BIOGRAPHICAL SKETCH

NAME: **Arend Sidow, PhD**

eRA COMMONS USER NAME: SIDOW.AREND

POSITION TITLE: **Associate Professor of Pathology and of Genetics**

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE	Completion Date	FIELD OF STUDY
Universität Göttingen, Germany	Vordiplom	05/1987	Biology
University of California, Berkeley, CA	PhD	11/1993	Genetics (Molecular Evolution)
Whitehead Institute, Cambridge, MA	Postdoc	08/1998	Mouse Genetics/Genomics

A. PERSONAL STATEMENT

My research program resides at the interface of computational and functional genomics, with a focus on the interactions between function, variation, and evolution. With a background in molecular evolution, model organism genetics, and functional genomics, I have shaped my group's interests to be broad, not focused on one particular field, and strive to apply concepts from one field to questions in another. As detailed in section C, we have published on a diverse range of topics in genomics and generated both significant amounts of data as well as analytic advances. I emphasize rigor both in the lab and analytically, and much of our research is highly collaborative. For the proposed project, my track record in innovation at the interface of genomics and evolution, most recently in cancer genome evolution, is evidence that the aims of the grant will be met and perhaps exceeded.

B. POSITIONS AND HONORS

Positions and Employment

1998-2005 Assistant Professor, Dept of Pathology and Dept of Genetics, Stanford University
2005 - Associate Professor, Dept of Pathology and Dept of Genetics, Stanford University

Other Experience (recent)

2014 - Member, Executive Committee of Joint Initiative for Metrology in Biology (NIST)
2010 - Director, Genome Sequencing, Department of Pathology
2010 - Director, Service Center of the Center for Genomics and Personalized Medicine
2010 - Member, Executive Committee of Stanford Center for Genomics and Personalized Medicine
2008 - Co-founder, DNAnexus
2007 - Member, Executive Committee of Stanford BioX
2000 - Co-director, (Director and PI 2008-11), Stanford Genome Training Program, NIH/NHGRI
1999-2011 Every Fall Quarter: Advanced Genetics (Gen203); Graduate course, enrollment ~80
2003-2010 NHGRI Sequencing Priorities Coordinating Committee
2003-2010 Joint Genome Institute: Sequencing Priorities Panel

Honors

2006 Stanford School of Medicine Graduate Teaching Award
2005 American Association for the Advancement of Science (AAAS) Fellow
2002-2003 Hume Faculty Scholar, Stanford University School of Medicine
1999-2002 Stanford University Frederick E. Terman Fellowship
1999-2000 Howard Hughes Medical Institute Junior Faculty Award
1994-1997 Helen Hay Whitney Postdoctoral Fellowship
1987-1988 Education Abroad Program Fellow, Göttingen – UC Berkeley

C. Contributions to Science

1. Molecular evolution of vertebrates. As a graduate student in the laboratory of the late Allan Wilson (of chimp-human and mitochondrial eve fame), I worked on the molecular evolution of the WNT gene family and was among the first to provide empirical evidence for the early vertebrate genome duplications. The work, published in the early 90's, contributed to the convergence of molecular evolution and developmental biology and to the beginnings of the evo-devo movement.

- Sidow A. (1992) Diversification of the Wnt gene family on the ancestral lineage of vertebrates. PNAS 89:5098-5102.
- Holland PW, Garcia-Fernandez J, Williams NA, Sidow A. (1994) Gene duplications and the origins of vertebrate development. Development Supplement 1994, 125-133.

2. Sequence first. Ask questions later. My postdoctoral work coincided with the implementation of the human genome project, one of the epicenters of which was the Whitehead Institute. It exposed me to the power of genomics, which was in its infancy then. When I established my lab at Stanford in 1999 I decided to base much of my research program on combining an evolution of my graduate work with genomics. We were among the first to recognize and apply the concept of evolutionary constraint to genome sequence comparisons. The work had widespread implications for our understanding of mammalian genome evolution and significantly contributed to the plan by the NIH to fund the sequencing of many mammalian genomes (with Sanger technology!) for the express purpose of annotating the human genome with evolutionarily constrained regions, the majority of which are noncoding. The case for this was laid out in a Cell minireview in 2002 and fortified in a Genome Research paper in 2003 in which we demonstrate the power of evolutionary tree-based rate estimates for the annotation of mammalian genomes.

- Sidow A. (2002) Sequence first. Ask questions later. Cell 111:13-16.
- Cooper GM, Brudno M, Green ED, Batzoglou S, Sidow A. (2003) Quantitative estimates of sequence divergence for comparative analyses of mammalian genomes. Genome Research 13:813-820.
- Stone EA, Cooper GM, Sidow A. (2005) Trade-offs in detecting evolutionarily constrained sequence by comparative genomics. Annual Review of Genomics and Human Genetics 6:143-164.

3. ENCODE pilot and evolutionary constraint in the mammalian genome. The value of comparative sequence analysis was harnessed on a large scale for the first time in the human ENCODE pilot project, which focused on 1% of the human genome. We showed definitively that a much larger fraction of mammalian genomes was constrained than could be explained by coding sequence. Later work addressed the correlation of evolutionary constraint with human intraspecific variation. A method we developed, GERP, has been used in many publications to identify evolutionarily constrained noncoding sequence or predict the impact of mutations of unknown significance. We also developed complementary methodology for mutation impact analyses of protein sequences.

- Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A. (2005) Distribution and intensity of constraint in mammalian genomic sequence. Genome Research 15:901-913
- Encode Project Consortium. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799-816.
- Marguiles EH, Cooper GM, ...Encode Comparative Analysis Group ... Sidow A. (2007) Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. Genome Research 17:760-774.
- Goode DL, Cooper GM, Schmutz J, Dickson M, Gonzales E, Tsai M, Karra K, Davydov E, Batzoglou S, Myers RM, Sidow A. (2010) Evolutionary constraint facilitates interpretation of genetic variation in resequenced human genomes. Genome Research 20:301-310.

4. Gene regulatory fundamentals of the human genome. In the mid 2000's, Agencourt Personal Genomics and Solexa invented competing approaches for massively parallel DNA sequencing. We were the first to adopt SOLiD technology (by then acquired by ABI) and also began to use Solexa sequencing (later acquired by Illumina). The field immediately put these technologies to use for functional genomic approaches and developed ChIP-seq, RNA-seq, and MNase-seq. We published the first ChIP-seq peak finder, QUEST, and contributed substantially to the analysis of ENCODE II data to reveal insights into the structure and function of regulatory elements, and to our understanding of nucleosome positioning in the human genome.

- Valouev A, Johnson DS, Sundquist A, Medina C, Anton E, Batzoglou S, Myers RM, Sidow A. (2008) Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. Nature Methods 5:829-835.
- Valouev A, Johnson SM, Boyd S, Smith C, Fire AZ, Sidow A. (2011) Determinants of nucleosome organization in primary human cells. Nature 474:516-520.
- Kundaje A, Kyriazopoulou-Panagiotopoulou S, Libbrecht M, Smith CL, Raha D, Winters EE, Johnson SM, Snyder M, Batzoglou S, Sidow A. (2012) Ubiquitous heterogeneity and asymmetry of the chromatin environment at regulatory elements. Genome Research 22:1735-1747.
- ENCODE Project Consortium. (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489:57-74.

5. Cancer genome evolution. All cells in our body are related by a bifurcating lineage tree whose root is the zygote. This concept permeates our work in cancer evolution, which we began in about 2009 when it became feasible to sequence whole human genomes with Illumina technology. We sequence multiple neoplastic or cancer samples per individual and then, using the same concepts of molecular evolution that I applied in graduate school and later in comparative genome analysis, we address fundamental questions in cancer evolution: How nonmalignant concurrent growths are related to the carcinoma, what genomic changes they harbor if any, the order of driver events in the evolution of the cancer, and more. Some results are provocative. For example, in luminal breast cancer, the earliest drivers are aneuploidies and not mutations in oncogenes or tumor suppressors; and remarkably, some driver mutations occur independently multiple times in the same patient, sometimes in neoplasias and not always in the carcinoma. It is early days in cancer genomics and cancer evolution in particular, but I am confident that our contributions in concepts, methodology, and insights, will be recognized as innovative and perhaps as pioneering as the field matures.

- Newburger DE, Kashef-Haghighi D, Weng Z, Salari R, Sweeney RT, Brunner AL, Zhu SX, Guo X, Varma S, Troxell ML, West RB, Batzoglou S, Sidow A. (2013) Genome evolution during progression to breast cancer. Genome Research 23:1097-1108.
- Valouev A, Weng Z, Sweeney RT, Varma S, Le QT, Kong C, Sidow A, West RB. (2014) Discovery of recurrent structural variants in nasopharyngeal carcinoma. Genome Research 24:300-309.
- Weng Z, Spies N, Zhu SX, Newburger DE, Kashef-Haghighi D, Batzoglou S, Sidow A, West RB. (2015). Cell-lineage heterogeneity and driver mutation recurrence in pre-invasive breast neoplasia. Genome Medicine, doi: 10.1186/s13073-015-0146-2.
- Sidow A, Spies N. (2015) Concepts in solid tumor evolution. TIGS, doi: 10.1016/j.tig.2015.02.001.

Much of our work has been collaborative, especially with Serafim Batzoglou's group in the Computer Science department. The breast cancer work is a collaboration with Batzoglou and pathologist Robert West. A full list of my papers can be found by searching for Sidow A [au] in entrez. A selective list is at http://mendel.stanford.edu/SidowLab.

D. Research Support

Ongoing Research Support

R01CA183904 (Multi-PI: Sidow [contact], West, Batzoglou) 03/01/15 - 02/28/19
NIH/NCI
Genomic Evolution of Breast Cancer
The goal of this project is to understand the molecular evolution and epigenetic transitions of HER2-positive breast cancer and associated neoplastic lesions.

Gift Ford, Sidow, Diehn (Multi-PI) 07/01/15 - 06/30/17
BRCA foundation
Understanding BRCA1 and BRCA2-associated cancers
The goal of this project is comprehensive mutation detection, molecular profiling, and evolutionary analysis of BRCA-positive breast and ovarian cancers.

1P01AG036695-01A1 (Rando PI; Sidow Co-I) 07/01/11-06/30/16
NIH/NIA
Molecular Regulation of Stem Cell Aging.
Core C: Genomics and Ultra High throughput Sequencing Core

Selected Completed Research Support

SPO 52520 (West, PI; Sidow, Co-I) 08/01/11-07/31/13
University of California Office of the President
Molecular classification of early breast neoplasia.
The goal of this project is to find the earliest genomic events that predispose proliferating breast tissue to develop a carcinoma

AEA Stanford-KAUST (Batzoglou, PI; Sidow, Co-PI) 01/01/12 – 12/31/13
KAUST
Comprehensive prediction of tissue-specific regulatory elements for human by using integrative analysis of high throughput functional assays. The goal of this project is to identifiy tissue-specific regulatory elements on the basis of ENCODE data

U54 HG004695 subaward (Birney PI; Sidow subaward PI) 07/01/10 - 03/31/12
NIH/NHGRI
EDAC: ENCODE Data Analysis Center
The goal of this project is to perform systems-level analyses on the ENCODE functional genomics data

U54 HG004576 (Myers, PI; Sidow, Co-PI) 09/27/07 – 06/30/10
NIH/NHGRI
Global Annotation of Regulatory Elements in the Human Genome
The major goal of this project is to functionally characterize genes and regulatory regions of the human genome.