

# JAHANGHEER SHAIK

8102 McDonnell Pediatric Research Building, 660 S. Euclid Ave., Saint Louis, MO 63110 |  
W:(314)362-3252 | C:(314)225-7392 | shaik@wustl.edu,jahangeer.shaik@gmail.com |  
<http://bioinformatics.micro.wustl.edu/>

## SUMMARY

- Experience contributing to applications such as Washington University Genomics and Pathology Services ([WUGPS](#)- also see [Pierian Dx](#)) to diagnose diseases such as cancers, cardiac disease and renal disease to facilitate personalized medicine through the use of Next-generation sequencing (NGS) data
- Multi-disciplinary data scientist with experience working with NGS, Microarray, Image and Text data
- Experienced Bioinformatics Professional with strong Leadership and collaboration-building skills
- Over 15 years of experience in Research, Development and Training
- Experience working with diseases spread by Micro organisms and by Cancers

## HIGHLIGHTS

- Bioinformatics Professional with Complex Problem Solving Skills using Big Data
- Experience Working on Diverse Datasets and Research Problems
- Track Record of Publishing in High Impact Journals
- Team Leadership
- Inter-Disciplinary Researcher
- Supervision and Training
- Fast Learner
- Computer-Savvy

## SKILLS

### Research

- Experience working and finding solutions to some of the most complex and diverse problems using Big data such as from NGS data, Microarray data, Image data and Free text as demonstrated by publications in high impact journals
- Proven ability to spearhead projects from multiple disciplines such as from Pathology and Immunology and Molecular Microbiology

### Leadership

- Trained and managed other faculty, post doctoral candidates, staff and students in The Department of Molecular Microbiology ([DMMB](#)) and supported the research performed by different labs
- Managed a group of scientists, computer professionals and staff to provide research support to investigators across the medical campus at Washington University and to build, support and improve tools and applications hosted by Center for Biomedical informatics in the Department of Pathology and Immunology

## WORK EXPERIENCE

### Instructor, DMMB

11/2011 to current

#### Washington University School of Medicine

Saint Louis, MO

- Spearheaded research to understand causal factors behind [Leishmaniasis](#), [Toxoplasmosis](#), and [Tuberculosis](#)
- Worked closely with the [Chairman](#) of the DMMB to understand the requirements and bioinformatics needs of DMMB
- Supervised Faculty, Students and other Scientists and Trained them in basic Bioinformatics Analysis and Approach
- Created Technical and Research documents and published them in high impact journals
- Co-drafted requests for Proposals and Statements of Work
- Developed Scientific Analysis Software to solve complex problems ([REDHORSE](#), [PAVED](#), and [AGELESS](#)) using NGS data
- Worked on genomes with various ploidies and developed unique strategies to handle these disparate data

### Staff Scientist, Department of Pathology and Immunology 01/2008 to 11/2011

#### Washington University School of Medicine

Saint Louis, MO

- Contributed to the bioinformatics segment of [WUGPS](#), an application to diagnose and facilitate personalized treatment for diseases such as cancers
- Supported research of several investigators by providing them with analysis and interpretation of their high throughput data
- Supervised Computer professionals, technologists and other scientists to develop new applications and to improve and support existing applications
- Developed tools and applications adhering to syntactic and semantic interoperability as required by cancer bioinformatics grid ([caBIG®](#)) for managing [array data](#), [tissue banking](#) and [clinical trials](#)
- Tested tools and applications developed by other caBIG® groups for syntactic and semantic interoperability
- Integrated information from several databases such as pathway databases, protein-interaction databases, transcription-target interaction databases, scientific literature data and array profiling data and generated applications [GRANITE](#) and [ERINA](#) for creating gene-relational networks

## EDUCATION

**Doctor of Philosophy: Electrical and Computer Engineering**      **December 2007**

**University of Memphis, Memphis, TN**

**Research Topic:** Knowledge Discovery in Microarrays

**Support:** Herff College of Engineering Fellowship, Travel grant from IEEE International joint conference in Neural Networks

**Master of Science: Bioinformatics**      **December 2007**

**University of Memphis, Memphis, TN**

**Research Topic:** Visualization Techniques for Microarray Data

**Highlights:** First student to graduate from the Department of Bioinformatics, Bioinformatics Summer grant-2006

**Master of Science: Electrical and Computer Engineering**      **August 2003**

**University of Memphis, Memphis, TN**

**Research Topic:** Hilbert-Wavelet transform based Rotated target Detection and Tracking

**Support:** Research Assistant in Intelligent Systems and Image Processing lab, Teaching assistant for Courses such as Image Processing and Computer Vision

**Bachelor of Technology: Electronics and Communications Engineering July 2001**

**Jawaharlal Technological University, Hyderabad, India**

**Research Topic:** Intra and inter- networks for scientific Development and Analysis

**Role and Award:** IEEE regional coordination committee member, best outgoing student of 2001.

## CERTIFICATIONS

Big Data and Hadoop Essentials from Udemy

Data Scientists Tool box from Johns Hopkins

## TOOLS AND APPLICATIONS BUILT

Only freely accessible online tools are listed here.

- 1) [REDHORSE](#)- tool to extract recombinations from next-generation sequencing data
- 2) [PAVED](#)- A tool to extract peaks and valleys from next-generation epigenomics data
- 3) [AGELESS](#)- A pipeline to extract parental inheritance from next-generation sequencing data

## PROFESSIONAL MEMBERSHIPS

- American Medical Informatics Association member (July 2009 to present)
- IEEE member, region 5 (Jan 2005 to Dec 2008)
- IEEE student regional coordination committee lead (April 1998 to Dec 2001)

- Active member of caBIG®. Leading architecture reviewer and member of analytical services best practices group (July 2008 to November 2011)
- Member of Bioinformatics organization ([www.bioinformatics.org](http://www.bioinformatics.org)) (Jan 2007 to present)
- Member of Open Bioinformatics Forum (<http://www.open-bio.org>) (March 2009 to present)

## TECHNICAL SKILLS

**Programming:** JAVA, C, R, PERL, Python, MATLAB and MAPLE

**Databases:** Oracle 9i/8i, SQL, PL/SQL, SQL plus and MS Access

**Operating Systems:** Windows, MAC and Linux (All distributions and versions)

## SCHOLARLY REFERENCES

**Co-first authorship is indicated by \***

### Manuscripts under Internal Review

1. B.A. Anderson, **J.S. Shaik\*** and S.M. Beverley, "The Chromatin Landscape of Early-Diverging Eukaryote *Leishmania major*" To be submitted to PLoS Genetics
2. **J.S. Shaik**, B.A. Anderson and S.M. Beverley, "PAVED- A Software suite for the analysis of epigenome-derived next-generation sequencing data" To be submitted to Genome Biology
3. **J.S. Shaik**, D.E. Dobson and S.M. Beverley, "Study of Aneuploidy and its inheritance using intra and inter-species experimental hybrids in *Leishmania*" to be submitted to PLoS genetics
4. E. Brettmann, H. Zangger, **J.S. Shaik**, L. Lye, M. Kuhlmann, K. Owens, S. Hickerson, N. Fasel and S.M. Beverley, "Elimination of a persistent endobiont virus in *Leishmania* by RNAi", to be submitted to PNAS

### Published Manuscripts

1. **Shaik JS**, Khan A, Beverley SM, Sibley DL. REDHORSE-REcombination and Double crossover detection in Haploid Organisms using next-generation SEquencing data. BMC Genomics. 2015;16:133
2. Romano A, Inbar E, Debrabant A, Charmoy M, Lawyer P, Gomes FR, Barhoumi M, Grigg M, **Shaik J**, Dobson DE, Beverley SM, Sacks D. Cross-species genetic exchange between visceral and cutaneous strains of *Leishmania* in the sand fly vector. PNAS. 2014;111(47):16808-13.
3. Khan A, **Shaik JS\***, Behenke M, Wang Q, Dubey JP, Lorenzi HA, Ajioka JW, Rosenthal BM, Sibley DL. NextGen sequencing reveals short double crossovers contribute disproportionately to genetic diversity in *Toxoplasma gondii*. BMC Genomics. 2014;15(1168):1-28.

4. **Shaik JS**, Yeasin M, Russomanno DJ. Evaluation of supervised and unsupervised 3D star visualisation algorithms. *International journal of data mining and bioinformatics*. 2013;8(4):443-61.
5. Inbar E, Akopyants NS, Charmoy M, Romano A, Lawyer P, Elnaiem DA, Kauffmann F, Barhoumi M, Grigg M, Owens K, Fay M, Dobson DE, **Shaik JS**, Beverley SM, Sacks D. The mating competence of geographically diverse *Leishmania* major strains in their natural and unnatural sand fly vectors. *PLoS Genetics*. 2013;9(7):1-14.
6. Huchtagowder V, **Shaik JS\***, Fiala M, Paasch J, Nagarajan R, Goldstein KE, DiPersio JF, Tomasson MH, R.Vij, Kulkarni S. Genomic Landscape of Immunoglobulin Light Chain (AL) Amyloidosis and Comparative Analyses with Related Malignant Plasma Cell Disorder-Multiple Myeloma. *Blood*. 2011;118(21):367.
7. Deshmukh H, Yu J, **Shaik JS**, MacDonald JT, Tobey J, Perry A, Payton JE, Gutmann DH, Watson MA, Nagarajan R. Identification of transcriptional regulatory networks specific to pilocytic astrocytoma. *BMC medical genomics*. 2011;4(1):57.
8. **Shaik JS**, Yeasin M. Fuzzy-adaptive-subspace-iteration-based two-way clustering of microarray data. *ACM Transactions on Computational Biology and Bioinformatics*. 2009;6(2):244-59.
9. **Shaik JS**, Iftexharuddin KM. Detection and tracking of targets in infrared images using Bayesian techniques. *Optics and Laser Technology*. 2009;41(6):832-42.
10. Grieselhuber NR, **Shaik JS\***, Chang LW, McGrath S, Wartman LD, Nagarajan R, Wilson RK, Mardis E, Ley TJ. Chromatin Immunoprecipitation of GFP-Tagged PML-Rara Coupled to High-Throughput Next Generation Sequencing. *AMER SOC HEMATOLOGY*. 2009:524.
11. Yeasin M, Vanteru BC, **Shaik JS**, Ahmed F. i-SEGOPubMed: a web interface for semantic enabled browsing of PubMed using Gene Ontology. *BMC Bioinformatics*. 2008;9(7):P20.
12. Yeasin M, Russomanno DJ, David J, Sorower MS, Smith M, **Shaik JS**. Robust Classification of Objects using a Sparse Detector Sensor. *IC-AI*. 2008:742-8.
13. Vanteru BC, **Shaik JS\***, Yeasin M. Semantically linking and browsing PubMed abstracts with gene ontology. *BMC Genomics*. 2008;9(1):S10.
14. **Shaik JS**, Yeasin M. A Unified Framework To Find Differentially Expressed Genes from Microarray Experiments. *International Joint Conference on Neural Networks*. 2007:2598--603.
15. **Shaik JS**, Yeasin M. Selection of best projection from 3d star coordinate projection space using energy minimization and topology preserving mapping. *International Joint Conference on Neural Networks*. 2007:2604-9.
16. **Shaik JS**, Yeasin M. Ranking function based on higher order statistics (RF-HOS) for two-sample microarray experiments. *Bioinformatics Research and Applications: Springer*; 2007. p. 97-108.
17. **Shaik JS**, Yeasin M. Two-way Clustering using Fuzzy ASI for Knowledge Discovery in Microarrays. *Computational Intelligence and Bioinformatics and Computational Biology*, 2007. 2007:39-45.
18. **Shaik JS**, Yeasin M. A unified framework for finding differentially expressed genes from microarray experiments. *BMC Bioinformatics*. 2007;8(1):1-21.
19. **Shaik JS**, Olusegum GE, Yeasin M. An Empirical CDF Approach to Estimate the Significance of Gene Ranking for Finding Differentially Expressed Genes. *Bioinformatics and Bioengineering*. 2007:1071-5.

20. **Shaik JS**, Yeasin M. Functionally Classifying Genes from Microarray Data Using Linear and Non-linear Data Projection. Proceedings of the IEEE International Conference on Computer Systems and Applications. 2006:608-15.
21. **Shaik JS**, Yeasin M. Performance Evaluation of Subspace-based Algorithm in Selecting differentially Expressed Genes and Classification of Tissue Types from Microarray Data. International Joint Conference on Neural Networks. 2006:2972-9.
22. **Shaik JS**, Yeasin M. A progressive framework for two-way clustering using adaptive subspace iteration for functionally classifying genes. International Joint Conference on Neural Networks. 2006:2980-5.
23. **Shaik JS**, Yeasin M. Visualization of high dimensional data using an automated 3D star co-ordinate system. International Joint Conference on Neural Networks. 2006:1339-46.
24. **Shaik JS**, Kothari CR, Russomanno DJ, Yeasin M. 3D Visualization of Relation Clusters from OWL Ontologies. SWWS. 2006:17-23.
25. Iftekharuddin KM, Islam MA, **Shaik JS**, Parra C, Ogg R. Automatic brain tumor detection in MRI: methodology and statistical validation. Optics and Photonics. 2005:2012-22.
26. **Shaik JS**, Iftekharuddin KM. Probabilistic detection and tracking of IR targets. Optical Science and Technology, SPIE 49th Annual Meeting. 2004:90-101.
27. **Shaik JS**, Iftekharuddin KM. Detection and tracking of rotated and scaled targets by use of Hilbert-wavelet transform. Optical Society of America. 2003;42(23):4718-35.
28. **Shaik JS**, Iftekharuddin KM. Automated tracking and classification of infrared images. International Joint Conference on Neural Networks. 2003;2:1201-6.
29. Iftekharuddin KM, **Shaik JS**. Distorted IR target detection and tracking using composite filters. Optical Science and Technology, SPIE 48th Annual Meeting. 2003:73-84.
30. Iftekharuddin KM, **Shaik JS**, Awwal AAS, Alam MS. Hilbert wavelet transform for recognition of image rotation. International Symposium on Optical Science and Technology. 2002:147-58.