function and Bioinformatics*, 61:94-104 (2005).
4. Chu W., Ghahramani, Z., Falciani, F. and Wild, D.L. **Biomarker Discovery in Microarray Gene Expression Data with Gaussian Processes.** *Bioinformatics* , 21: 3385-3393 (2005).
5. Beal, M.J., Falciani, F., Ghahramani, Z., Rangel C. and Wild, D.L. **A Bayesian approach to reconstructing genetic regulatory networks with hidden factors.** *Bioinformatics*, 21: 349-356 (2005).
6. Saqi, M.A.S. and Wild, D.L. **Expectations from structural genomics revisited: an analysis of structural genomics targets.** *American Journal of Pharmacogenomics* (2005) (in press).
7. Rangel, C., Angus, J., Ghahramani, Z., Lioumi, M., Sotheran, E., A., Gaiba, A.,.Wild, D.L. and Falciani, F. **Modeling T-cell activation using gene expression profiling and state space models.** *Bioinformatics*, 20(9):1361-1372 (2004).
8. Dubey, A., Hwang, S., Rangel, C., Rasmussen, C.E., Ghahramani, Z. and Wild, D.L. **Clustering protein sequence and structure space with infinite Gaussian mixture models.** *Pacific Symposium on Biocomputing* 2004. Ed. R.B. Altman, A.K. Dunker, L. Hunter and T.E. Klein. World Scientific Publishing, Singapore, 399-410 (2004).
9. Bourne, P.E., Allerston, C.J.K., Krebs, W., Li, W., Shindyalov, I.N., Godzik, A., Friedberg, I., Liu, T., Wild, D., Hwang, S., Ghahramani, Z., Chen, L. and Westbrook, J. **The Status of Structural Genomics Defined Through the Analysis of Current Targets and Structures.** *Pacific Symposium on Biocomputing* (2004). ed. R.B. Altman, A.K. Dunker, L. Hunter and T.E. Klein. World Scientific Publishing, Singapore, 375-386 (2004).
10. Wild, D.L. and Saqi, M.A.S. **Structural Proteomics: Inferring function from protein structure.** *Current Proteomics* 1, 59-65 (2004).
11. Mainguy J.-P., MacDonnell G., Bund, S. and Wild, D.L. **KMD: an open source port of the ArrayExpress microarray database.** *Applied Bioinformatics*, 3(4), 257-60 (2004).
12. Chu W., Ghahramani Z. and Wild D.L. **A Graphical Model for Protein Secondary Structure Prediction.** Proceedings of the 21st International Conference on Machine Learning, Banff, Canada, 2004.
13. Chu W., Ghahramani Z. and Wild D. **Predicting Protein Secondary Structure using Sigmoid Belief Networks to Parameterize Segmental Semi-Markov Models.** Proceedings of the 12th European Symposium on Artificial Neural Networks (2004). D-side publications, Evere, Belgium.
14. Raval, A., Ghahramani, Z. and Wild, D. L., **A Bayesian network model for protein fold and remote homologue recognition.** *Bioinformatics* 18(6):788-801 (2002).
15. Rangel, C., Wild, D. L. Falciani, F., Ghahramani, Z., and Gaiba, A., **Modeling biological responses using gene expression profiling and linear dynamical systems.** *Proceedings of the 2nd International Conference on Systems Biology*. Madison, WI: OmniPress, pp 248-256 (2001).
16. Shneerson, V. L., Wild, D. L. & Saldin, D. K., **An Exponential Modeling Algorithm for Protein Structure Completion by X-Ray Crystallography.** *Acta. Cryst. A* **57**, 163-175 (2001).
17. Saqi, M.A., Wild, D. L. and Hartshorn, M.J. **Protein Analyst: a distributed object environment for protein sequence and structure analysis**. *Bioinformatics,* **15**:521-522 (1999).
18. Saldin, D. K. Shneerson, V. L. & Wild, D. L., **A Maximum Entropy Algorithm for Holographic Structure Completion in Macromolecular X-ray Crystallography.** *Journa l of Imaging Science and Technology*, 41:482-487 (1997).
19. Saldin, D. K., Shneerson, V. L., Vamvakas, J. A. & Wild, D. L., **Holographic Methods for X-ray and Electron Crystallography.** *Micron*, 28: 321-329 (1997).

20. Bhattacharya, P. & Wild, D. L., **A New Edge Detector for Grey Volumetric Data.** *Computers in Biology and Medicine*, 26:315- 328 (1996).
21. Wild, D. L., Tucker, P. A. & Choe, S., **A Visual Data Flow Environment for Macromolecular Crystallographic Computing.** *Journal of Molecular Graphics*, **13:** 291-298 (1995).
22. Timmins, P. A., Wild, D. & Witz, J., **The three-dimensional distribution of RNA and protein in the interior of tomato bushy stunt virus: a neutron low-resolution single-crystal diffraction study.** *Structure*, **2(12):** 1191-1201 (1994).
23. Wild, D. L. & Blanshard, J. M. V., **The Relationship of the Crystal Structure of Amylose Polymorphs to the Structure of the Starch Granule.** *Carbohydrate Polymers* **6:** 121-143 (1986).
24. Wilson, K. S., Stura, E. A., Wild, D. L., Todd, R. J., Stuart, D. I., Babu, Y. S., Jenkins, J. A., Standing, T. S., Johnson, L. N., Fourme, R., Kahn, R., Gadet, A., Bartels, K. S. & Bartunik, H. D., **Macromolecular crystallography with synchrotron radiation, II. Results**. *J. Appl. Cryst.* **16:** 28-41 (1983).
25. Skarnulis, A. J., Wild, D. L., Anstis, G. R., Humphreys, C. J. & Spence, J. C. H., **Multislice calculations using an Array processor.** *Inst. Phys. Conf. Ser. No. 61.* (Bristol: Institute of Physics) (1982).
26. Weber, I. T., Johnson, L. N., Wilson, K. S., Wild, D. L., Yeates, D. G. R. & Jenkins, J. A., **Crystallographic studies on the activity of glycogen phosphorylase b.** *Nature* **274:** 433-437 (1978).
27. Johnson, L. N., Weber, I. T., Wild, D. L., Wilson, K. S. & Yeates, D. G. R., **Crystallographic analysis at low resolution of metabolite binding sites on phosphorylase b.** *J. Mol.. Biol.* **118:** 579-591 (1978).
28. Johnson L.N., Wilson K.S., Weber I.T., Wild D.L., Jenkins J.A., Stura E.A., **Crystallographic studies on the structure and function of glycogen phosphorylase b.** Biochem Soc Trans 6(6):1108-11 (1978).
29. Johnson, L. N., Weber, I. T., Wild, D. L., Wilson, K. S. & Yeates, D. G. R., **The crystal structure of glycogen phosphorylase b.** *Proc. 11th FEBS Meeting, Copenhagen,* 1977, **Vol. 42:** 185-194. (Oxford: Pergammon) (1978).

Book Chapters

1. Rangel, C. Angus, J., Ghahramani, Z. and Wild, D.L. **Modeling genetic regulatory networks using gene expression profiling and state space models.** In Husmeier, D., Roberts, S. and Dybowski, R. (Eds.) , ***Applications of Probabilistic Modelling in Medical Informatics and Bioinformatics.*** Springer Verlag, (2004), pp. 269-293.

2. Saldin, D.K, Shneerson, V.L. and Wild,D.L. **A maximum entropy algorithm for holographic structure completion in macromolecular X-ray crystallography.** In J.-M. Fournier (Ed.), ***Holography: The first 50 years.*** Springer-Verlag (2004). In press.

Selected Conference Papers and Articles:

1. Beal, M.J., Rangel, C., Falciani, F., Ghahramani, Z. and Wild, D.L. **Classical and Bayesian approaches to reconstructing genetic regulatory networks.** *12th International Conference on Intelligent Systems for Molecular Biology*; Glasgow, Scotland, 2004 (poster).
2. Chu W., Ghahramani, Z. and Wild, D.L **Bayesian Segmental Semi-Markov Models for Protein Secondary Structure Prediction** *12th International Conference on Intelligent Systems for Molecular Biology*; Glasgow, Scotland, 2004 (poster).
3. Wild, D.L., Rasmussen, C.E., Ghahramani, Z., Cregg, J., de la Cruz, B.J., Kan C-C., and Scanlon, K. **A Bayesian approach to modelling uncertainty in gene expression clusters**. *3rd International Conference on Systems Biology,* Stockholm, Sweden (2002) (extended poster abstract). http://www.ki.se/icsb2002/pdf/ICSB_227.pdf
4. Wild, D.L., Raval A. and Ghahramani Z. **A Bayesian network model for protein fold and remote homologue recognition**. *Structural Genomics: From Gene to Structure to Function, Molecular Graphics and Modelling Society Meeting,Cambridge* (2000) (poster).
5. Goryanin, I., Saqi, M.A.S. and Wild, D.L. **An approach to the analysis of bacterial genome data which combines structural and pathway information.** *Data Mining for Bioinformatics - Towards In Silico Biology, European Bioinformatics Institute, Cambridge* (1999) (poster).
6. Dov J Stekel, D.J., Wild, D., Falciani, F. and Zanders, E.D. **Mathematical analysis of inflammatory gene expression pathways induced by Il-1 beta and Tnf-alpha in human fibroblast-like synoviocytes**. *Data*

Mining for Bioinformatics - Towards In Silico Biology, European Bioinformatics Institute, Cambridge (1999) (poster).

7. Wild, D., and Ghahramani, Z. **A Bayesian network approach to protein fold recognition**. *6th International Conference on Intelligent Systems for Molecular Biology*; Montreal, Canada, 1998 (poster).
8. Saqi, M., Wild, D. and Hartshorn, M. **Protein Analyst: a distributed object environment for protein sequence and structure analysis.** *6th International Conference on Intelligent Systems for Molecular Biology*; Montreal, Canada, 1998 (poster).
9. Wild, D.L. & Choe, S., **An AVS/Express Interface to CCP4**. *in: Crystallographic Computing 7*, (eds: Bourne, P. and Watenpaugh, K.). IUCr. (1996) http://www.iucr.org/iucrtop/comm/ccom/School96/iucr.html
10. Bhattacharya, P. & Wild, D. **An Edge Detector for Gray Volumetric Data and Its Applications to Bioimaging**, *Proc. of the Conf. on Applications of Digital Image Processing XVIII*, San Diego, CA, pp. 506-510. (Vol. 2564, SPIE Proc. Series).
11. Wild, D. L., Chen, X. & Saldin, D. K., **Holographic Reconstruction for Macromolecular Structure Completion in X-Ray Crystallography by Iterative Applications of Linear Programming.** *in: Signal Recovery and Synthesis V, Vol. 11, 1995*, OSA Technical Digest Series, (Optical Society of America, Washington DC, 1995), pp 64-67.
12. Wild, D. L., **Developing an X-windows based user interface for protein crystallographic software.** *Joint CCP4 and ESP-EACBM Newsletter on Protein Crystallography*; Daresbury Laboratory, **25:** 78 (1991).
13. Wild, D. L., **An X-windows based user interface for protein crystallographic software.** *in: Crystallographic Computing 5*, (eds: Moras, D., Podjarny, A. D. & Thierry, J. C.) O.U.P, p. 471 (Abstract) (1991).

Recent Invited Presentations:

- Conway Institute, Dublin, Seminar, May 2005. **Graphical Models and Bayesian Methods in Systems Biology and Clinical Bioinformatics.**
- Mayo Clinic, Rochester, Seminar, April 2005. **Graphical Models and Bayesian Methods in Systems Biology and Clinical Bioinformatics.**
- Indiana University, Informatics Colloqium, February 2005. **Graphical Models and Bayesian Methods in Systems Biology and Bioinformatics.**
- The Jackson Laboratory, Maine, Seminar February 2005. **Modeling genetic regulatory networks in T-cell activation.**
- Hutchison/MRC Research Centre, Cambridge, UK. Seminar, July 2004. **Gaussian processes for microarray cancer gene expression classification.**
- University of Birmingham, School of Biosciences. Seminar, July 2004. **Reconstructing genetic regulatory networks with hidden factors.**
- Virginia Bioinformatics Institute, Blacksburg. Seminar, June 2004. **Reconstructing genetic regulatory networks with hidden factors.**
- University of Hawaii, Cell and Molecular Biology Graduate Program Seminar, June 2004. **Reconstructing genetic regulatory networks with hidden factors.**
- 14th Annual Meeting on Mathematical and Statistical Aspects of Molecular Biology, Cambridge, March 2004. **Classical and Bayesian approaches to reconstructing genetic regulatory networks with hidden factors.**
- National e-Science Centre, Edinburgh, January 2004. **Graphical models and Bayesian methods in bioinformatics.**
- Pacific Symposium on Biocomputing 2004, Hawaii, January 2004. **Clustering protein sequence and structure space with infinite Gaussian mixture models.**
- University of Southern California, Computational Biology Colloquim, November 2003. **Graphical models and Bayesian methods in bioinformatics.**
- University College London, Centre for Computational Science, October 2003. **Bayesian approaches to protein secondary structure prediction and family clustering.**
- GlaxoSmithKline Medicines Research Centre, Stevenage, July 2003. **Modelling T-cell activation using gene expression profiling and state space models.**
- CCP11 Meeting, *Towards the Functional Analysis of Microarrays*, Manchester, March 2002 **Modelling biological responses using gene expression profiling and linear dynamical systems.**

- Institute for Adaptive and Neural Computation, University of Edinburgh, June 2002. **Modelling biological responses using gene expression profiling and linear dynamical systems.**
- MRC Laboratory of Molecular Biology, Cambridge, January 2002. **Modelling biological responses using gene expression profiling and linear dynamical systems.**
- 6th Biennial Conference of the European Society for Engineering and Medicine, Belfast, May 2001. **A Bayesian network model for protein fold and remote homologue recognition.**
- 2nd International Conference on Systems Biology, Caltech, December 2001 **Modeling biological responses using gene expression profiling and linear dynamical systems.**

Ph. D. theses supervised:

Modeling Biological Responses using Gene Expression Profiling and Linear Dynamical Statistical Models. Claudia Rangel Escareno. Claremont Graduate University, School of Mathematical Sciences, 2003.