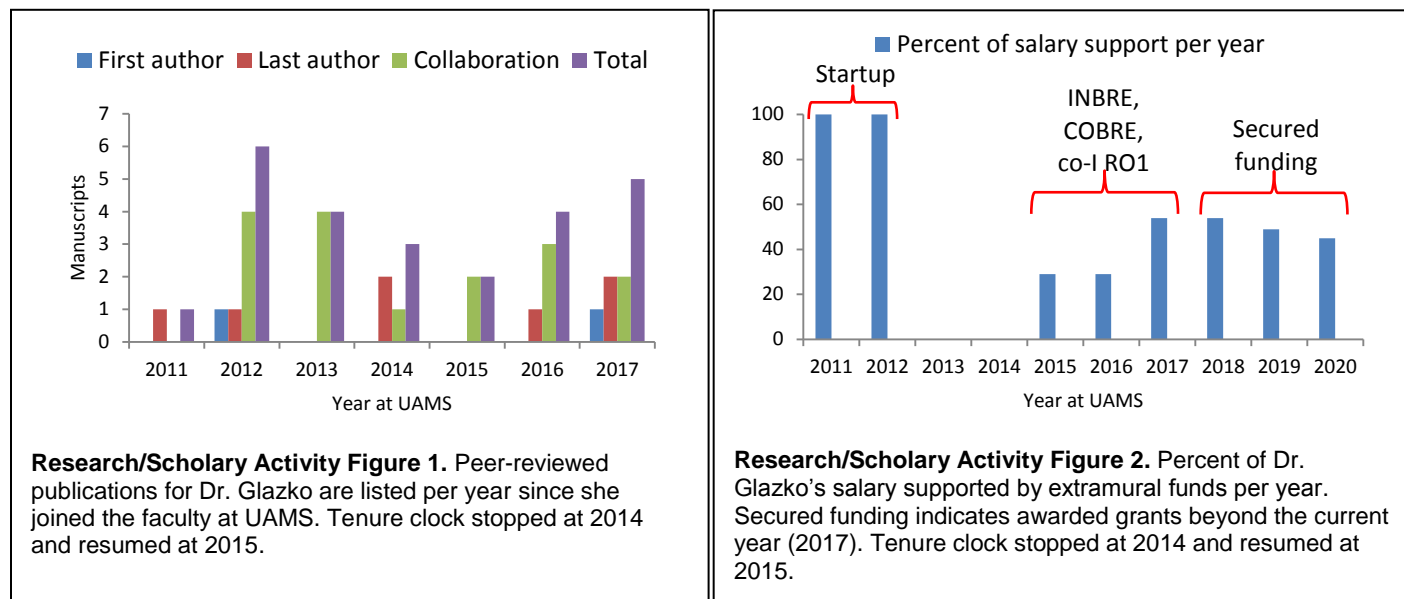


RESEARCH/SCHOLARLY ACTIVITIES

A. Overview

Detailed below are Research/Scholarly Activities for Dr. Glazko that account for **80%** of her time/effort as a faculty member in the Basic Scientist Tenure Track at UAMS. The highlights of this section include:



- Published **71** papers that were cited more than **3600** times: **47** prior to joining the faculty at UAMS in 2011, **24** as Assistant Professor at UAMS (Research/Scholarly Activities **Figure 1**)
- **Seven** senior author publications since 2011 at UAMS (Research/Scholarly Activities **Figure 1**)
- Increase over the years of Dr. Glazko's efforts covered by NIH-funding: from **29%** in 2015 toward **54%** in 2017 (Research/Scholarly Activities **Figure 2**)
- Produced one Bioconductor software package "Gene Set Analysis in R" (GSAR) which has been downloaded by other investigators more than **6000** times
- Serves currently as Editorial Board Member and Associate Editor for *BMC Genomics* and *Frontiers in Genetics*, the section of *Bioinformatics and Computational Biology*
- 32 hours/week for Research/Scholarly Activities (40 hr/wk * 80% time/effort = 32 hr/wk)
 - 10 hr/wk grant-related work
 - 18 hr/wk research, code development, data analysis, publications
 - 4 hr/wk presentations, editorial work

B. Research Interests

Dr. Glazko's area of expertise is Bioinformatics, an interdisciplinary field that develops methods and software tools for understanding large-scale high-throughput omics (transcriptomics, proteomics, metagenomics, to name just a few) data sets. These data sets are most frequently collected with the aim of elucidating the molecular basis of various human diseases as well as answering fundamental biological questions. Bioinformatics implements data preprocessing and data analysis for these studies and, frequently, new Bioinformatics methods generate new clinical and biological knowledge. Hence, Bioinformatics approaches and pipelines serve as a cornerstone for all omics-oriented clinical and basic science research projects. Bioinformatics is similar to Biostatistics in that it is highly collaborative, i.e., the efforts in the field of Bioinformatics are equally distributed between methods development for omics data preprocessing and analysis and data analysis per se. As an interdisciplinary field of science, Bioinformatics combines computer science, statistics, mathematics and biology to analyze and interpret omics data.

Dr. Glazko develops and applies statistical and computational methods to analyze high dimensional omics data to gain a better understanding of the molecular causes of diseases and to discover new biological phenomena. Her major methodological interests are in developing new and comprehensive approaches for pathway analysis, network inference, data integration and statistical methods for single-cell omics analysis.

Dr. Glazko major collaborative projects include 1) developing new platform-independent molecular classifiers to distinguish polyps with and without colorectal cancer risk with Dr. Curt H. Hagedorn (Professor of Internal Medicine/Gastroenterology, UAMS), 2) building a new pipeline for metaproteomics data analysis in chronic kidney disease with Dr. John M. Arthur (Professor of Internal Medicine/Nephrology, UAMS) and 3) designing new statistical tool for the analysis of mass spectrometry count data in proteomics with Dr. Boris L. Zybailov (Assistant Professor of Biochemistry and Molecular Biology).

Additionally, she directs the Arkansas INBRE (<http://brin.uams.edu/>) Bioinformatics Research Support Core at UAMS (<http://brin.uams.edu/bioinformatics.asp>). In the last two years the INBRE Bioinformatics Research Support Core has implemented several Bioinformatics pipelines, in particular for RNA-seq, epigenomic and 16S rRNA sequencing (metagenomics) data analysis and provided Bioinformatics Research Support for more than ten INBRE-supported investigators since its foundation in 2015.

Dr. Glazko's methodological and collaborative projects are state-of-the-art in the field of Bioinformatics (her publications have been cited more than 3600 times, Google Scholar, retrieved August 1st) and in high demand, as confirmed by high popularity of the software tool she developed, Gene Set Analysis in R (GSAR), that was downloaded more than 6,000 times (see 'Contributed software tools' Section for more detail). This explains her unique role in collaboration with UAMS clinical and basic scientists, and is reflected in her service as the UAMS INBRE Bioinformatics Research Support Core Director, Co-Director of Arkansas INBRE Bioinformatics Core and Co-Director of Systems Biology and Bioinformatics Core for COBRE Center for Translational Pediatric Research.

C. Current Grants (Dr. Glazko uses 10 hrs/wk of her time to serve as a Co-Investigator on extramurally/intramurally funded grants and to write new grants and progress reports)

1. NIH/NIGMS P20GM121293, P.I., Dr. Alan Tackett, UAMS; Role: Co-Director, Systems Biology and Bioinformatics Core, 07/11/2017 to 06/30/2022. Title: Center for Translational Pediatric Research. (20% salary support for Dr. Glazko FY2017-2022; \$1,500,000 annual direct grant cost). Funded July 11, 2017
2. NIH/NIGMS P20GM103429; P.I., Dr. Lawrence E. Cornett, UAMS; Role: Director, Bioinformatics Research Support Core, Co-Director, Bioinformatics Core, 05/1/2015 to 04/30/2020. Title: Partnerships for Biomedical Research in Arkansas. (25% salary support for Dr. Glazko FY2015-2017; \$3,217,687 annual direct grant cost)
3. NIH GM081766-07; P.I., Dr. Wayne P. Wahls, UAMS; Role: Co-Investigator, 08/10/2007 to 5/31/2019. Title: Biochemistry of Recombination in Meiosis. (4% salary support for Dr. Glazko FY2015-2017; \$294,612 annual direct grant cost)
4. NIH/NIGMS P20GM103429-15S1; P.I., Dr. Lawrence E. Cornett, UAMS; Role: Co-Investigator, 08/15/2016 to 08/14/2017. Title: Partnerships for Biomedical Research in Arkansas Supplement, to

create the National Resource for Proteomics (5% salary support for Dr. Glazko FY2017; \$306,604 annual direct grant cost)

D. Previous Grants

1. 1 R21 CA155832; NIH/Funded by Clinical and Translational Science Institute CTSI, University of Rochester Medical Center. P.I., Dr. Galina Glazko, 07/01/2010 – 06/30/2012. Title: Gene Expression Profiles in Tumors: Local and Global Expression Differences (not transferred to UAMS)
2. 5 R01 DC009439; P.I., Dr. Loisa Bennetto, University of Rochester Medical Center. Role: Biostatistician, 02/18/09 – 01/31/14. Title: Taste, Smell, and Feeding Behavior in Autism: A Quantitative Traits Study
3. 2 R01 HL062826-09A2; P.I., Dr. Bradford C. Berk, University of Rochester Medical Center; Role: Biostatistician, 08/01/10 – 04/30/14. Title: Genetics of Vascular Remodeling
4. 5 R01 AA016178; P.I., Dr. Jim Fry, University of Rochester Medical Center; Role: Biostatistician, 04/02/2007 – 03/31/2011. Title: Genetic Basis of Clinical Variation in Ethanol Tolerance in *Drosophila melanogaster*
5. 5 R00 LM009477; P.I., Dr. Peter A. Salzman, University of Rochester Medical Center; Role: Biostatistician, 01/01/2009 – 08/17/2011. Title: Statistical & Computational Tools for Reconstruction of Gene Regulatory Networks
6. 5 R21 GM079259; P.I., Dr. Xing Qui, University of Rochester Medical Center; Role: co-investigator, 07/01/2007 – 06/30/2010. Title: Correlation Vectors in Gene Expression Profiling
7. **Sloan Research Fellowship** (<https://sloan.org/fellowships> awarded yearly to 126 early-career, tenure-track faculty in the U.S. or Canada, in recognition of distinguished performance and outstanding promise in one of eight specific fields); P.I., Dr. Galina Glazko, University of Rochester Medical Center; 07/2007-07/2009, \$130,000

E. Collaborative Grant Effort

Dr. Glazko has been actively involved as a Co-Investigator on grants listed in Sections C and D above. The sum of direct costs for the grants listing Dr. Glazko as a co-investigator above (Drs. Cornett, Wahls and Tackett) is **~\$5,318,903** annual to UAMS.

F. Publications (18 hrs/wk of Dr. Glazko's time are used for publication of data)

Please refer to the "Publications Section" of this packet for details. In brief, Dr. Glazko has published **71** peer-reviewed manuscripts in her field of study. These manuscripts were cited more than **3600** times. Of these, Dr. Glazko published **47** prior to her faculty appointment at UAMS in 2011 and **24** as Assistant Professor at UAMS (**7** out of **24** are senior author publications).

G. Presentations/Seminars

1. "Pathway Analysis Online", UAMS, IDeA National Resource for Proteomics Workshop, invited presentation (May 25 2017)
2. "Pathway Analysis in Practice", UAMS, IDeA National Resource for Proteomics Workshop, invited presentation (May 16 2017)
3. "Pathway Analysis: Theory and Practice", UAMS, IDeA National Resource for Proteomics Workshop, invited presentation (April 2017)
4. "Bioinformatics For Genomics and Proteomics", UAMS, The Biomedical Informatics Research & Application Seminar (September 2016)
5. "Statistical Modelling For Biological Systems", Invited presentation for Bioinformatics graduate students, UALR (October 2014)

6. “Pathway Analysis of Expression Data”, Invited presentation for Bioinformatics graduate students, UALR (February 2012)
7. “Pathway Analysis of Expression Data: Deciphering Functional Building Blocks of Complex Diseases” International Society for Computational Biology meeting, Boston, Invited tutorial presentation (June 2010)
8. “Gene Vectors: Analysis Strategies and Emerging Phenomena”, Invited seminar, University of Arkansas for Medical Sciences, Division of Biomedical Informatics (August 2010)
9. “Similarity Searches in Genome-Wide Numerical Data Sets”, Invited seminar, University of Rochester Medical Center, Department of Biostatistics and Computational Biology (October 2005)
10. “The Use of Generalized Expression Profiles for Prediction of Proteins Related to Merozoite Invasion”, CAMDA (Critical Assessment of Microarray Data Analysis) conference, Hilton Hotel, Durham North Carolina (November 2004)

H. Symposiums

1. IDeA National Resource for Proteomics Workshop, UAMS, May 25, 2017, workshop Instructor
2. IDeA National Resource for Proteomics Workshop, UAMS, May 16, 2017, workshop Instructor
3. IDeA National Resource for Proteomics Workshop, UAMS, April, 2017, workshop Instructor
4. The Midsouth Computational Biology and Bioinformatics Society (MCBIOS) conference, Clinton Presidential Library, Little Rock AR, March 2015, conference Co-organizer

I. Editorial Work

1. Editorial Board member, Associate Editor, *Frontiers in Genetics*, section *Bioinformatics and Computational Biology* (2012-present)
2. Editorial board member, Associate Editor, *BMC Genomics* (2014-present)

J. NIH Grant Reviewer

1. Stage 1 reviewer for grant applications submitted in response to RFA-10-005 (Recovery Act Limited Competition: Director’s Opportunity for Research in Five Thematic Areas), 05/24/2010-05/25/2010

K. Non-NIH Grant Reviewer

1. Bioinformatics pilot projects for the Arkansas INBRE, one proposal, 02/15/2017
2. NSF Electronic Proposal Review, Proposal No. 1616878, 01/06/2016
3. Bioinformatics pilot projects for the South Carolina INBRE, three proposals, 6/2/2016

L. Contributed Software Tools

1. **CLUSTERM**, Mutational Spectrum Analysis software
What it does? Given a mutational spectrum as an input CLUSTERM software finds mutational hotspots using mixture of binomial distributions model.
Reference paper: **Glazko G.V.**, Milanese L., Rogozin I.B. (1998). The subclass approach for mutational spectrum analysis: application of the SEM algorithm. *J Theor Biol*, **192**:475-487.
2. **CHRCLASS**, Scaffold/Matrix Attachment regions (S/MARs) Prediction Tool
What it does? Given a DNA region as an input CHRCLASS software finds matrix/scaffold attachment regions and for each region estimates its probability of being S/MARs.
Reference paper: **Glazko G.V.**, Rogozin I.B., Glazkov M.V. Comparative study and prediction of DNA fragments associated with various elements of the nuclear matrix. (2001). *Biochim Biophys Acta*, **1517**: 351-364.

Highlight: This publication motivated other authors to propose several alternative approaches, but an independent analysis by Japanese colleagues (Purbowasito et al., DNA Res., 2004) showed a higher specificity/sensitivity of the ChrClass method as compared to other MARs *in silico* prediction tools.

3. **GADIST**, R script to compute Generalized-Average Based Distances

What it does? It calculates Generalized-Average Based distance between binary vectors.

Reference paper: **Glazko G.**, Gordon A., Mushegian, A. (2005). The choice of optimal distance measure in genome-wide data sets. *Bioinformatics*, **21**:Suppl 3:iii3-iii11.

4. **PSI-SQUARE**, software for local similarity searches

What it does? It implements similarity search in omics data sets based on new BLAST-like algorithm.

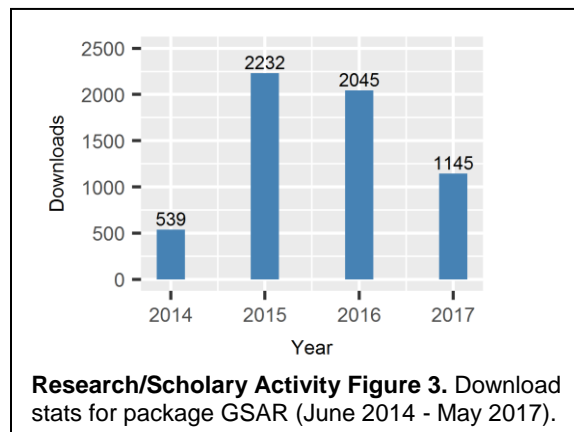
Reference paper: **Glazko G.**, Coleman M., Mushegian, A. (2006). Similarity searches in genome-wide numerical data sets. *Biology Direct*, **1**: 13.

5. **GSAR**, Gene Set Analysis in R: a Bioconductor package that provides a set of multivariate statistical tests for omics data.

What it does? It implements several multivariate statistical tests for pathway analysis in omics data.

Reference paper: Rahmatallah Y., Zybaylov B., Emmert-Streib F., **Glazko G.** (2017). GSAR: Bioconductor package for Gene Set Analysis in R. *BMC Bioinformatics*, **18**: 61.

Highlight: The package is freely available on the Bioconductor website www.bioconductor.org (more than 6,000 total downloads since 2014, Figure 3)



M. Patents

1. U.S. Provisional Application No.: 62/465,588. Title: Compositions and Methods for Detecting Sessile Serrated Adenomas/Polyps. UAMS Reference No.: 2017-17 Inventors: Rahmatallah, Glazko and Hagedorn

N. Peer Review

1. *Bioinformatics* (5 manuscripts)
2. *BMC Bioinformatics* (1 manuscript)
3. *Genome Biology* (1 manuscript)
4. *PLOS Computational Biology* (1 manuscript)
5. *Scientific Reports* (2 manuscripts)
6. *PLOS ONE* (4 manuscripts)
7. *Gene* (4 manuscripts)
8. *Genome Research* (2 manuscripts)
9. *Molecular Biology and Evolution* (2 manuscripts)
10. *Complexity* (1 manuscript)
11. *Neurobiology of Aging* (1 manuscript)
12. *BMC Genomics* (2 manuscripts)
13. *Biostatistics* (1 manuscript)
14. *Cellular & Molecular Biology Letters* (1 manuscript)

15. *IET Systems Biology* (1 manuscript)