

CURRICULUM VITAE

DATE LAST UPDATED: January 4, 2011

FULL NAME: Indra Neil Sarkar

CURRENT POSITIONS: Director of Biomedical Informatics,
Assistant Professor of Clinical and Translational Science,
Assistant Professor of Microbiology and Molecular Genetics,
& Assistant Professor (Secondary) of Computer Science

UNIVERSITY ADDRESS: Center for Clinical and Translational Science
University of Vermont
89 Beaumont Avenue, Given Courtyard N309
Burlington, VT 05405 USA
+1 (802) 656 8283
E-Mail: neil.sarkar@uvm.edu
Web: <http://www.uvm.edu/~insarkar>

EDUCATION

1999 **BSc** (Microbiology)
Advisor: Douglas B. Luckie, PhD
Michigan State University, East Lansing, MI

2002 **MPhil** (Biomedical Informatics)
Advisor: David H. Figurski, PhD
Columbia University, New York, NY

2004 **PhD** (Biomedical Informatics)
Advisor: David H. Figurski, PhD
Columbia University, New York, NY

2008 **MLIS** (Library & Information Science)
Advisor: Jian Qin, PhD
Syracuse University, Syracuse, NY

2008 **Certificate of Advanced Study** (Digital Libraries)
Syracuse University, Syracuse, NY

MAJOR ADMINISTRATIVE RESPONSIBILITIES

2006-2008 *Informatics Manager*, MBLWHOI Library,
Marine Biological Laboratory, Woods Hole, MA

2009-Pres *Director of Biomedical Informatics*,
Center for Clinical and Translational Science, University of Vermont,
Burlington, VT

ACADEMIC APPOINTMENTS

- 2004-2006 **Bioinformatics Associate**, Division of Invertebrate Zoology, American Museum of Natural History, New York, NY
- 2004-2006 **Instructor**, Downstate Medical Center, State University of New York, New York, NY
- 2007-2008 **Assistant Research Scientist**, MBLWHOI Library, Marine Biological Laboratory, Woods Hole, MA
- 2009-Pres **Assistant Professor**, Center for Clinical and Translational Science, University of Vermont, Burlington, VT
- 2009-Pres **Assistant Professor**, Department of Microbiology and Molecular Genetics, College of Medicine, University of Vermont, Burlington, VT
- 2009-Pres **Assistant Professor (Secondary Appointment)**, Department of Computer Science, College of Engineering and Mathematical Sciences, University of Vermont, Burlington, VT

MAJOR COMMITTEE ASSIGNMENTS

- 1999-2003 **Ex-Officio** (2003); **Chair** (2002); **Vice-Chair** (2001); **Member** (1999-2002), Graduate Student Advisory Group, Columbia University
- 2000-2002 **Student Representative**, Department of Biomedical Informatics, College of Physicians and Surgeons, Columbia University, New York, NY
- 2002-2003 **Member**, Advisory Committee, Columbia Speakers Fund, Columbia University, New York, NY
- 2009-Pres **Member**, Center for Clinical and Translational Science – Agency for Human Services, State of Vermont, Waterbury, VT
- 2009-Pres **Member**, Informatics Advisory Group, Fletcher Allen Health Care, Burlington, VT
- 2009-Pres **Member**, Regional Health Information Technology Advisory Board, State of Vermont, Montpelier, VT
- 2009-Pres **Member**, Health Information Technology Higher Education Advisory Group, State of Vermont, Montpelier, VT
- 2009 **Delegate**, Vermont Delegation to the National Governors Association for Health Information Technology and Implementation, Montpelier, VT
- 2009 **Member**, Health Care Payment Reform Committee (Convened by Senator Bill Carris and Representative Anne O'Brien, Co-Chairs), Montpelier, VT
- 2009-Pres **Member**, EHR Cause of Death Working Group, National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD
- 2009 **Member**, Public Health Working Group, Transdisciplinary Research Initiative, University of Vermont, Burlington, VT
- 2009-2010 **Member**, Clinical Epidemiologist Faculty Search Committee, Department of Surgery, College of Medicine, University of Vermont, Burlington, VT
- 2009-Pres **Chair**, Microbiology and Molecular Genetics Retreat Committee, University of Vermont, Burlington, VT
- 2010-Pres **Member**, Gund Professor for Ecological Economics & Director of the Gund Institute for Ecological Economics Search Committee, Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, VT

- 2010-Pres **Co-Chair**, Ad Hoc Committee on Scholarly Communication, University of Vermont, Burlington, VT
- 2010-Pres **Member**, Animal Genetics/Genomics Faculty Search Committee, Department of Animal Science, College of Agriculture and Life Sciences, University of Vermont, Burlington, VT
- 2010-Pres **Member**, Work Group on Privacy, Information Security and Data Retention (convened by the Chief Compliance Officer and Chief Information Officer), University of Vermont, Burlington, VT

HOSPITAL OR AFFILIATED INSTITUTION APPOINTMENTS

- 2004-Pres **Affiliated Research Scientist**, Department of Biomedical Informatics, Columbia University, New York, NY
- 2008-Pres **Research Associate**, American Museum of Natural History, New York, NY
- 2009-Pres **Adjunct Assistant Scientist**, MBL/WHOI Library, Marine Biological Laboratory, Woods Hole, MA

OTHER PROFESSIONAL POSITIONS

- 1995 **Lab Practicum**, Institut fuer Klinische und Molekulare Virologie, Frederick Alexandr Universitaat Erlangen-Nurenburg, Erlangen, Bavaria, Germany
- 1996 **Research Intern**, Research Computing Division, Genetics Institute, Cambridge, MA
- 1996-1998 **Research Assistant**, Department of Microbiology, Michigan State University, East Lansing, MI
- 1997-1999 **Teaching Assistant & Student Instructor**, Lyman Briggs College Michigan State University, East Lansing, MI
- 1998-1999 **Research Assistant**, Department of Physiology, Michigan State University, East Lansing, MI
- 1999-2004 **Graduate Research Fellow**, Department of Biomedical Informatics, College of Physicians and Surgeons, Columbia University, New York, NY
- 2001 **Research Assistant**, Lister Hill Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Bethesda, MD
- 2006-2008 **Lecturer**, NLM/MBL Biomedical Informatics Short Course, Woods Hole, MA

PROFESSIONAL SOCIETIES

- 1995-Pres **Member**, American Association for the Advancement of Science
2000-Pres **Member**, American Medical Informatics Association
Meetings Committee Member (2009-Pres)
Public Policy Committee Member (2009-Pres)
2000-Pres **Member**, International Society for Computational Biologists
Conferences Committee Member (2005-Pres)
2004-Pres **Member**, International Society for Infectious Diseases
2005-Pres **Member**, Medical Library Association
2005-Pres **Member**, Computer Society, Institute of Electrical and Electronics Engineers
2005-2010 **Member**, Association for Computing Machinery
2007-2009 **Member**, Gerontological Society of America

HONORS AND AWARDS

- 1999 Medical Informatics Training Fellowship, National Library of Medicine, National Institutes of Health
2001 Summer Research Fellowship, Lister Hill Center for Biomedical Communications, National Library of Medicine, National Institutes of Health
2002 Who's Who in America (56th Edition), Marquis Publishing Group
2003 International Lipari School Scholarship, Italian Ministry of Research and Universities
2006 Donald A.B. Lindberg Research Fellowship, Medical Library Association
2007 Who's Who of Emerging Leaders (1st Edition), Marquis Publishing Group
2007 The Honor Society of Phi Kappa Phi
2008 Award of Recognition for "outstanding contributions to the launch and success of the Encyclopedia of Life Project," Alfred P. Sloan Foundation
2008 Golden Key Society
2009 Beta Phi Mu (International Society for Librarianship and Information Science)

RESEARCH

My work is dedicated to furthering a biomedical informatics research agenda across the entire spectrum of biomedicine, from molecules to populations. Specific research involves the development and use of a range of computational techniques (including knowledge gathering and discovery methods, phylogenetics, information theory, ontology development, semantic indexing, and natural language processing) to facilitate the analysis and linking of molecular and public health data. Ultimately, my research aims to enable the creation of testable models of disease and provide a framework to enable the assessment of comparative hypotheses across the spectrum of biomedicine and health care.

GRANT AND CONTRACT SUPPORT

Current

- 2010-2012 **Co-Investigator**
R00LM009825: “Informatics for Zoonotic Disease Surveillance: Combining Animal and Human Data.” The goal of this project is to develop an infrastructure for zoonotic disease surveillance.
PI: Matthew Scotch (ASU)
National Library of Medicine, National Institutes of Health
\$38,018 Total Award to UVM (1.0 CM Current Effort)
- 2010-2013 **Co-PI (Biomedical Informatics)**
Children’s Health Insurance Program Reauthorization Act of 2009 (CHIPRA): Maine & Vermont. “CMS Children’s Quality Demonstration Grant.” This project aims to (1) allow research and evaluation of new evidence-based measures to provide an opportunity to identify the challenges in gathering and reporting data to providers and consumers; (2) promote health information technology in children’s health care delivery; and, (3) evaluate of provider-based models that improve delivery of children’s health care.
PI: Judith Shaw (UVM)
\$11,277,362 Total Award (1.2 CM Current Effort)
- 2008-2011 **PI**
R01LM009725: “Enhancing Organism Based Disease Knowledge Via Name Based Taxonomic Intelligence.” The goal of this project is to develop linkages between biodiversity and biomedical knowledge through semantic technologies to identify plausible hypotheses within the context of infectious diseases.
National Library of Medicine, National Institutes of Health
\$1,053,540 Total Award (2.4 CM Current Effort)
- 2007-2011 **Consultant (2009-Pres); Project Director (2007-2008)**
“Bridging Biomedical and Biodiversity Knowledge Via Taxonomic Information.” The goal of this project is to develop an informational portal for consumers and experts in the molecular biology of aging that interfaces with the Encyclopedia of Life project.
PI: Gary G. Borisy (MBL)
Ellison Medical Foundation
\$1,954,885 Total Award (0.0 CM Current Effort; Consultant only)

Previous

2010

PI (Biomedical Informatics)

Vermont Agency for Human Services: "Project 3: Create Decision Support Systems, quality and cost data analysis to support state sponsored and other program beneficiaries in receiving the highest quality and most cost effective services." This 9-month project established infrastructure for an extensible research data warehouse that will ultimately serve researchers and government officials for the State of Vermont. The initial data will include vital record data, claims (Medicaid as well as all-payer claims data) and clinical data from Fletcher Allen Health Care.

Overall PI: Judith Shaw (UVM)

Vermont Department of Health

\$1,074,307 Total Award (2.0 CM Current Effort)

2008-2010

Co-Investigator (Biomedical Informatics)

H79SM05880: "Mental Health Intergovernmental Service System Interactive Online Network for Vermont" Through an intergovernmental partnership that strives to build infrastructure for mental health, the goal of this project is to address the needs of Vermont veterans and other adults with trauma spectrum illness who are also involved with the criminal justice system.

PI: Thomas Simpatico (UVM)

Substance Abuse and Mental Health Services Administration

\$2,100,000 Total Award (0.6 CM Current Effort)

2004-2010

PI (2008-2010); **Co-PI** (2006-2008); **Investigator** (2004-2006)

IIS 0241229: "Development of New Digital Library Applications in the Context of a Basic Ontology for Biosystematics Information Using the Literature of Entomology (Ants)." As part of an international collaboration (German and the United States), the goal of this project is to digitize a museum collection and develop ontology driven methods for its organization, focusing on literature pertaining to ants.

National Science Foundation

\$559,245 Total Award

2006-2007

PI

2006 DAB Lindberg Fellowship: "Bridging Biomedical and Biodiversity Knowledge via Taxonomic Information." The goal of this project was to organize taxonomic literature from biomedical and biodiversity resources into a form that can be browsed from a centralized Web interface.

Medical Library Association

\$25,000 Total Award

2004-2007 **Investigator**
“Virtual Center for Plant Evolutionary Genomics.” The goal of this project was to develop systematic approaches for organizing and curating plant genomes using high-throughput data (e.g., ESTs). Particular emphasis was placed in the advancement of phylogenetic techniques and guiding the development of automated techniques to facilitate high-throughput phylogenetic analyses.
PI: Gloria M. Coruzzi (NYU)
\$4,994,582.00 Total Award

TEACHING

Full Courses

1994 *Chemistry* [Accelerated High School Course]
Educ. Stud. Prog., Massachusetts Institute of Technology, Cambridge, MA

1994 *Calculus I (AB)* [Accelerated High School Course]
Educ. Stud. Prog., Massachusetts Institute of Technology, Cambridge, MA

2004-2006 *Introduction to Medical Informatics* [400 Graduate Level]
Downstate Medical Center, State University of New York, New York, NY

2004-2006 *Clinical Decision Support Systems* [400 Graduate Level]
Downstate Medical Center, State University of New York, New York, NY

2004-2006 *Internet Integration into Health Care* [400 Graduate Level]
Downstate Medical Center, State University of New York, New York, NY

2009-Pres *Introduction to Biomedical Informatics (CTS 395)*
Center for Clinical and Translational Science, University of Vermont,
Burlington, VT

2010 *Programming for Bioinformatics (MMG 231; Offered Springs of Even Years)*
Department of Microbiology and Molecular Genetics, University of Vermont,
Burlington, VT

2011 *Methods in Bioinformatics (MMG 296; Offered Springs of Odd Years)*
Department of Microbiology and Molecular Genetics, University of Vermont,
Burlington, VT

Individual Lectures / Course Units

1997-1999 *Biocomputing* [Undergraduate Lab Course]
Lyman Briggs College, Michigan State University, East Lansing, MI

2005 *Organizing and Understanding the Biological Data Deluge through Phylogenetics* [Conference Tutorial]
IEEE Computational Systems Bioinformatics Conference. Stanford, CA.

2006-2008 *Bioinformatics* [CME Level Survey Course]
NLM/MBL Biomedical Informatics Short Course, Woods Hole, MA

2007 *Biodiversity Informatics: Managing Knowledge Beyond Humans and Model Organisms* [Conference Tutorial]
Pacific Symposium on Biocomputing. Maui, HI.

2009 *Introduction to Biodiversity Informatics* [Graduate Level Course]
Annelid Tree of Life – Clitellata Lecture Series, Gothenburg, Sweden

- 2009 *Finding Needles in the Biodiversity Haystack: Secondary Uses of Biological Information* [Graduate Level Course]
Annelid Tree of Life – Clitellata Lecture Series, Gothenburg, Sweden

MENTORING

Thesis Committees

- 2009-2010 Comm. Chair– Violet Roskens (MSc, Biology, UVM)
2009-2010 Comm. Member– Sasi Kunta (MSc, Computer Science, UVM)
2009-Pres Comm. Member– Saurav Acharya (PhD, Computer Science, UVM)
2009-Pres Comm. Member– Sebastian Kvist (PhD, Comparative Biology, AMNH)
2010-Pres Primary Advisor– Ahmed Nabhan (PhD, Computer Science, UVM)
2010-Pres Comm. Member– Peter Durda (PhD, Clinical and Translational Science, UVM)
2010-Pres Comm. Member– W. Gabe Tharpe (PhD, Clinical and Translational Science, UVM)
2010-Pres Comm. Member– Adam Sateriale (PhD, Cell and Molecular Biology, UVM)

Undergraduate Researchers

- 2009-Pres Patrick Benson (Biomedical Informatics/Problem Oriented Medical Record)
2010-Pres Benjamin Earle (Computational Biology/Natural Language Processing)
2010-Pres Ryan Dundon (Computational Biology/Phylogenetics)

High School Students

- 2004-2006 Atin Saha (Bioinformatics Research Project, AMNH)
2005-2006 Judy Ri (Bioinformatics Research Project, AMNH)
2005-2006 Subashis Paul (Bioinformatics Research Project, AMNH)
2005-2006 Lakshman Sankar (Bioinformatics Research Project, AMNH)
2005-2006 Tanu Chauhan (Bioinformatics Research Project, AMNH)

SERVICE / ADMINISTRATION

Editorships

- 2005-2007 **Guest Editor**, *Journal of Biomedical Informatics* (Topic: Phylogenetics)
2007-Pres **Editorial Board Member**, *Recent Patents in Computer Science*
2009-Pres **Editorial Board Member**, *Journal of Biomedical Informatics*
2009-Pres **Editorial Board Member**, *International Knowledge Discovery in Bioinformatics*
2009-Pres **Editorial Board Member**, *Briefings in Bioinformatics*
2008 **Guest Editor**, *BMC Bioinformatics (Selected Papers from the 2008 AMIA Summit on Translational Bioinformatics)*
2009 **Guest Editor**, *BMC Bioinformatics (Selected Papers from the 2009 AMIA Summit on Translational Bioinformatics)*
2009 **Guest Editor**, *BMC Bioinformatics* (Topic: Biodiversity Informatics)
2010-Pres **Guest Editor**, *Briefings in Bioinformatics* (Topic: Education in Bioinformatics)
2011-Pres **Editorial Board Member**, *Journal of the American Medical Informatics Association*
2011-Pres **Associate Editor**, *Methods of Information in Medicine*

Journal / Book / Conference Reviews

2004	Pacific Symposium on Biocomputing
2005-Pres	Bioinformatics
2006-Pres	Journal of Biomedical Informatics
2006-Pres	BMC Bioinformatics
2006	World Scientific Press
2007	Association for Computational Linguists: BioNLP
2007	International Symposium on Languages in Biology & Medicine
2007-Pres	American Medical Informatics Association Annual Symposium
2007-Pres	BMC Source Code for Biology and Medicine
2008-Pres	IMIA Yearbook of Medical Informatics
2009	Artificial Intelligence in Medicine
2009-Pres	PLoS ONE
2010	Journal of Translational Medicine

Conference Organizing

2006	Symposium Session Chair , AAAS Annual Symposium. St. Louis, MO.
2007	Track Chair , Pacific Symposium on Biocomputing. Big Island, Hawaii.
2007-2008	Founding Track Chair , AMIA Summit on Translational Bioinformatics. San Francisco, CA.
2008	Workshop Organizer , Data Analysis Working Group 2008 Hack-a-thon. EOL BioSynC, Field Museum of Natural History. Chicago, IL.
2008-2009	Vice Chair & Track Chair , AMIA Summit on Translational Bioinformatics. San Francisco, CA.
2008-2009	Member , Technical Implementation Board, Third International Conference for the Barcode of Life. Mexico City, Mexico.
2009-2010	Track Chair , AMIA Summit on Translational Bioinformatics. San Francisco, CA.
2010	Member , Program Committee, IEEE ICDM-2010 workshop on Biological Data Mining and its Applications in Healthcare. Sydney, Australia.
2010-Pres	Chair , Scientific Program Committee. AMIA Summit on Translational Bioinformatics. San Francisco, CA.

External Advisory Panels / Boards

1991-Pres	Member , Board of Directors, Mosaic Data Systems, Inc., Bedford, MA
1996-1999	President & Co-Founder , Board of Directors, InterDream Corporation, Bedford, MA
1999-2008	Vice-President (2004-2008); Assistant Treasurer (2003-2004); Director (1999-2008), Lyman Briggs Alumni Association, Michigan State University, East Lansing, MI
2002-2005	Member , Advisory Committee, Wiley Interscience, Hoboken, NJ
2006-2007	Member , Data Curation Education Program, University of Illinois Urbana-Champaign
2007-Pres	Chair , Data Analysis Working Group, Barcode of Life Initiative, Consortium for the Barcode of Life, Washington, DC

- 2007-Pres **Member**, Implementation Board, Barcode of Life Initiative, Consortium for the Barcode of Life, Washington, DC
- 2009-Pres **Member**, Electronic Health Record - System, Vital Records Functional Work Group, HL7 EHR Technical Committee, National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD

Grant Review

- 2007 Information and Intelligent Systems/Information Integration and Informatics, National Science Foundation
- 2008 Information and Intelligent Systems/Cyber Enabled Discovery and Innovation, National Science Foundation
- 2009 Emerging Technologies and Training in Neurosciences, Challenge Grants Internet Assisted Review, National Institutes of Health
- 2009 Bioengineering Sciences and Technologies, Challenge Grants Internet Assisted Review, National Institutes of Health
- 2010 Revisionary Syntheses in Systematics, National Science Foundation
- 2010 Assembling the Tree of Life, National Science Foundation

PUBLICATIONS

Patents

1. Lussier YA, **Sarkar IN**, Cantor MN. *System and Method for Generating an Amalgamated Database*. Provisional Patent: Columbia University Science and Technology Ventures, Health Sciences Division; US PTO 20060074991; Priority: November 2002.
2. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. *CAOS: Method and System for Recognizing One or More Characteristic Attributes of One or More Groups of Information Items*. PCT/US02/03540; 60/267,972; Priority: February 2001.

Theses

1. **Sarkar IN**. Automated Techniques for the Identification and Classification of Evolutionarily Significant Genomic Features. [Doctoral Dissertation] Graduate School of Arts and Sciences at the College of Physicians and Surgeons, Columbia University, 2004.
2. **Sarkar IN**. Biocomputing: Addressing the Black-Box Syndrome. [Bachelor's Capstone Essay] Lyman Briggs College, Michigan State University, 1999.

Software Applications

1. **Sarkar IN**, Schenk R. *LigerCat*. (2007)
2. Koning D, **Sarkar IN**, Moritz TD. *TaxonGrab*. (2005)
3. **Sarkar IN**. *ORFcurator*. (2004)
4. Cantor M, **Sarkar IN**, Lussier YA. *GenesTrace*. (2003)
5. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. *Phylogenomenclature*. (2002)
6. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. *CAOS*. (2002)

Peer-Reviewed Journal

1. Min G-S, **Sarkar IN**, Siddall ME. "Salivary transcriptome of the North American medicinal leech, *Macrobdella decora*." *Journal of Parasitology*. 2010 (In Press).
2. Kvist SB, **Sarkar IN**, Erséus C. "Genetic variation, phylogeny, and invasiveness of the cosmopolitan marine genus *Tubificoides* (Annelida: Clitellata: Naididae: Tubificinae)." *Molecular Phylogenetics and Evolution*. 2010 Nov;57(2):687-702.
3. Scotch M, Mei C, Brandt C, **Sarkar IN**, Cheung K. "At the intersection of public health informatics and bioinformatics: using advanced Web technologies for phylogeography." *Epidemiology*. 2010 Nov; 21(6): 764-8.
4. **Sarkar IN**. "Biomedical Informatics and Translational Medicine." *J Transl Med*. 2010 Feb 26;8:22.
5. Chen ES, **Sarkar IN**. "MeSHing Molecular Sequence and Clinical Trials: A Feasibility Study." *Journal of Biomedical Informatics*. 2010 Jun; 43(3):442-450.
6. Melton GB, Raman N, Chen ES, **Sarkar IN**, Pakhomov S, Madoff RD. "Evaluation of Family History Information within Clinical Documents and Adequacy of HL7 Clinical Statement and Clinical Genomics Family History Models for Its Representation." *J Am Med Inform Assoc*. 2010 May 1; 17(3):337-340.
7. Lienau EK, DeSalle R, Allard M, Brown EW, Swofford D, Rosenfeld JA, **Sarkar IN**, Planet PJ. "The mega-matrix tree of life: using genome-scale horizontal gene transfer and sequence evolution data as information about the vertical history of life." *Cladistics*. 2010 August 26; 26: 1-11.
8. Miller H, Norton CN, **Sarkar IN**. "GenBank and PubMed: How connected are they?" *BMC Res Notes*. 2009 Jun 9;2(1):101.
9. **Sarkar IN**, Planet PJ, DeSalle R. "CAOS Software for Use in Character Based DNA Barcoding." *Molecular Ecology Resources*. 2008;8(6):1256-1259.
10. **Sarkar IN**, Schenk R, Norton CN. "Exploring Historical Trends Using Taxonomic Name Metadata." *BMC Evolutionary Biology*. 2008 8:144.
11. Garcia-España A, Chung PJ, **Sarkar IN**, Stiner E, Sun TT, Desalle R. "Appearance of new tetraspanin genes during vertebrate evolution." *Genomics*. 2008 91(4):326-34.
12. **Sarkar IN**, Egan MG, Coruzzi GM, Lee EK, DeSalle R. "Automated Simultaneous Analysis Phylogenetics (ASAP): An Enabling Tool for Phylogenomics." *BMC Bioinformatics*. 2008 19;9:103.
13. Rach J, DeSalle R, **Sarkar IN**, Schierwater B, Hadrys H. "Character-based DNA barcoding allows discrimination of genera, species, and populations in Odonata." *Proceedings of the Royal Society B: Biology*. 2008 275(1632):237-47.
14. **Sarkar IN**. "Biodiversity Informatics: Organizing and Linking Information Across the Spectrum of Life." *Briefings in Bioinformatics*. 2007 8(5):347-57.
15. Leary PR, Remsen DP, Norton CN, Patterson DJ, **Sarkar IN**. "uBioRSS: Tracking Taxonomic Literature Using RSS." *Bioinformatics*. 2007 23(11):1434-1436.
16. Perkins SL, **Sarkar IN**, Carter R. "The Phylogeny of Rodent Malaria Parasites: Simultaneous Analysis Across Three Genomes." *Infection, Genetics and Evolution*. 2007 7(1):74-83.
17. Kelly RP, **Sarkar IN**, Eernisse DJ, DeSalle R. "DNA Barcoding Using Chitons (genus *Mopalia*). *Molecular Ecology Notes*. 2007 7(2): 177-183.
18. Chiu JC, Lee EK, Egan MG, **Sarkar IN**, Coruzzi GM, DeSalle R. "OrthologID: Automation of Genome Scale Ortholog Identification within a Parsimony Framework."

Bioinformatics 2006 22(6):699–707.

19. Planet PJ, **Sarkar IN**. “mILD: A Tool for Constructing and Analyzing Matrices of Pairwise Character Incongruence Tests.” *Bioinformatics* 21:4423–4224. 2005.
20. Koning D, **Sarkar IN**, Moritz TD. “TaxonGrab: Extracting Taxon Names from Text.” *Journal of Biodiversity Informatics* 2;79–82. 2005.
21. **Sarkar IN**, Rosenfeld JA, Planet PJ, Figurski DH, DeSalle R. “ORFcurator: Molecular Curation of Genes and Gene Clusters in Prokaryotic Organisms.” *Bioinformatics* 20: 3462–3465. 2004.
22. **Sarkar IN**, Cantor MN, Hartel F, Bodenreider O, Lussier YA. “An Evaluation of Hybrid Methods for Matching Biomedical Terminologies: Mapping the Gene Ontology to the UMLS.” *Stud Health Technol Inform* 95:62–7. 2003.
23. **Sarkar IN**, Planet PJ, Bael TE, Stanley SE, Siddall M, DeSalle R, Figurski DH. “Characteristic Attributes in Cancer Microarrays.” *Journal of Biomedical Informatics* Apr/May;35(2):111–122. 2002.
24. **Sarkar IN**, Thornton J, Planet PJ, Schierwater B, and DeSalle R. “A systematic method for classification of novel homeoboxes.” *Molecular Phylogenetics and Evolution* Sep;24(3):388–399. 2002.
25. **Sarkar IN** and Starren JB. “Desiderata for Personal Electronic Messaging in Clinical Systems.” *Journal of the American Medical Informatics Association* May/Jun;9:209–216. 2002.
26. Planet PJ, DeSalle R, Sidall ME, Bael T, **Sarkar IN**, Stanley SE. “Systematic Analysis of DNA Microarray Data: Ordering and Interpreting Patterns of Gene Expression.” *Genome Research* Jul;11(7):1149–1155. 2001.

Peer-Reviewed Conference

1. **Sarkar IN**. “Leveraging Biomedical Ontologies and Annotation Services to Organize Microbiome Data from Mammalian Hosts.” *AMIA Annu Symp Proc*. 2010. (In Press)
2. Chen ES, Melton G, Engelstad M, **Sarkar IN**. “Standardizing Clinical Document Names Using the HL7/LOINC Document Ontology and LOINC Codes.” *AMIA Annu Symp Proc*. 2010. (In Press)
3. **Sarkar IN**, Schenk R, Miller H, Norton CN. “LigerCat: Using ‘Mesh Clouds’ from Journal, Article, or Gene Citations to Facilitate the Identification of Relevant Biomedical Literature.” AMIA Annual Meeting, San Francisco. *AMIA Annu Symp Proc* 563-567. 2009.
4. **Sarkar IN**, Agrawal A. “Literature Based Discovery of Gene Clusters Using Phylogenetic Methods.” *AMIA Annu Symp Proc* 689–693. 2006.
5. **Sarkar IN**, Cantor MN, Bodenreider O, Lussier YA. “GenesTrace: Phenomic Knowledge Discovery Via Structured Terminology.” *Pacific Symposium Biocomputing* 10:103–114. 2005.
6. **Sarkar IN**, Cantor MN, Gelman R, Hartel F, Lussier YA. “Linking Biomedical Language Information and Knowledge Resources: GO and UMLS.” *Pacific Symposium Biocomputing* 8:427–450. 2003.
7. Lussier YA, **Sarkar IN**, Cantor MN. “An Integrative Model for In-Silico Clinical Genomics Discovery Science.” *AMIA Annu Symp Proc* 469–473. 2002.
8. **Sarkar IN** and Rindflesch TC. “Discovering Protein Similarity Using Natural Language Processing.” *AMIA Annu Symp Proc* 677–681. 2002.

Conference Abstracts

1. **Sarkar IN**. "Using Biomedical Ontologies To Enable Morphology Based Phylogenetics: A Feasibility Study for Fishes." *Bio-Ontologies Special Interest Group 2010, Intelligent Systems for Molecular Biology*, Boston, MA. 2010.
2. **Sarkar IN**. "Deriving and Using Barcode Based Diagnostics for Information Retrieval." *Second International Barcode of Life Conference*, Taipei, Taiwan. 2007.
3. **Sarkar IN**. "Public Health Through the Lens of Biodiversity." *AMIA Spring Congress*, Orlando, FL. 2007.
4. **Sarkar IN**, Leary PR, Norton CN. "A Customizable 'Mash-up' for Model and Disease Organisms." *Medical Library Association Annual Meeting*, Philadelphia, PA. 2007.
5. Umejei C, **Sarkar IN**. "zMedline: A Medline-Based Content Management System." *AMIA Annual Meeting*, Washington, DC: 469–473. 2006.
6. **Sarkar IN**, Remsen DP. "Enabling Biological Knowledge Integration Through Scientific Nomenclature." *International Society for Phylogenetic Nomenclature Conference*, Yale University, New Haven, CT. 2006.
7. **Sarkar IN**, DeSalle R. "Automated Barcoding Using the Characteristic Attribute Organization System." *CBoL DAWG*, National Museum of Natural History, Paris, France. 2006.
8. Rach J, DeSalle R, **Sarkar IN**, Schierwater B, Hadrys H. "Character Based DNA-Barcoding for Identifying Conservation Units in Odonata." *CBoL DAWG*, National Museum of Natural History, Paris, France. 2006.
9. Catapano T, Agosti D, Sautter G, Koning D, Boehm K, Johnson NF, Heidorn PB, Moritz TD, **Sarkar IN**, Stephenson C. "TaxonX: A Lightweight and Flexible XML Schema for Markup of Taxonomic Treatments." *TDWG*, St Louis, MO. 2006.
10. **Sarkar IN**, Planet PJ, DeSalle R. "TEPC: Total Evidence Phylogenetic Correlation of Microbial Phenotypes and Genotypes." *ISMB 2005: 13th Annual International Conference on Intelligent Systems for Molecular Biology*, Detroit, MI. 2005.
11. **Sarkar IN**, Cantor MN, Bodenreider O, Lussier YA. "GenesTrace: Biological Knowledge Discovery via Structured Terminology." *MedInfo 2004: 11th World Congress on Medical Informatics*, San Francisco, CA. 2004.
12. Cantor MN, **Sarkar IN**, Gelman R, Lussier YA. "Hybrid Lexical Methods for Mapping the Gene Ontology to the UMLS." *Pacific Symposium on Biocomputing*, Kaua'i, HI. 2003.
13. **Sarkar IN**, Cantor MN, DeSalle R, Lussier YA. "Exploring SNOMED Using Phylogenetic Analysis Tools." *AMIA Annual Meeting*, Washington, DC. 2002.
14. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. "Knowledge Acquisition of Organized Sets (KAOS) of Clinical Data." *AMIA Annual Meeting*, Washington, DC. 2001.
15. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. "Knowledge Aggregation and Separation Through a Novel Classification Algorithm." *AAAS Annual Meeting and Innovation Exposition*, San Francisco, California. 2001.
16. **Sarkar IN**, Malesewski JJ, Luckie DB. "Biocomputing in the Undergraduate Classroom." *AAAS Annual Meeting and Innovation Exposition*, San Francisco, California. 2001.
17. **Sarkar IN**, Luckie DB. "Biocomputing: Learning How to Write Programs to Solve Problems." *Michigan State University Undergraduate Research Forum*, East Lansing, Michigan. 1998.

18. **Sarkar IN**, Koslowsky DJ. “Mitochondrial Cleavage mRNA Cleavage/Polyadenylation Regulation in Pathogenic African Trypanosomes.” *Michigan State University 1997 Undergraduate Research Forum*, East Lansing, Michigan. 1997.

Book Chapter

1. Egan MG, Lee EK, Chiu JC, **Sarkar IN**, Coruzzi GM, DeSalle R. “Gene Orthology Assessment With OrthologID.” in *Bioinformatics for DNA Sequence Analysis*. Humana Press, Editor: David Posada. 2009.

Invited Presentations

1. University of Vermont. Department of Obstetrics and Gynecology. “An Overview of UVM CCTS Biomedical Informatics Activities.” 9 December 2010.
2. National Association of Health Data Organizations, All-Payer Claims Database Workshop, Salt Lake City, UT. “Piecing Together Health Care Reform: The Vermont Story.” 15 October 2010.
3. University of Vermont. Department of Medicine Fellows’ Research Seminar Series. “Translational Bioinformatics: Guiding Bench Side Innovations to the Clinic.” 23 July 2010.
4. University of Vermont. Inaugural Informatics Grand Rounds. “Biomedical Informatics Across the University and State of Vermont.” 29 June 2010.
5. University of Utah. “Biomedical Informatics Approaches for Exploring Biodiversity Knowledge.” 15 April 2010.
6. Third International Barcode of Life Conference. Plenary Address. “Informatics and Data Analytics for DNA Barcoding: Wherefrom, Where now, and Whereto?” Mexico City, MX. 11 November 2009.
7. Systems Biology of Aging. “Enabling a Systems View of the Biology of Aging through Semantic Technologies.” Phoenix, AZ. 6 December 2008.
8. Gerontological Society of America Annual Meeting. “Development of an Automated Framework to Identify Biology of Aging Literature.” National Harbor, MD. 22 November 2008.
9. European Molecular Biology Network. Keynote address. “Biodiversity Informatics: Enabling a Macroscopic View of Biology.” 18–20 September 2008.
10. Cambridge Semantic Web Group. “Harnessing the Power of the Semantic Web to Manage Biodiversity Data.” 10 June 2008.
11. Harvard Medical School. “Biodiversity Informatics: Information Integration Across the Spectrum of Life.” 25 February 2008.
12. Plant and Animal Genome XVI Conference. “Towards Developing Practical Taxonomic Ontologies.” 14 January 2008.
13. 5th Genome Standards Consortium (GSC) Workshop. “Gaz – an open source community-developed Gazetter.” Cambridge, UK. 12–14 December 2008.
14. Coalition for Networked Information, Fall 2007 Task Force Meeting. “Briefing: The Biodiversity Heritage Library.” Washington, DC. 10–11 December 2007.
15. Lyman Briggs College 40th Anniversary Celebration. “Biodiversity Informatics and Assembly of the Encyclopedia of Life.” East Lansing, MI. 4 October 2007.

16. National Marine Fisheries Service, National Oceanic and Atmospheric Administration (NOAA). "A Fish By Any Other Name: Organizing and Navigating Taxonomic Content From Literature." Woods Hole, MA. 4 April 2007.
17. MITRE Corporation. "Biomedical Insights Through the Lens of Biodiversity." Bedford, MA. 14 March 2007.
18. Marine Biological Laboratory Bay Paul Center. "Informatics Meets Biodiversity: Introductions." Woods Hole, MA. 9 March 2007.
19. Pacific Symposium on Biocomputing (PSB 07). "Invited Panelist: New Frontiers in Natural Language Processing." Maui, HI. 5 January 2007.
20. Columbia University Department of Biomedical Informatics, Department Retreat. Alumni Speaker. Hudson Valley, NY. 28–30 August 2006.
21. National Museum of Natural History, Paris, France. "Automated Barcoding Using the Characteristic Attribute Organization System." 7 July 2006.
22. Yale University, New Haven, CT: International Society for Phylogenetic Nomenclature (PhyloCode). "Enabling Biological Knowledge Integration Through Scientific Nomenclature." 29 June 2006.
23. University of Illinois Urbana–Champaign, Urbana–Champaign, IL. "Bridging Biomedical and Biodiversity Knowledge Through Literature." TeleNature Workshop on Refactoring Natural History Literature. 17–18 April 2006.
24. ACM Joint Conference on Digital Libraries (JCDL 2005); Networked Knowledge Organization Systems/Services (NKOS) Workshop. "Invited Panelist: Using the Existing Biomedical Infrastructure to Guide the Design and Development of Biological Ontologies." Denver, CO. 10 June 2005.
25. Quantitative Biology (QuBi) Symposium at Hunter College, City University of New York. "Contemporary Topics in Bioinformatics: Organizing and Curating the Genome Deluge." New York, NY. 6 May 2005.
26. Pacific Symposium on Biocomputing (PSB 05). "Invited Panelist: Enabling Resources for Biomedical Discovery." Big Island, HI. 6 January 2005.
27. National Library of Medicine. Training Directors' Meeting. "The Characteristic Attribute Organization System." Washington, DC. 9 July 2003.
28. Michigan State University – Lyman Briggs School. 2002 Fall Commencement. Alumni Speaker. East Lansing, MI. 7 December 2002.
29. Michigan State University – Lyman Briggs School. 35th Anniversary Alumni Guest Speaker. "Physiognomonics." East Lansing, MI. 26 September 2002.