Sonia Leach

National Jewish Health
Center for Genes, Environment, and Health
1400 Jackson St, A656
Denver C0 80206

leachs@njhealth.org
sonia.leach@gmail.com
compbio.ucdenver.edu/Hunter_lab/Leach

RESEARCH INTERESTS

Computational Biology, Next-Generation Sequencing, Genomic Data Integration, Probabilistic Graphical Models, Biological Interaction Networks, Gene Prioritization

PROGRAMMING EXPERIENCE

Languages: C/C++, Perl, Matlab, Splus/R, BioConductor, Mathematica

Tools: PhenoGen Informatics Website (phenogen.uchsc.edu), first design and implementation of the underlying gene expression data analysis pipeline
Hanalyzer analysis tool (hanalyzer.sourceforge.net), first design and implementation of the underlying network construction and analysis tools

EDUCATION

Brown University May 2006, Ph.D Computer Science
Thesis Title: Informed Structural Priors for Bayesian Networks: Applications in Molecular Biology Using Heterogeneous Data Sources
Advisors: Thomas Dean and Lawrence Hunter

Brown University May 1996, Sc.M. Computer Science
Thesis Title: Learning Dynamical Systems Using Hidden Markov Models
Advisors: Leslie Pack Kaelbling and Thomas Dean

Bucknell University May 1994, B.S. (magna cum laude) Computer Science and Engineering with Honors
Thesis Title: D-resolution: A Semantically Based Query Processing Procedure for Annotated Logic Programming
Advisor: James Lu

PROFESSIONAL EXPERIENCE

National Jewish Health, Center for Genes, Environment and Health   Denver CO
Assistant Professor, Director of Bioinformatics
Supervise and develop data analysis, interpretation, and visualization pipelines for next-generation sequencing data, and integration with other genetic, epigenetic, genomic, proteomic and metabolomic data. Oversee high-performance computing resources. Develop algorithms for relevant subgraph extraction, multi-loci prioritization and heterogeneous data integration.

Katholieke Universiteit Leuven Postdoctoral Fellow   Leuven Belgium
Developed tools for disease gene prioritization using interaction networks built by genomic data fusion (with Yves Moreau).

University of Colorado Health Sciences Center Postdoctoral Fellow   Aurora CO
Built graphical network-based models of gene interactions by integrating heterogeneous data sources, to aid high-throughput dataset analysis (with Larry Hunter).

University of Colorado Health Sciences Center Research Assistant   Denver CO
    Developed gene expression analysis pipeline for University-wide collaborations (with
    Larry Hunter).
National Cancer Institute Research Assistant   Bethesda MD
    Applied statistical methods to analyze gene expression data (with Larry Hunter).
National Library of Medicine Research Assistant   Bethesda MD
    Examined techniques for clustering gene expression data (with Larry Hunter).
NASA Ames Research Center Research Assistant   Mountain View CA
    Exploited qualitative information in decompositional model based learning (with Brian
    Williams).
Brown University, Research Assistant   Providence RI
    Helped develop algorithms for solving planning problems as Markov Decision
    Processes in domains with uncertainty and large state spaces (with Leslie Kaelbling and
    Tom Dean).
Bucknell University, Research Assistant   Lewisburg PA
    Created an interpreter for a multi-valued logic language (with James Lu).

TEACHING AND OUTREACH EXPERIENCE

University of Colorado Denver Computational Bioscience Program   Aurora CO
    Instructor NHLBI Genomics and Proteomics Hands-on Workshop
    Instructor UCD CPBS7711 and CPBS7712, lectures on Next-Generation Sequencing, Hidden
        Markov Models, Relevant Subgraph Extraction, Gene Prioritization
    PhD Thesis Committee Member
    Ronald Schuyler, PhD July 2012, thesis title 'Multi-component Genetic Associations.'
    David Knox, PhD student, project title 'Modeling Transcription As Regulation.'
    Thanh Le, PhD student, project title 'A Machine Learning Approach to Expression Analysis.'
    Student Rotation Supervisor
    Jaime Merlano, PhD student, gene expression regulatory models from clinical attributes.
Katholieke Universiteit Leuven   Leuven Belgium
    Student project supervisor
    Balazs Almassy, Erasmus student intern, expression data analysis in BioConductor.
University of Colorado Health Sciences Center   Denver and Aurora CO
    Student project supervisor
    Ramalakshmi Ramachandruin continuing education, expression analysis in Perl.
    Roman Brunecky Pharmacology PhD lab rotation, clustering in matlab.
    Ben Elias high school summer intern, gene expression analysis pipeline in perl/matlab.
    Shailaja Sheel Masters, expression network visualization Cytoscape plugin.
    Steven Glidewell Human Medical Genetics PhD lab rotation, time series in Bioconductor.
    Kemeni Tenku STEP student, expression analysis in matlab.
Brown University   Providence RI
    Teaching Assistant, Artificial Intelligence under Tom Dean.
        Prepared and presented lectures in professor's absence, held office hours.
    Course Assistant, Learning Dynamical Systems under Tom Dean.
        Contributed tutorials and code, and co-developed web-pages for course.
Bucknell University   Lewisburg PA
    Teaching Assistant, Programming Languages under James Lu.
        Supervised weekly labs and graded assignments.
NASA Space Grant   Providence RI
    K-12 Educational Outreach to inspire students in math and science space-related topics

through interactive presentations and discussions in schools.

Big Brother Big Sister   Denver CO
School-based Mentoring Program to build relationships and provide general tutoring.

GRANT EXPERIENCE
National Jewish Health

Co-Investigator - Metabolic Profiles of COPD Phenotypes	NIH P20 HL113445-01 (Bowler, PI)	4/01/12-3/31/17
Co-Investigator - Mechanisms promoting translocations and mature B cell lymphomas	NIH R01 CA166325-01 (Wang, PI)	4/01/12-3/31/17
Co-PI - Role of the Sf3a mRNA splicing complex in innate immunity regulation	Butcher Program Seed Grant (Alper, Co-PI)	6/1/12-5/31/14
Co-Investigator - Lung Genomics Research Consortium	NIH RC2 HL101715 (D. Schwartz, PI)	9/30/09-9/29/11

Brown University
Augmentation Awards for Science and Engineering Research Training (AASERT) Grant, Air Force Office of Scientific Research, applied and awarded self-funding 3-year grant entitled *Representing and Solving Air Campaign Planning Problems as Partially Observable Markov Decision Processes* under parent grant F30602-95-1-0020 *Time-Critical Planning and Scheduling for Aircraft Maintenance and Deployment.*
NASA Space Grant, applied and awarded 1-year individual fellowship requiring 25% time commitment to K-12 science education outreach activities in own research area.

University of Colorado Health Sciences Center
National Library of Medicine (NLM) R01 Grant Renewal, contributed methodology and use-case results on genomic data fusion network-based interpretation of expression data for grant R01-LM008111 *Technology Development for a MolBio Knowledge Base*

Katholieke Universiteit Leuven
Research Foundation Flanders (FWO), a major Belgian funding channel, formulated a PhD research project on design of algorithms for exploring, visualizing and extracting information in large sparse graphs with external node/edge annotations.
Concerted Research Program (GOA), a major Belgian funding channel, the FWO-project description was contributed as one work package for grant *MaNet: Mathematical engineering tools for Networks: Data driven mining, identification, control and optimization* which was awarded to the department.

PROFESSIONAL ACTIVITIES
Faculty Service:
Future Ideas Group member, National Jewish, 2011-2012
Comprehensive Exam Committee Chair, UCD CPBS Program 2012
Recruitment Committee, UCD CPBS 2011 and 2012
Program Retreat Committee Chair, UCD CPBS Program 2012
Reviewing:
Journals: BMC Bioinformatics, BMC Research Notes, BioTechniques, Journal of Biomedical Informatics, Nucleic Acids Research
Conferences: UAI, AAAI, ISMB, PSB, ECCB, AMIA Summit on Translational Bioinformatics
Program Committees:
ECCB 2010
Meeting Organization:

AAAI Spring Symposium on Learning Dynamical Systems, 1996. Helped with organizing program and logistics for symposium chair, Tom Dean.

European Conference on Computational Biology, 2010. Helped in venue selection and budget planning for conference organizer, Yves Moreau.

PUBLICATIONS

Yang IV, Luna LG, Cotter J, Talbert J, Leach SM, et al. (2012) The Peripheral Blood Transcriptome Identifies the Presence and Extent of Disease in Idiopathic Pulmonary Fibrosis. *PLoS ONE* 7(6): e37708.

Bin L, Kim BE, Hall CF, Leach SM, Leung DYM, "Inhibition of Transcription Factor Specificity Protein 1 Alters the Gene Expression of Keratinocytes Leading to Up-regulation of Kallikrein-related Peptidases and TSLP", *Journal of Investigative Dermatology*, 2011 Nov; 131(11):2213-2222.

Yu S, Tranchevent L, Leach SM, De Moor B, Moreau Y. A kernel based framework for cross-species candidate gene prioritization, *Machine Learning in Systems Biology*, 2011.

Feng W, Leach SM, Tipney H, Phang T, Geraci M, Spritz R, Hunter L, Williams T. Spatial and Temporal Analysis of Gene Expression During Growth and Fusion of the Mouse Facial Prominences. *PLoS ONE*. 2009; 4(12): e8066.

Leach SM, Tipney H, Feng W, Baumgartner WA, Kasliwal P, Schuyler RP, Williams T, Spritz RA, Hunter L. Biomedical Discovery Acceleration, with Applications to Craniofacial Development. *PLoS Comput Bio*. 2009; 5(3): e1000215.

Tipney HJ, Leach SM, Feng W, Spritz R, Williams T, Hunter L. Leveraging existing biological knowledge in the identification of candidate genes for facial dysmorphology. AMIA Summit on Translational Bioinformatics *BMC Bioinformatics*. 2009; 10(Suppl2): S12.

Karimpour-Fard A, Leach SM, Gill RT, Hunter LE. Predicting protein linkages in bacteria: which method is best depends on task. *BMC Bioinformatics*. 2008; Sep 24;9:397.

Karimpour-Fard A, Leach SM, Hunter LE, Gill RT. The topology of the bacterial co-conserved protein network and its implications for predicting protein function. *BMC Genomics*. 2008; Jun 30;9:313.

Gabow AP, Leach SM, Baumgartner WA, Hunter LE, Goldberg DS. Improving protein function prediction methods with integrated literature data. *BMC Bioinformatics*. 2008; Apr 15;9:198.

Bhave SV, Hornbaker C, Phang TL, Saba L, Lapadat R, Kechris K, Gaydos J, McGoldrick D, Dolbey A, Leach S, Soriano B, Ellington A, Ellington E, Jones K, Mangion J, Belknap JK, Williams RW, Hunter LE, Hoffman PL, Tabakoff B. The PhenoGen Informatics website: tools for analyses of complex traits. *BMC Genetics*. 2007; 8:59.

Leach SM, Gabow A, Hunter L, Goldberg DS. Assessing and Combining Reliability of Protein Interaction Sources. *Pacific Symposium of Biocomputing*. 2007; 12:433-444.

Lowes BD, Zolty R, Minobe WA, Robertson AD, Leach S, Hunter L, Bristow MR. Serial gene expression profiling in the intact human heart. *Journal of Heart and Lung Transplant*. 2006; 25(5):579-88.

Acquaah-Mensah GK, Leach SM, Guda C. Predicting the subcellular localization of human proteins using machine learning and exploratory data analysis. *Genomics Proteomics Bioinformatics*. 2006; 4(2):120-33.

Leach SM, Lu JJ, Murray NV, Rosenthal E. Mho-Resolution and Hybrid Knowledge Bases. *Journal of the Franklin Institute*. 2001; 338(5):583-600.

Hunter LE, Taylor RC, Leach SM, Simon R. GEST: A Gene Expression Search Tool Based on a Novel Bayesian Similarity Metric. *Bioinformatics*. 2001; 17:S115-S122.

Givan R, Leach SM, Dean T. Bounded Parameter Markov Decision Processes. *Artificial*

Intelligence. 2000; 122(1):71-109.

Leach SM, Lu JJ, Murray NV, Rosenthal E. Mho-Resolution: An Inference Rule for Regular Multiple-Valued Logics. *European Workshop on Logics in AI*. 1998.

Givan R, Leach S, Dean T. Bounded Parameter Markov Decision Processes. *European Conference on Planning*. 1997; 234-246.

Dean T, Givan R, Leach SM. Model Reduction Techniques for Computing Approximately Optimal Solutions for Markov Decision Processes. *Uncertainty in AI*. 1997; 124-131.

Leach SM, Lu JJ. Query Processing in Annotated Logic Programming: Theory and Implementation. *Journal of Intelligent Information Systems*. 1996; 6:33-58.

Leach SM, Lu JJ. Computing Annotated Logic Programs. *International Conference on Logic Programming*. 1994; 257-271.