BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors in the order listed on Form Page 2.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME	POSITION TITLE
Suzanna E. Lewis	Staff Scientist
eRA COMMONS USER NAME selewis	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
University of Michigan with highest honors	B.S.	1974	Microbiology
University of Michigan	M.S.	1978	Microbiology
University of Michigan	M.S.	1979	Computer Science

A. Positions and Honors:

1979-80 Database design and development, Systar Corporation
1980-83 Software project leader, Tak Automation
1983-86 Software project leader, Integrated Automation
1986-87 Software design, Hewlett-Packard
1987-91 Software project leader, Lawrence Berkeley National Laboratory
1991-92 Director of Informatics, Human Genome Center, Lawrence Berkeley National Laboratory
1992-2006 Head of Informatics, Berkeley Drosophila Genome Center, University of California, Berkeley
2006-present Staff Scientist, Lawrence Berkeley National Laboratory

Honors

2000 Co-winner, American Association for the Advancement of Science Newcomb Cleveland Prize for "The genome sequence of *D. melanogaster*".
2005 Fellow of the American Association for the Advancement of Science

Other Experience and Professional Memberships

1998-present Gene Ontology Consortium (co-founder, co-PI)
2000-2006 Organizer, Cold Spring Harbor Laboratory/Wellcome Trust conference on Genome Informatics
2001 Lecturer, Cold Spring Harbor Laboratory course on *Drosophila* genetics
2001 Lecturer, Cold Spring Harbor Laboratory course on Computational Biology
2001-present Generic Model Organism Database Project (co-founder)
2002-present Sequence Ontology (co-founder)
2002-present Instructor, Cold Spring Harbor Laboratory course on Bioinformatics
2002-2005 Reactome (co-PI)
2003 Lecturer, Jackson Laboratory course on Genome Sequence Analysis: Theory and Practice
2003-present Advisory Board Member for Rat Genome Database
2003-present Advisory Board Member for *Gramene* Database
2004-present Open Biomedical Ontologies (co-founder)
2005-present Advisory Board Member for RNA Ontology Consortium

B. Selected peer-reviewed publications (in chronological order).

Ashburner M, Misra S, Roote J, **Lewis SE**, Blazej R, Davis T, Doyle C, Galle R, George R, Harris N, Hartzell G, Harvey D, Hong L, Houston K, Hoskins R, Johnson G, Martin C, Moshrefi A, Palazzolo M, Reese MG, Spradling A, Tsang G, Wan K, Whitelaw K, Celniker S, et al. An exploration of the sequence of a 2.9-Mb region of the genome of Drosophila melanogaster: the Adh region. Genetics 1999;153:179-219.

Rubin GM, Yandell MD, Wortman JR, Gabor Miklos GL, Nelson CR, Hariharan IK, Fortini ME, Li PW, Apweiler R, Fleischmann W, Cherry JM, Henikoff S, Skupski MP, Misra S, Ashburner M, Birney E, Boguski MS, Brody T, Brokstein P, Celniker SE, Chervitz SA, Coates D, Cravchik A, Gabrielian A, Galle RF, Gelbart WM, George RA, Goldstein LS, Gong F, Guan P, Harris NL, Hay BA, Hoskins RA, Li J, Li Z, Hynes RO, Jones SJ, Kuehl PM, Lemaitre B, Littleton JT, Morrison DK, Mungall C, O'Farrell PH, Pickeral OK, Shue C, Vosshall LB, Zhang J, Zhao Q, Zheng XH, **Lewis S**. Comparative genomics of the eukaryotes. Science 2000;287:2204-15.

Reese MG, Hartzell G, Harris NL, Ohler U, Abril JF, **Lewis SE**. Genome annotation assessment in Drosophila melanogaster. Genome Res 2000;10:483-501.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, **Lewis S**, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25-9.

Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva A, **Lewis S**. The generic genome browser: a building block for a model organism system database. Genome Res 2002;12:1599-610.

Mungall CJ, Misra S, Berman BP, Carlson J, Frise E, Harris N, Marshall B, Shu S, Kaminker JS, Prochnik SE, Smith CD, Smith E, Tupy JL, Wiel C, Rubin GM, **Lewis SE**. An integrated computational pipeline and database to support whole-genome sequence annotation. Genome Biol 2002;3:RESEARCH0081.

Lewis SE, Searle SM, Harris N, Gibson M, Lyer V, Richter J, Wiel C, Bayraktaroglu L, Birney E, Crosby MA, Kaminker JS, Matthews BB, Prochnik SE, Smith CD, Tupy JL, Rubin GM, Misra S, Mungall CJ, Clamp ME. Apollo: a sequence annotation editor. Genome Biol 2002;3:RESEARCH0082.

Misra S, Crosby MA, Mungall CJ, Matthews BB, Campbell KS, Hradecky P, Huang Y, Kaminker JS, Millburn GH, Prochnik SE, Smith CD, Tupy JL, Whitfied EJ, Bayraktaroglu L, Berman BP, Bettencourt BR, Celniker SE, de Grey AD, Drysdale RA, Harris NL, Richter J, Russo S, Schroeder AJ, Shu SQ, Stapleton M, Yamada C, Ashburner M, Gelbart WM, Rubin GM, **Lewis SE**. Annotation of the Drosophila melanogaster euchromatic genome: a systematic review. Genome Biol 2002;3:RESEARCH0083.

Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, **Lewis SE**, Rubin GM, Ashburner M, Celniker SE. The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. Genome Biol 2002;3:RESEARCH0084.

Tomancak P, Beaton A, Weiszmann R, Kwan E, Shu S, **Lewis SE**, Richards S, Ashburner M, Hartenstein V, Celniker SE, Rubin GM. Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol 2002;3:RESEARCH0088.

Mi H, Vandergriff J, Campbell M, Narechania A, Majoros W, **Lewis S**, Thomas PD, Ashburner M. Assessment of genome-wide protein function classification for Drosophila melanogaster. Genome Res 2003;13:2118-28.

Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, **Lewis S**, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la Cruz N, Tonellato P, Jaiswal P, Seigfried T, White R; Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 2004;32(Database issue):D258-61.

Lewis SE. Gene Ontology: looking backwards and forwards. *Genome Biology* 2004, 6:103

Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath GR, Wu GR, Matthews L, **Lewis S**, Birney E, Stein L. Reactome: a knowledgebase of biological pathways. Nucleic Acids Res 2005;33(Database issue):D428-32.

Drysdale RA, Crosby MA, Gelbart W, Campbell K, Emmert D, Matthews B, Russo S, Schroeder A, Smutniak F, Zhang P, Zhou P, Zytkovicz M, Ashburner M, de Grey A, Foulger R, Millburn G, Sutherland D, Yamada C, Kaufman T, Matthews K, DeAngelo A, Cook RK, Gilbert D, Goodman J, Grumbling G, Sheth H, Strelets V, Rubin G, Gibson M, Harris N, **Lewis S**, Misra S, Shu SQ; FlyBase Consortium. FlyBase: genes and gene models. Nucleic Acids Res 2005;33(Database issue):D390-5.

Eilbeck K, **Lewis SE**, Mungall CJ, Yandell M, Stein L, Durbin R, Ashburner M. The Sequence Ontology: a tool for the unification of genome annotations. Genome Biol 2005;6:R44.

Leontis NB, Altman R, Berman HM, Brenner SE, Brown J, Engelke J, Harvey SC, Holbrook SR, Jossinet F, **Lewis SE**, Major F, Mathews DH, Richardson J, Williamson JR, and Westhof E. The RNA Ontology Consortium: An Open Invitation to the RNA Community. RNA 2006 12:533-541.

Rubin DL, **Lewis SE**, Mungall CJ, Misra S, Westerfield M, Ashburner M, Sim I, Chute CG, Solbrig H, Storey MA, Smith B, Day-Richter J, Noy NF, Musen MA. National Center for Biomedical Ontology: advancing biomedicine through structured organization of scientific knowledge. OMICS 2006 Summer;10(2):185-98.

Yandell MD, Mungall CJ, Prochnik S, Smith C, Kaminker J, Hartzell G, **Lewis S**, Rubin GM. Comparative analyses of gene-structure in 11 animal genomes. PLoS Comput Biol. 2006 3;2(3):e15 [Epub]

Swedlow JR, **Lewis SE**, Goldberg IG. Modeling data across labs, genomes, space and time. Nat Cell Biol. 2006 Nov;8(11):1190-4.

Thomas PD, Mi H, **Lewis S**. Ontology annotation: mapping genomic regions to biological function. Curr Opin Chem Biol. 2007 Feb;11(1):4-11.

Vastrik I, D'Eustachio P, Schmidt E, Joshi-Tope G, Gopinath G, Croft D, de Bono B, Gillespie M, Jassal B, **Lewis S**, Matthews L, Wu G, Birney E, Stein L. Reactome: a knowledgebase of biological pathways and processes. Genome Biol. 2007 Mar 16;8(3):R39

Day-Richter J, Harris MA, Haendel M; Gene Ontology OBO-Edit Working Group, **Lewis S.** OBO-Edit--an ontology editor for biologists. *Bioinformatics*. 2007 Aug 15;23(16):2198-200.

Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K, Ireland A, Mungall CJ; OBI Consortium, Leontis N, Rocca-Serra P, Ruttenberg A, Sansone SA, Scheuermann RH, Shah N, Whetzel PL, **Lewis S.** The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol.* 2007 Nov;25(11):1251-5.

C. Research Support

Ongoing Research Support

P41 HG002273 (Blake) 3/1/07-2/28/12
Jackson Laboratory, NHGRI (prime)
Gene Ontology Consortium

The major goal of this project is to provide and implement shared, structured vocabularies for the biological description of gene products in any organism.
Role: Co-Investigator

U54 HG004028 (Musen) 9/23/05-7/31/10
NIH/Roadmap
National Center for Biomedical Ontology

The major goal of this project is to develop tools and methods for assimilating, archiving, and accessing machine-processable representations of biomedical domain objects, processes, and relations to assist in the management, integration, visualization, analysis, and interpretation of the huge, distributed data sets that are now the hallmark of biomedical research.
Role: Co-Investigator

U41 HG004269 (Stein) 03/01/07-02/28/08
Cold Spring Harbor Laboratory, NHGRI (prime)
A data coordination center for the model organism ENCODE project

The ModENCODE Data Coordination Center will compile and track data being generated by different modENCODE research projects and will implement quality assurance methods to ensure that data meet the relevant data standards.
Role: Co-Investigator

OVERLAP: None

R01 NS058296 (Martone) 04/01/07-03/31/08
UCSD, NINDS (prime)
In Process pending activation of subcontract
Ontology-Based Multiscale Investigation of Neurological Disease

The major goal of this proposal is to develop tools that bridge between information contained in ontologies, distributed relational databases and images.
Role: Co-Investigator
OVERLAP: None

1R01GM083649-01A1 (Hunter, L.) 09/30/2007-08/31/2011
University of Colorado, NLM
Ontologies and Biomedical Language Processing

The goal of this project is to explore both the potential for, and obstacles to, the mutual application of biomedical ontologies and biomedical language processing.
Role: Subcontract PI
Overlap: None.

Completed Research Support

R01 HG002273-04 (Blake) 1/1/04-2/28/07
NIH/NHGRI
Gene Ontology Consortium
The major goal of this project is to provide and implement shared, structured vocabularies for the biological description of gene products in any organism.
Role: Co-Investigator

2P41 HG00739 (Gelbart) 12/01/98-11/30/05
NIH/NHGRI
FlyBase: A Drosophila Genomic and Genetic Database
This grant supports development of the FlyBase database, which contains the core genetic and genomic information for the fruitfly *Drosophila melanogaster* and related organisms of the family Drosophilidae.

5P50 HG00750 (Rubin) 12/24/98-11/30/04
NIH/NHGRI
Drosophila Genome Center
This grant supports a Drosophila Genome Center coordinating a scientific collaboration among the University of California, Berkeley, Lawrence Berkeley National Laboratory, Baylor College of Medicine and the Carnegie Institute of Washington at Baltimore. The goals of this project are to sequence and biologically annotate the *Drosophila* genome.