May 7, 2008

CURRICULUM VITAE

David Johnson States, M.D., Ph.D.

Position Professor of Human Genetics
Director of Bioinformatics Training Program
University of Michigan
Ann Arbor, MI 48109

Address

Work: 100 Washtenaw Rd.
Ann Arbor, MI 48109

Home: 2254 Belmont Rd.
Ann Arbor, MI 48104

Telephone: (734) 615-5510

Fax: (734) 615-6553

Email: dstates@umich.edu
David_J_States@post.harvard.edu

URL: http://www.stateslab.org

Date of Birth: July 12, 1953 **Place of Birth:** Boston, MA

Citizenship: U.S.A.

Marital Status: Married - Angel W. Lee, M.D., Ph.D.

Education

1967 - 1971	Graduate with honors, Brighton High School, Rochester, NY
1971 - 1975	B.A. magna cum laude in Biochemistry Harvard University, Cambridge, MA
1975 - 1983	M.D. and Ph.D. in Biophysics Harvard University and Harvard Medical School, Boston, MA

Research and Professional Experience

1973-1975	Undergraduate honors thesis, Harvard University, Cambridge, MA Advisor: Prof. J. Woodland Hastings Thesis: Selection of circadian rhythm mutants in Chlamydomonas rheinhardtii.
1975	Research Assistant to Prof. Walter Gilbert Harvard University, Cambridge, MA Research: Sequencing the L-1 deletion of the *lac* operon.
1976-1977	Research Assistant to Prof. Lawrence Kedes, Stanford University, Palo Alto, CA Research: Heterogeneity in sea urchin histone genes.
1977-1983	Doctoral thesis, Harvard University, Cambridge MA Advisors: Profs. Martin Karplus and Christopher Dobson Thesis: Magnetic Resonance and Theoretical Studies of the Refolding of Bovine Pancreatic Trypsin Inhibitor.

1983-1984	Staff Scientist, Francis Bitter National Magnet Laboratory Massachusetts Institute of Technology, Cambridge, MA Supervisor: Leo J. Neuringer, Ph.D., Director of Biophysics Research: Magnetic resonance imaging theory and instrumentation.
1984-1985	Consultant to IBM Instruments Inc., Danbury, CT Supervisor: Colin Harrison, Ph.D.
1984-1986	Internship and Residency in internal medicine University of California at San Diego Medical Center, San Diego CA Chief of Medicine: Helen Ranney, M.D.
1986-1989	Medical Staff Fellow, pulmonary medicine National Heart Lung and Blood Institute, National Institutes of Health Clinical director: Harry R. Keiser, M.D. Laboratory director: Ronald G. Crystal, M.D. Research: Molecular genetics and biophysics of proteases in the lung.
1989-1992	Senior Staff Fellow, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health Research: Biophysical approaches to the analysis of molecular sequence data.
1992-1995	Director, Biomedical Engineering Program, Washington University School of Engineering & Applied Science, St. Louis, MO
1992-2000	Associate Professor of Biomedical Computing, Joint appointments: Biochemistry and Molecular Biophysics, Biomedical Engineering, Computer Science Washington University Medical School, St. Louis, MO Research: Computational genome analysis.
1992-2000	Director, Institute for Biomedical Computing, Washington University Medical School, St. Louis, MO
1997- 1999	Director, Center for Genetics in Medicine, Washington University Medical School, St. Louis, MO
2000-2001	Associate Professor of Genetics, Joint appointments: Biochemistry and Molecular Biophysics, Biomedical Engineering, Computer Science Washington University Medical School, St. Louis MO Research: Computational genome analysis.
2001-present	Professor of Human Genetics, Director of Bioinformatics Training Program University of Michigan, Ann Arbor, MI Research: Computational genomics and proteomics
2006-present	Senior Scientist, National Center for Integrative Biomedical Informatics University of Michigan, Ann Arbor, MI

Medical License

State of Maryland D34107 (1986-1992)

Board Certification

American Board of Internal Medicine, 1987

Awards and honors

Harvard College Scholar, graduate of Harvard College with *magnu cum laude* honors
Dreyfus Foundation Summer Undergraduate Research Fellowship, Harvard College
NIH Medical Scientist Training Program Fellowship for M.D., Ph.D., Harvard Medical School
Howard Hughes Medical Research Fellowship with Prof. Laurence Kedes, Stanford University
Member of Strathmore's Who's Who
Fellow of the American College of Medical Informatics (elected 2001)

Professional Activities

Scientific Boards

Scientific Advisory Board (chair), Protein Information Resource, Georgetown University, Washington, DC
Scientific Advisory Board, Cytoscape, Univ. of California at San Diego., CA.
Scientific Advisory Board, Gene Ontology, Jackson Laboratory, Bar Harbor, ME.
Hartwell Center for Biotechnology, St Jude Children's Research Hospital, Memphis, TN
Scientific Advisory Board, Blueprint Canada, Toronto, CA
National Biomedical Research Foundation, Georgetown University, Washington, DC
DzGenes Corp., St. Louis, MO
MGI Corp., Los Angeles, CA
Gene Expression Analysis Core, Siteman Cancer Center, Washington University, St. Louis, MO

Study Sections

1991, 1992	National Institutes of Health, National Center for Human Genome Research ad hoc study section for informatics
1990, 1991	Department of Energy, Office of Health and Environmental Research Genome study section
1997	National Institutes of Health, National Center for Research Resources ad hoc study section for computational biology
1998	National Institutes of Health, National Library of Medicine ad hoc study section for Genome Informatics
1998	National Science Foundation ad hoc study section for Plant Genomics
2000	National Heart Lung and Blood Institute Special panel on genomics
2000-2002	Sloan Foundation Advisory Committee for Postdoctoral Fellowships in Computational Molecular Biology
2001-2004	National Institutes of Health, National Heart Lung and Blood Institute External Scientific Panel Programs For Genomic Applications
2001-2004	National Institutes of Health, National Institute for General Medical Sciences Graduate training study section
2007-2011	National Institutes of Health, National Library of Medicine Biomedical Library and Informatics study section

Grant review

1990-present	Division of Computational Biology, National Science Foundation
1990-present	Office of Health and Environmental Research (Genome), Department of Energy

Editorial Boards

2007-present	PLoS Computational Biology (ad hoc Associate Editor)
2004-present	Cancer Proteomics
2001-present	Applied Bioinformatics
1996-present	Journal of Computational Biology
2001-2002	Associate Editor for North America, Bioinformatics
1992-1996	GENE-COMBIS

Journal review panels

Science, Nature, Nature Biotechnology, Nature Genetics, Proceedings of the National Academy of Science, Nucleic Acids Research, Genome Research, Genomics, Bioinformatics, Journal of Proteome Research, Proteomics, Journal of Molecular Biology, Journal of Theoretic Biology, IBM Systems Journal.

Conferences and workshops

1992	Co-chair – Computational Biology Mini-track 27th Hawaiian International Conference on Systems Science, Maui, HI
1993	Program committee, Intelligent Systems for Molecular Biology, National Library of Medicine, Bethesda, MD.
1994	Co-chair for Informatics, Genome Mapping and Sequencing Meeting Cold Spring Harbor Laboratory, NY
1994	Organizer, Software sharing workshop Cold Spring Harbor Laboratory, NY
1994	Program committee, Intelligent Systems for Molecular Biology Stanford University, Palo Alto, CA
1995	Organizer, Protein classification workshop Washington University, St. Louis, Missouri
1995	Program committee, Intelligent Systems for Molecular Biology Cambridge University, Cambridge, U.K.
1996	Chair, Organizing Committee, Intelligent Systems for Molecular Biology Washington University, St. Louis, Missouri
1996	Co-organizer Java/Corba workshop on databases in genome analysis Washington University, St. Louis, Missouri
1999	Organizing committee, BioPerl Workshop Heidelberg, Germany

1999	Invited participant, Workshop on Chromosomal Aberrations National Cancer Institute, Bethesda, MD
1999-2001	Organizing committee, Computational Genomics Baltimore, MD
2005	Chair, Organizing Committee, Intelligent Systems for Molecular Biology Detroit, MI

Professional societies

International Society for Computational Biology

1997-2003	Board of Directors (founding member)
1997-2000	Treasurer
2004-	Conference Committee (member)
2004-	Public Affairs Committee (member)
2004-2005	Chair of the Organizing Committee for ISMB 2005
2008-	Chair of the Public Affairs Committee

European Conference on Computational Biology

2007-2008	Computational Systems Biology Scientific Program Committee (member)

Life Science Society

2007-2008	Computational Systems Biology Scientific Program Committee (member)

American Medical Informatics Association

2001-	Fellow of the American College of Medical Informatics
2008-	Conference Committee (member)

American Society for Hematology (member)

Protein Society (member)

American Association for the Advancement of Science (member)

Washington University

1992-2000	Executive Committee, School of Engineering and Applied Science
1997-2001	Executive Committee, Institute for Biological and Medical Engineering
1998-1999	Program Committee for Molecular Genetics, Division of Biology and Biomedical Science
2000-present	Program Committee for Computational Biology, Division of Biology and Biomedical Science

University of Michigan

2001-2005	Basic Science Chairs Committee, University of Michigan Medical School
2002-2005	Endowment for the Basic Sciences Operating Committee
2001-2005	Program in Biological Sciences Operating Committee
2001-2005	Program in Biological Sciences Admissions Committee

2004-2005	T-FORE Task Force for the Research Enterprise
2005-2006	Brehm Diabetes Center Planning Group
2001-2005	Program in Bioinformatics Executive Committee (Chair)
2001-2006	Bioinformatics Graduate Affairs Committee
2005-2006	Information Technology Faculty Advisory Committee (ITFAC)
2004-2008	General Clinical Research Center Advisory Committee
2005-present	Center for Computational Medicine and Biology Executive Committee
2005-present	National Center for Integrative Biomedical Informatics Executive Committee

Training Activities

Training Grants

Principal Investigator
Genome Analysis Institutional Training Grant (5T32HG0004504).
National Institute for Genome Science/National Institutes of Health,
1997 through 2002, (newly established program).

Principal Investigator
Bioinformatics Training Grant. (T32-GM-070449-01-a1)
National Institute for General Medical Science/National Institutes of Health,
2005 through 2010, (newly established program).

Courses created

Bio 5495	Computational Molecular Biology – a graduate level course that served as the core of the Washington University Computational biology training program
BME 537a	Intensive Short Course in Computational Biology – an intensive course for professional scientists in molecular biotechnology and pharmaceutical research.
Bio 5496	Journal Club in Computational Biology – a weekly graduate literature review course
Bioinfo 526	Fundamentals of Bioinformatics – a graduate level course that served as the core of the University of Michigan Bioinformatics training program
Bioinfo 602	Journal Club in Bioinformatics
Pharm 618	Cancer Systems Biology (with Dr. Angel W. Lee)
Bioinfo 575	Programming Laboratory in Bioinformatics

Courses taught

1996-1998	Course master for Bio5495 Computational Molecular Biology
1996-2001	Course master for Bio5496 Computational Biology Journal Club
1997-2001	Course master for BME537a Intensive Short Course on Computational Molecular Biology

1999-2001	Co-instructor CS546T Advanced Algorithms for Computational Biology
2001-2006	Course master for Bioinfo 602 Journal Club in Bioinformatics
2002-2004	Course master for Bioinfo 526 Fundamentals of Bioinformatics
2005-present	Lecturer in Bioinfo 526/527
2002-2004	Lecturer in Human Genetics 804 Methods in Molecular Genetics
2006-present	Co-course master, Human Genetics 802/803 Student Seminar
2005-2006	Pharm/Bioinfo 618 Cancer Systems Biology (co-taught with Dr. Angel Lee)
2006-present	Course master for Bionfo 575/800.3 Programming Laboratory in Bioinformatics

Internships and mentoring

2007-	Google Summer of Code Maital Ashkenazi, Hebrew University, Jerusalem, Israel Enhanced Search Plugin for Cytoscape
2008-	Google Summer of Code Patrick McSweeney, Syracuse University Random Network Generators
2008-	Google Summer of Code Gang Su, University of Michigan Community Structure Analysis in Biological Network

Students and Postdoctoral Trainees

Postdoctoral

D'vorah Graeser - postdoc completed '94
Nengbing Tao - postdoc completed '95
William Reisdorf - postdoc completed '97
Geetha Vasanthakumar - postdoc completed ‘98
Lisa Gu - postdoc completed ’98
Marcin Adamski – postdoc 2002-2005
Anura Hewagama – postdoc 2002-2007
Tom Blackwell- 1996-present

Doctoral

David Politte - doctoral complete, 1999
Zhengyan Kan – doctoral complete, 2002
Rongxiang Liu – doctoral complete, 2002
Eric Rouchka – doctoral complete, 2002
Richard McEachin – doctoral complete, 2004
Orkun Soyer – doctoral complete, 2004 (co-mentor with Richard Goldstein)

Bin Qian – doctoral complete, 2004 (co-mentor with Richard Goldstein)
Ji Chen – doctoral complete, 2006
Damian Fermin – doctoral compete, 2007
Yili Chen – doctoral compete, 2008
Carlos Santos – doctoral compete, 2008 (co-mentor with Brian Athey)
Arvind Rao – doctoral compete, 2008 (co-mentor with Doug Engel)
Xing Li – doctoral candidate (co-mentor with Deborah Gomucio)
Sirarat Sarntivijai- doctoral candidate (co-mentor with Brian Athey)
Gang Su– doctoral candidate
Yu-Hsuan Lin – doctoral candidate (co-mentor with Doug Engel)
Junguk Hur – doctoral candidate (co-mentor with Eva Feldman)

Wie Liu - rotation student
German Leparc - rotation student
Weidong Tian - rotation student
Yongmei Ji - rotation student
Kai Tan - rotation student
Long Lu - rotation student
Todd Lowe - rotation student (member of thesis committee)
Tom Nishino - rotation student
Nilesh Ron - rotation student
Chad Creighton – rotation student
Geogi Kostov – rotation student
Manjusha Pande – rotation student
Yongsheng Huang – rotation student
Yan Zhang – rotation student
Chunchao Zhang – rotation student

Mark Benson – committee member
Adrian Chapman – committee member
Jeffery Hou – committee member
Jessica Lehoczky – committee member
Colleen McCabe – committee member
Arnab Nandi – committee member
Viktoria Strumba – committee member
Shankar Subramaniam – committee member
Yuanyuan Tian – committee member
Cong Yu – committee member

Masters

Xiaobing Shi - masters complete ’99
Keith Doolittle - masters completed ‘98
Nomi Harris - masters completed '92
Reece Hart - masters completed '94 (chair of doctoral thesis committee)
Tina Seawell - masters completed '96
Patrica Widder - masters completed ‘98
Sirarat Sarntivijai- masters complete 2006

Interns

Regina Patel – biological and engineering science
Katrin Wormit – information science

Daniella Eggle – bioinformatics
Maital Ashkenazi – bioinformatics (sponsored by Google Summer of Code 2007)

Undergraduate

Peter Chiu – computer science
Ron Knickerbocker – computer science
Surya Rednam – biomedical engineering
John McCrow – mechanical engineering
Kelly Carter – biomedical engineering
Trevor Harmon – computer engineering
Teddi Tejda - biomedical engineering
Carlos Santos - computer science/biology
Casey Overby – bioinformatics (individual major)
Jeremy Philips – computer science
Kristin Veresh – Industrial and Operations Engineering
Heather Grifka – Biology (honors thesis)
Suellen Yin – Biology/Math
Rachel Harrison – Biology/Math
Jennifer MacDonald – Biochemistry/Math

Grant support

Principal investigator

1994-1997	Department of Energy (DE-FG02-94ER61910) Analysis and Annotation of Nucleic Acid Sequence $750,000
1995-1999	National Institutes of Health (HG-01391) Information Systems for Very High Throughput Sequencing $576,144 $94,400 (supplement)
1997-2002	National Institute of Health (HG00045) Genome Analysis Institutional Training Grant $1,01,827
1997-2002	National Science Foundation NCSA PACI Molecular Biology application team Efficient tools for sequence analysis $290,000
1997-2002	National Science Foundation SDSC NPACI Molecular Sciences thrust Large-scale sequence classification $125,000
1998-2001	Department of Energy (DE-FG02-94ER61910) Analysis and Annotation of Nucleic Acid Sequence $700,000
2000-2002	Merck Foundation for Genome Research (grant #225) Structural Modeling of Genomic Regulatory Complexes $293,000

2002-2004	Alfred P. Sloan Foundation (2002-5-59-BCMB) Professional Masters Degree in Bioinformatics $64,990
2000-2005	Howard Hughes Medical Institute (76200-561301) 1999 Biomedical Research Support Program for Medical Schools – Bioinformatics $3,875,000
2002-2005	National Institutes of Health (R01 LM005770) Computational Approaches to Protein Sequence Analysis $932,336
2002-2007	National Institute of Health (LM008106) Representing and Acquiring Knowledge of Genome Regulation $1,757,690
2005-2010	National Institutes of Health (T32-GM-070449-01-a1) Bioinformatics Training Grant $1,225,579
2007-2011	National Institute of Health (LM008106 - pending) Representing and Acquiring Knowledge of Genome Regulation $1,757,690

Co-investigator

1993-1997	National Institute of Health (HG00956, R. Waterston PI) Large-scale Genome Sequencing Role: co-investigator $33,500 (States lab component)
1994-1997	National Institutes of Health (HG00201, D. Schlessinger, PI) Vertically Integrated Genome Mapping and Sequencing Role: director of the Informatics Core $1,275,650 (informatics core)
1994-2000	National Science Foundation (NCR-9318178, J. Cox, PI) Distributed Networking of Gigabit Networks
1994-2000	**Principal investigator:** 4-D Computational Optical Sectioning Microscopy $2,118,180 (COSM component)
1997-2000	National Institute of Health (HG001720, P Kwok PI) Diallelic marker discovery and testing Role: co-investigator for informatics $298,980 (informatics component)
2000-2003	National Institute of Health (HG001720, P Kwok PI) Diallelic marker discovery and testing Role: co-investigator for informatics $266,112 (informatics component)
2000-2006	National Institute of Health (M01 RR 000042-44, Robert Kelch, PI) General Clinical Research Center (member of the GCRC Advisory Committee) $30,592,412

2001-2005	National Institute of Health (pending, T Simon PI) Endometerial Cancer Program Project Role: co-investigator $136,000 (States lab component)
2002-2005	Michigan Economic Development Council (Gilbert S. Omenn, PI) Proteomics Alliance for Cancer $3,126,588
2003-2008	National Institute of Health (P41-RR-018627, Philip Andrews, PI) National Pathway Mapping Center (director of the bioinformatics core) $12,073,929
2004-2007	National Institutes of Health (BAA RM-04-23 Daniel Clauw, PI) Michigan Clinical Research Collaboratory (MCRC) Project
2005-2007	Michigan Economic Development Council (Gilbert S. Omenn, PI) Proteomics Alliance for Cancer Research $2,363,605
2005-2007	National Cancer Institute (Gilbert Omenn, PI) Mouse Models for Cancer, Eastern Consortium $3,629,357
2005-2008	National Institutes of Health (P20 HG003890-01, Kerby Shedden, PI) MACE – Michigan Alliance for Cheminformatics Exploration $610,000
2005-2010	National Institutes of Health (U54 DA21519-01A1, Brian Athey PI) National Center of Integrative Biomedical Informatics (senior scientist and director of the bioinformatics and education cores) $18,698,966
2006-2008	American Diabetes Association (Jessica Schwartz, PI) Dissecting a Genetic Program for GH-Induced Insulin Resistance $200,000
2007-2010	Yale University (subcontract on ENCODE proposal) An Integrative Approach Towards Complete Definition of the Transcriptome $227,579
2008-2012	National Institute of Health (P41-RR-018627, Philip Andrews, PI - pending) National Pathway Mapping Center (director of the bioinformatics core) $12,778,108

Collaborators

Pankaj Agarwal, SmithKline Beecham
Russ Altman, Stanford University/San Diego Supercomputer Center
Sarah Elgin, Washington University
Philip Green, University of Washington
Steven Gullans, Harvard University
Lawrence Hunter, University of Colorado
Mark Johnston, Washington University

Pui-Yan Kwok, Washington University
Angel W. Lee, Washington University
James Lindelien, Time Logics Inc.
Michael Lovett, Washington University
David Schlessinger, Washington University
Robert Schreiber, Washington University
Martin McIntosh, University of Washington
Robert Murphy, Carnegie Mellon University
Mark Musen, Stanford University
Judith Blake, Jackson Laboratory
Catherine Wu, Georgetown University
Jessica Schwartz, University of Michigan
Zhaohui Qin, University of Michigan
Cun-yun Wang, University of Michigan
J. Douglas Engel, University of Michigan
Alfred Hero, University of Michigan
Hosagrahar V. Jagadish, University of Michigan
Jignesh Patel, University of Michigan
Dragomir Radev, University of Michigan
Deborah Gumucio, University of Michigan
Kirby Shedden, University of Michigan
Gustavio Rosania, University of Michigan
Brian Athey, University of Michigan
Gilbert S. Omenn, University of Michigan
Samir Hanash, University of Washigton
Raju Kucherlapati, Harvard Medical School
David Saracino, Harvard Partners
Melvin McInnis, University of Michigan
Haimeng Chen, University of Michigan

Invited lectures, seminars and national meeting presentations

1983 Francis Bitter National Magnet Laboratory, Massachusetts Insitute of Technology
NMR and theoretical studies of pancreatic trypsin inhibitor folding

1983 University of Massachusetts at Boston, Department of Biochemistry
Novel approaches to high resolution multi-dimensional NMR

1991 Cold Spring Harbor Genome Mapping and Sequencing Meeting
Analysis of error prone genomic sequence data

1992 Hawaiian International Conference on Systems Science, Maui, HI
Bayesian classification of protein structural elements

1992 Hawaiian International Conference on Systems Science, Maui, HI
Tutorial on molecular biology and genomics

1992 National Science Foundation Workshop on Scientific Infrastructure
Future of biocomputing

1992 Baylor College of Medicine, Human Genome Center
Computational Genome Analysis

1992 Yale University, Department of Human Genetics
Computational Genome Analysis

1992 Washington University, Biomedical Computer Laboratory
Computational Genome Analysis

1992 American Society for Crystallographic Research
Bayesian classification of protein sequence and structure

1993 Biophysical Society
Protein sequence megaclassification

1994 Washington University, Biophysical Evenings
Protein sequence and structure classification

1994 Department of Energy Human Genome Grantees and Contractors Meeting
Analysis and Annotation of Nucleic Acid Sequence

1995 Genome Sequence Analysis Conference, Hilton Head, S.C
Basecalling accuracy and the reliability of genome sequence assembly

1995 Johns Hopkins University, Genome Database Seminar Series
Automating the Analysis and Annotation of Nucleic Acid Sequence

1995 Washington University, Department of Physics
Physics of protein folding

1996 Intelligent Systems for Molecular Biology Conference
Optimal encoding strategies for sequence similarity search

1996 Department of Energy Human Genome Grantees and Contractors Meeting
Data accuracy and the analysis of genomic sequence data

1996 Monsanto Corporation
Computational issues in genome sequencing and annotation

1997 University of Chicago, Department of Biochemistry
Computational Genome Analysis

1997 Department of Energy Human Genome Grantees and Contractors Meeting
Automated genome annotation

1997 Department of Energy Human Genome Grantees and Contractors Meeting
Panel discussion on genome sequence assembly accuracy

1998 San Diego Supercomputer Center, Molecular Sciences Workshop
Protein sequence megaclassification

1998 Pfizer Corporation
Automated Genome Annotation: Keeping up with the Information Explosion

1996 Monsanto Corporation
Automated Genome Annotation: Keeping up with the Information Explosion

1998 University of Pennsylvania, Department of Genetics
Structural models for DNA binding protein specificity

1999 Novartis Corporation
Computational genome annotation

1999 Research in Computational Biology, New York, NY.
Estimation of allele frequencies from color-multiplexed electropherograms

1999 Intelligent Systems for Molecular Biology
Identity by descent segmentation and single nucleotide polymorphism distributions in the human genome

1999 Intelligent Systems for Molecular Biology
Tutorial on relational database abstraction and XML

1999 Cold Spring Harbor Workshop on human genome polymorphism databases
Databases for single nucleotide polymorphisms in the human genome

1999 Computational Genomics Conference
Single nucleotide polymorphism clustering and human population genetics

2000 Mallinkrodt Chemical Corporate Research Day
Post-genome science

2000 Cold Spring Harbor Genome Mapping and Sequencing Meeting
Transcript reconstruction using genomically aligned EST sequences

2000 Cold Spring Harbor Genome Mapping and Sequencing Meeting
Identifying transcription factor binding site clusters in the human genome

2000 University of Pennsylvania, Department of Genetics
Identifying transcription factor binding site clusters in the human genome

2000 Washington University, Department of Genetics
Identifying transcription factor binding site clusters in the human genome

2000 Vanderbilt University School of Medicine, Division of Human Genetics
Identifying transcription factor binding site clusters in the human genome

2000 University of Maryland Supercomputer Center, Distinguished Lecture Series
Computational dissection of genomic regulatory elements

2001 Pfizer Bioinformatics Program Seminar
Computational Systems Biology

2002 University of Michigan Biophysics Seminar
Statistical Mechanics of Chromatin and Gene Regulation

2002 University of Michigan Pharmacology and Toxicology Retreat
Is scientific data open? Data sharing in biomedical research

2002 University of Michigan Physics Seminar
Statistical Mechanics of Chromatin and Gene Regulation

2003 University of Michigan Biomedical Engineering Seminar
Computational Systems Biology

2004 W3C Life Sciences Working Group
Life Science Ontology Issues

2004 University of Michigan Toxicology Training Program
Bioinformatics for Toxicology

2004 National Cancer Institute Mouse Models for Cancer Consortium Meeting
Proteomics data analysis and integration

2004 University of Michigan Obstetrics and Gynecology Grand Rounds
Computational Systems Biology

2005 University of Georgia, Computational Systems Biology Symposium
Data integration in proteomics and systems biology

2006 National Center for Integrative Biomedical Informatics
First Annual Research Workshop
Data integration and natural language processing in systems biology

2006 National Cancer Institute Mouse Models for Cancer Consortium Meeting
Proteomics Data Analysis and the Assessment of Significance

2006 National Institute of Health
National Centers for Biomedical Computing All Hands Meeting, Bethesda, MD
Data integration and natural language processing in systems biology

2007 US HUPO Annual Meeting, Seattle, WA
Identifying novel translation products in human blood by mass spectrometry

2007 National Center for Integrative Biomedical Informatics
Second Annual Research Workshop
Data integration and natural language processing in systems biology

2007 NCI Mouse Models for Cancer/Clinical Proteomic Technology Assessment Program
Identifying novel translation products in mouse models for cancer by mass spectrometry

2007 Carnegie Mellon University, Computational Biology Program
Integrating Genomics and Proteomics: Novel Translation Products Identified by Mass Spectrometry

2007 Ohio Collaborative Conference on Bioinformatics (OCCBIO) Keynote address
Integrating Genomics and Proteomics: Novel Translation Products Identified by Mass Spectrometry

2007 University of Pittsburgh Medical Center
Integrating Genomics and Proteomics: Novel Translation Products Identified by Mass Spectrometry

2008 AMIA Translational Bioinformatics Summit
Panel organizer: High Dimensionality Data in Translational Bioinformatics

2008 Arizona State University/University of Arizona School of Medicine
Integrating Genomics and Proteomics: Novel Translation Products Identified by Mass Spectrometry

2008 University of Chicago Medical School
Integrating Genomics and Proteomics: Novel Translation Products Identified by Mass Spectrometry

2008 National Center for Integrative Biomedical Informatics
Third Annual Research Workshop
Data integration and natural language processing in systems biology

Publications

Peer reviewed articles (including peer reviewed conference publications)

1. **States, D.J.**, Dobson, C.M., Karplus, M. and Creighton, T.E. (1980) A conformational isomer of bovine pancreatic trypsin inhibitor protein produced by refolding. *Nature*, **286**, 630-632.
2. **States, D.J.**, Haberkorn, R.A. and Ruben, D.J. (1982) A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants. *Journal of Magnetic Resonance*, **48**, 286-292.
3. Brooks, B.R., Bruccoleri, R., E., D., O.B., **States, D.J.**, Swaminathan, S. and Karplus, M. (1983) CHARMM: A Program for Macromolecular Empirical Energy Modelling. *Journal of Computational Chemistry*, **4**, 187-230.
4. **States, D.J.** (1984) Number Crunching on IBM's new S9000, Application of a Micro-Computer System to Scientific Data Processing. *Byte*, **9**.
5. **States, D.J.**, Dobson, C.M. and Karplus, M. (1984) A new two-disulphide intermediate in the refolding of reduced bovine pancreatic trypsin inhibitor. *J Mol Biol*, **174**, 411-418.
6. Weiss, M.A., Eliason, J.L. and **States, D.J.** (1984) Dynamic filtering by two-dimensional 1H NMR with application to phage lambda repressor. *Proc Natl Acad Sci U S A*, **81**, 6019-6023.
7. Delepierre, M., Dobson, C.M., Karplus, M., Poulsen, F.M., **States, D.J.** and Wedin, R.E. (1987) Electrostatic effects and hydrogen exchange behaviour in proteins. The pH dependence of exchange rates in lysozyme. *J Mol Biol*, **197**, 111-130.
8. Simons, M. and **States, D.J.** (1987) Coronary thrombosis in non-Q-wave myocardial infarction. *N Engl J Med*, **316**, 106-107.
9. **States, D.J.**, Creighton, T.E., Dobson, C.M. and Karplus, M. (1987) Conformations of intermediates in the folding of the pancreatic trypsin inhibitor. *J Mol Biol*, **195**, 731-739.
10. **States, D.J.** and Karplus, M. (1987) The Calculation of Electrostatic Effects in Proteins. *Journal of Molecular Biology*, **197**.
11. Takahashi, H., Nukiwa, T., Yoshimura, K., Quick, C.D., **States, D.J.**, Holmes, M.D., Whang-Peng, J., Knutsen, T. and Crystal, R.G. (1988) Structure of the human neutrophil elastase gene. *J Biol Chem*, **263**, 14739-14747.
12. Crystal, R.G., Brantly, M.L., Hubbard, R.C., Curiel, D.T., **States, D.J.** and Holmes, M.D. (1989) The alpha 1-antitrypsin gene and its mutations. Clinical consequences and strategies for therapy. *Chest*, **95**, 196-208.
13. Hunter, L. and **States, D.J.** (1991), *HICSS-24*. IEEE Computer Society Press, Los Alamitos, CA, , Maui, HI, Vol. 24, pp. 595-604.
14. Hunter, L. and **States, D.J.** (1991), *Proceedings of the IEEE Conference on Applications of Artificial Intelligence*, pp. 2-17.
15. **States, D.J.** and Botstein, D. (1991) Molecular sequence accuracy and the analysis of protein coding regions. *Proc Natl Acad Sci U S A*, **88**, 5518-5522.

16. **States, D.J.**, Gish, W. and Altschul, S.F. (1991) Improved Sensitivity of Nucleic Acid Database Searches Using Application-Specific Scoring Matrices. *Methods*, **3**, 66-70.

17. Harris, N., Hunter, L. and **States, D.J.** (1992), *Proceedings of the Tenth National Conference on Artificial Intelligence*. AAAI Press/MIT Press, Menlo Park, CA, , pp. 837-842.

18. Hunter, L., Harris, N. and **States, D.J.** (1992) In Sleeman, D. and Edwards, P. (eds.), *Proceedings of the Ninth International Conference on Machine Learning*. Morgan Kaufman Publishers, San Mateo, CA, , pp. 224-232.

19. Hunter, L. and **States, D.J.** (1992) Bayesian Classification of Protein Structure. *IEEE Expert*, **7**, 67-75.

20. **States, D.J.** (1992) Molecular sequence accuracy: analysing imperfect data. *Trends Genet*, **8**, 52-55.

21. **States, D.J.**, Walseth, T.F. and Lee, H.C. (1992) Similarities in amino acid sequences of Aplysia ADP-ribosyl cyclase and human lymphocyte antigen CD38. *Trends Biochem Sci*, **17**, 495.

22. Claverie, J.M. and **States, D.J.** (1993) Information Enhancement Methods for Large Scale Sequence Analysis. *Computers in Chemistry*, **17**, 191-201.

23. Gish, W. and **States, D.J.** (1993) Identification of protein coding regions by database similarity search. *Nat Genet*, **3**, 266-272.

24. Green, P., Lipman, D., Hillier, L., Waterston, R., **States, D.J.** and Claverie, J.M. (1993) Ancient conserved regions in new gene sequences and the protein databases. *Science*, **259**, 1711-1716.

25. **States, D.J.**, Harris, N.L. and Hunter, L. (1993) Computationally efficient cluster representation in molecular sequence megaclassification. *Proc Int Conf Intell Syst Mol Biol*, **1**, 387-394.

26. **States, D.J.**, Hunter, L., Harris, N. and Lipman, D.J. (1993), *Genome Mapping and Sequencings*. Cold Spring Harbor Press, Cold Spring Harbor, NY, Cold Spring Harbor, NY, , Vol. 6.

27. Agarwal, P. and **States, D.J.** (1994) The Repeat Pattern Toolkit (RPT): analyzing the structure and evolution of the C. elegans genome. *Proc Int Conf Intell Syst Mol Biol*, **2**, 1-9.

28. **States, D.J.** and Gish, W. (1994) Combined use of sequence similarity and codon bias for coding region identification. *J Comput Biol*, **1**, 39-50.

29. Agarwal, P. and **States, D.J.** (1996) A Bayesian evolutionary distance for parametrically aligned sequences. *J Comput Biol*, **3**, 1-17.

30. Cooper, M.L., Maffitt, D.R., Parsons, J.D., Hillier, L. and **States, D.J.** (1996) Lane tracking software for four-color fluorescence-based electrophoretic gel images. *Genome Res*, **6**, 1110-1117.

31. **States, D.J.** and Agarwal, P. (1996) Compact encoding strategies for DNA sequence similarity search. *Proc Int Conf Intell Syst Mol Biol*, **4**, 211-217.

32. Huang, W., Yin, Z., Fuhrmann, D.R., **States, D.J.** and Thomas, L.J., Jr. (1997) A method to determine the filter matrix in four-dye fluorescence-based DNA sequencing. *Electrophoresis*, **18**, 23-25.

33. Nagaraja, R., MacMillan, S., Kere, J., Jones, C., Griffin, S., Schmatz, M., Terrell, J., Shomaker, M., Jermak, C., Hott, C. *et al.* (1997) X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. *Genome Res*, **7**, 210-222.

34. Agarwal, P. and **States, D.J.** (1998) Comparative accuracy of methods for protein sequence similarity search. *Bioinformatics*, **14**, 40-47.

35. Huang, W., Fuhrmann, D.R., Politte, D.G., Thomas, L.J., Jr. and **States, D.J.** (1998) Filter matrix estimation in automated DNA sequencing. *IEEE Trans Biomed Eng*, **45**, 422-428.

36. Liu, R., Blackwell, T.W. and **States, D.J.** (1998), *Proceedings of the Second Annual Conference of Computational Biology (RECOMB98)* ACM Press., New York, Vol. 2, pp. 173-181.

37. Polite, D.G., Maffit, D.R. and **States, D.J.** (1998), *Proceedings of the Second Annual Conference on Computational Biology (RECOMB98)* ACM Press., New York, Vol. 6, pp. 202-206.

38. Rouchka, E.C. and **States, D.J.** (1998) Sequence assembly validation by multiple restriction digest fragment coverage analysis. *Proc Int Conf Intell Syst Mol Biol*, **6**, 140-147.

39. Blackwell, T.W., Rouchka, E. and **States, D.J.** (1999) Identity by descent genome segmentation based on single nucleotide polymorphism distributions. *Proc Int Conf Intell Syst Mol Biol*, 54-59.

40. Kan, Z., Gish, W., Rouchka, E., Glasscock, J. and **States, D.J.** (2000) UTR reconstruction and analysis using genomically aligned EST sequences. *Proc Int Conf Intell Syst Mol Biol*, **8**, 218-227.

41. Lee, A.W. and **States, D.J.** (2000) Both src-dependent and -independent mechanisms mediate phosphatidylinositol 3-kinase regulation of colony-stimulating factor 1-activated mitogen-activated protein kinases in myeloid progenitors. *Mol Cell Biol*, **20**, 6779-6798.

42. Lee, A.W. and **States, D.J.** (2000) In Shaw, S., Turni, L. A. and Katz, K. S. (eds.), *Protein Review on the Web.* .

43. Kan, Z., Rouchka, E.C., Gish, W.R. and **States, D.J.** (2001) Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res*, **11**, 889-900.

44. Liu, R., Blackwell, T.W. and **States, D.J.** (2001) Conformational model for binding site recognition by the E.coli MetJ transcription factor. *Bioinformatics*, **17**, 622-633.

45. **States, D.J.** (2001) Time to defend what we have won. *Bioinformatics*, **17**, 299.

46. **States, D.J.** (2001) Describing the release of sequence data. *Science*, **292**, 1066-1067.

47. **States, D.J.**, Nowotny, V. and Blackwell, T.W. (2001) Probabilistic approaches to the use of higher order clone relationships in physical map assembly. *Bioinformatics*, **17 Suppl 1**, S262-269.

48. Kan, Z., **States, D.J.** and Gish, W. (2002) Selecting for functional alternative splices in ESTs. *Genome Res*, **12**, 1837-1845.

49. Liu, R. and **States, D.J.** (2002) Consensus promoter identification in the human genome utilizing expressed gene markers and gene modeling. *Genome Res*, **12**, 462-469.

50. Rouchka, E.C., Gish, W. and **States, D.J.** (2002) Comparison of whole genome assemblies of the human genome. *Nucleic Acids Res*, **30**, 5004-5014.

51. **States, D.J.** (2002) Bioinformatics code must enforce citation. *Nature*, **417**, 588.

52. Liu, R., McEachin, R.C. and **States, D.J.** (2003) Computationally identifying novel NF-kappa B-regulated immune genes in the human genome. *Genome Res*, **13**, 654-661.

53. Adamski, M., Blackwell, T., Menon, R., Martens, L., Hermjakob, H., Taylor, C., Omenn, G.S. and **States, D.J.** (2005) Data management and preliminary data analysis in the pilot phase of the HUPO Plasma Proteome Project. *Proteomics*, **5**, 3246-3261.

54. Martens, L., Hermjakob, H., Jones, P., Adamski, M., Taylor, C., **States, D.J.**, Gevaert, K., Vandekerckhove, J. and Apweiler, R. (2005) PRIDE: the proteomics identifications database. *Proteomics*, **5**, 3537-3545.

55. Omenn, G.S., **States, D.J.**, Adamski, M., Blackwell, T.W., Menon, R., Hermjakob, H., Apweiler, R., Haab, B.B., Simpson, R.J., Eddes, J.S. *et al.* (2005) Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics*, **5**, 3226-3245.

56. Ping, P., Vondriska, T.M., Creighton, C.J., Gandhi, T.K., Yang, Z., Menon, R., Kwon, M.S., Cho, S.Y., Drwal, G., Kellmann, M. *et al.* (2005) A functional annotation of subproteomes in human plasma. *Proteomics*, **5**, 3506-3519.

57. Santos, C., Blake, J. and **States, D.J.** (2005) Supplementary data need to be kept in public repositories. *Nature*, **438**, 738.

58. Santos, C., Eggle, D. and **States, D.J.** (2005) Wnt pathway curation using automated natural language processing: combining statistical methods with partial and full parse for knowledge extraction. *Bioinformatics*, **21**, 1653-1658.

59. Fermin, D., Allen, B.B., Blackwell, T.W., Menon, R., Adamski, M., Xu, Y., Ulintz, P., Omenn, G.S. and **States, D.J.** (2006) Novel gene and gene model detection using a whole genome open reading frame analysis in proteomics. *Genome Biol*, **7**, R35.

60. Huo, J.S., McEachin, R.C., Cui, T.X., Duggal, N.K., Hai, T., **States, D.J.** and Schwartz, J. (2006) Profiles of growth hormone (GH)-regulated genes reveal time-dependent responses and identify a mechanism for regulation of activating transcription factor 3 by GH. *J Biol Chem*, **281**, 4132-4141.

61. Lee, A.W. and **States, D.J.** (2006) Colony-stimulating factor-1 requires PI3-kinase-mediated metabolism for proliferation and survival in myeloid cells. *Cell Death Differ*. 13(11):1900-14.

62. Rauch, A., Bellew, M., Eng, J., Fitzgibbon, M., Holzman, T., Hussey, P., Igra, M., Maclean, B., Lin, C.W., Detter, A. **States, D.J.** and McIntosh, M. (2006) Computational Proteomics Analysis System (CPAS): an extensible, open-source analytic system for evaluating and publishing proteomic data and high throughput biological experiments. *J Proteome Res*, **5**, 112-121.

63. **States, D.J.**, Omenn, G.S., Blackwell, T.W., Fermin, D., Eng, J., Speicher, D.W. and Hanash, S.M. (2006) Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nature Biotechnology*, **24**, 333-338.

64. Carayol N, Chen J, Yang F, Jin T, Jin L, **States D. J.**, Wang CY. (2006) "A dominant function of IKK/NF-kappaB signaling in global lipopolysaccharide-induced gene expression." Journal of Biological Chemistry 281(41):31142-51. Epub 2006 Aug 16.

65. Tian Y, McEachin RC, Santos C, **States D. J.**, Patel JM. (2007) "SAGA: a subgraph matching tool for biological graphs." Bioinformatics. 23(2):232-9.

66. Jayapandian M, Chapman A, Tarcea VG, Yu C, Elkiss A, Ianni A, Liu B, Nandi A, Santos C, Andrews P, Athey B, **States D. J.**, Jagadish HV. (2007) "Michigan Molecular Interactions (MiMI): putting the jigsaw puzzle together." Nucleic Acids Res. 2007 Jan;35(Database issue):D566-71.

67. Li X, Madison BB, Zacharias W, Kolterud A, **States D. J.**, Gumucio DL. (2007) "Deconvoluting the intestine: molecular evidence for a major role of the mesenchyme in the modulation of signaling crosstalk." Physiological Genomics. 29(3):290-301.

68. Rosania GR, Crippen G, Woolf P, **States D. J.**, Shedden K. (2007) "A Cheminformatic Toolkit for Mining Biomedical Knowledge." Pharmaceutical Research 24(10):1791-1802.

69. Chen J, Blackwell TW, Fermin D, Menon R, Chen Y, Gao J, Lee AW, **States D. J.**, (2007) "Evolutionary-conserved gene expression response profiles across mammalian tissues." OMICS. 2007 Spring;11(1):96-115.

70. Rao A., Hero, A.O. III, **States D.J.**, and Engel, J.D. (2007) "Inferring Time-Varying Network Topologies from Gene Expression Data." EURASIP Journal on Bioinformatics and Systems Biology, 2007:51947, 12 pages

71. Chen, Y., Blackwell, T.W., Chen, J., Gao, J., Lee, A.W. and **States, D.J.** (2007) Integration of Genome and Chromatin Structure with Gene Expression Profiles to Predict c-MYC Recognition Site Binding and Function. PLoS Computational Biology Apr 6;3(4):e63.

72. Rao A., Hero, A.O. III, **States D.J.**, and Engel, J.D. (2007) "Using directed information to build biologically relevant influence networks." Computional Systems Bioinformatics Conference 6:145-56.

73. Rao A., Hero, A.O. III, **States D.J.**, and Engel, J.D. (2007) "Motif discovery in tissue-specific regulatory sequences using directed information." EURASIP J Bioinform Syst Biol. 2007:13853.

74. Mathivanan S, Ahmed M, Ahn NG, Alexandre H, Amanchy R, Andrews PC, Bader JS, Balgley BM, Bantscheff M, Bennett KL, Björling E, Blagoev B, Bose R, Brahmachari SK, Burlingame AS, Bustelo XR, Cagney G, Cantin GT, Cardasis HL, Celis JE, Chaerkady R, Chu F, Cole PA, Costello CE, Cotter RJ, Crockett D, DeLany JP, De Marzo AM, DeSouza LV, Deutsch EW, Dransfield E, Drewes G, Droit A, Dunn MJ, Elenitoba-Johnson K, Ewing RM, Van Eyk J, Faca V, Falkner J, Fang X, Fenselau C, Figeys D, Gagné P, Gelfi C, Gevaert K, Gimble JM, Gnad F, Goel R, Gromov P, Hanash SM, Hancock WS, Harsha HC, Hart G, Hays F, He F, Hebbar P, Helsens K, Hermeking H, Hide W, Hjernø K, Hochstrasser DF, Hofmann O, Horn DM, Hruban RH, Ibarrola N, James P, Jensen ON, Jensen PH, Jung P, Kandasamy K, Kheterpal I, Kikuno RF, Korf U, Körner R, Kuster B, Kwon MS, Lee HJ, Lee YJ, Lefevre M, Lehvaslaiho M, Lescuyer P, Levander F, Lim MS, Löbke C, Loo JA, Mann M, Martens L, Martinez-Heredia J, McComb M, McRedmond J, Mehrle A, Menon R, Miller CA, Mischak H, Mohan SS, Mohmood R, Molina H, Moran MF, Morgan JD, Moritz R, Morzel M, Muddiman DC, Nalli A, Navarro JD, Neubert TA, Ohara O, Oliva R, Omenn GS, Oyama M, Paik YK, Pennington K, Pepperkok R, Periaswamy B, Petricoin EF, Poirier GG, Prasad TS, Purvine SO, Rahiman BA, Ramachandran P, Ramachandra YL, Rice RH, Rick J, Ronnholm RH, Salonen J, Sanchez JC, Sayd T, Seshi B, Shankari K, Sheng SJ, Shetty V, Shivakumar K, Simpson RJ, Sirdeshmukh R, Siu KW, Smith JC, Smith RD, **States D. J.**, Sugano S, Sullivan M, Superti-Furga G, Takatalo M, Thongboonkerd V, Trinidad JC, Uhlen M, Vandekerckhove J, Vasilescu J, Veenstra TD, Vidal-Taboada JM, Vihinen M, Wait R, Wang X, Wiemann S, Wu B, Xu T, Yates JR, Zhong J, Zhou M, Zhu Y, Zurbig P, Pandey A. (2008) "Human Proteinpedia enables sharing of human protein data." Nature Biotechnology 26(2):164-7.

75. **States, D.J.**, Ade AS, Wright ZC, Bookvich AV, Athey BD. (2008) "MiSearch Adaptive PubMed Search Tool." Bioinformatics. [Epub ahead of print]

76. Du X, Yang F, Manes NP, Stenoien DL, Monroe ME, Adkins JN, **States, D.J.**, "Purvine SO, Camp Ii DG, Smith RD. (2008) "Linear Discriminant Analysis-Based Estimation of the False Discovery Rate for Phosphopeptide Identifications." J Proteome Res. [Epub ahead of print]

77. Ashkenazi M, Bader GD, Kuchinsky A, Moshelion M, **States, D.J.** (2008) "Cytoscape ESP: simple search of complex biological networks." Bioinformatics [Epub ahead of print]

78. Elkiss, A, Shen, S; Fader, A, Erkan, G., **States, D.J.** and Radev D. (2008) "Blind men and elephants: What do citation summaries tell us about a research article?" Journal of the American Society for Information Science and Technology 59:51-62.

Book chapters

1. **States, D. J.**, and Boguski, M. S. (1991), "Similarity and Homology in Molecular Sequence," in A Sequence Analysis Primer, eds. M. Gribskov and J. Devereux, Stockton Press (New York, London).
2. Boguski, M. S., Ostell, J., and **States, D. J.** (1992), "Molecular Sequence Databases and Their Uses," in Protein Engineering: A Practical Approach, eds., A.R. Rees, R. Wetzel, and M.J.E. Sternberg, IRL Press (Oxford, UK and Washington, D.C.).
3. **States, D.J.**, and Reisdorf, Jr., W.C., (1998) "Classification-Based Molecular Sequence Analysis," Chapter 6, in *Computational Methods in Molecular Biology*, S.L. Salzberg, D.B. Searls, S. Kasif, Eds., Elsevier Science B.V., pp. 87-106.
4. Gilbert S. Omenn, **David J. States**, Marcin Adamski, Thomas W. Blackwell, Rajasree Menon, Henning Hermjakob, Rolf Apweiler, et. al (2006) "Overview of the HUPO Plasma Proteome Project: Results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database," in *Exploring the Human Proteome*. Gilbert S. Omenn ed. Wiley-VCH Wienheim.
5. Peipei Ping, Thomas M. Vondriska, Chad J. Creighton, TKB Gandhi, Ziping Yang, Rajasree Menon, Min-Seok Kwon, Sang Yun Cho, Garry Drwal, Markus Kellmann, Suraj Peri, Shubha Suresh, Mads Gronborg, Henrik Molina, Raghothama Chaerkady, B. Rekha, Arun S. Shet, Robert E. Gerszten, Haifeng Wu" Mark Raftery, Valerie Wasinger, Peter Schulz-Knappe, Samir M. Hanash, Young-ki Paik, William S. Hancock, **David J. States**, Gilbert S. Omenn and Akhilesh Pandey (2006) "A functional annotation of subproteomes in human plasma," in *Exploring the Human Proteome*. Gilbert S. Omenn ed. Wiley-VCH Wienheim.

Non-reviewed articles and letters

1. **States, D.J.** [Book Review] "Calculating the Secrets of Life". *Computers and Chemistry*, 21(4):279-280, 1997.
2. **States, D.J.** (1996), "New Software - Dead or Alive, " [Letter] *Byte*, 21(12):19.