OMB No. 0925-0001 and 0925-0002 (Rev. 03/2020 Approved Through 02/28/2023)

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Ganapathiraju, Madhavi K

eRA COMMONS USER NAME: madhavi

POSITION TITLE: Associate Professor

EDUCATION/TRAINING

| INSTITUTION AND LOCATION | DEGREE | Completion DateMM/YYYY | FIELD OF STUDY |
| --- | --- | --- | --- |
| Delhi University, New Delhi, India | B.Sc. | 05/1993 | Mathematics, Physics, & Electronics |
| Indian Institute of Science, Bangalore, India | M.Eng. | 05/1998 | Electrical and Communications Engg |
| Carnegie Mellon University, Pittsburgh, PA USA | Ph.D. | 05/2007 | Language and Information Technologies (Computer Science) |

# A. Personal Statement

**Research**: I have training in Computer Science and Communications Engineering. Aided by knowledge and experience in areas of machine learning, signal processing, information technologies and high-performance computing, I have been successful in developing algorithms for bioinformatics, computational systems biology, network biology and biomedicine. I received the NIMH BRAINS award (Biobehavioral Research Awards for Innovative New Scientists) with $1.79M funding to study the protein-protein interactions (PPIs) in the human interactome. I also received supplemental fundingto develop a website (Wiki-Pi) to disseminate these results. (See C.3). Our results from this project are making significant translational impact on biology.

Through our work on schizophrenia interactome we closed the apparent gap between GWAS genes and historically considered genes of SZ. We identified repurposable for schizophrenia two of which are in clinical trials. It showed the interactome surrounding the DNA variants that may be conferring a negative association of rheumatoid arthritis in families of patients with schizophrenia. During COVID19 pandemic, we constructed the interactome of proteins that are targeted by SARS-CoV-2 virus and analyzed the interactome and released our findings immediately via a preprint publication. Our extensive work on 5 neuropsychiatric disorders, results of which are presented in the proposal, is in preparation for publication.

**Collaborations for translations of computational work**: I have established collaborations regionally and internationally centered around the translation of research outcomes to biology. These joint efforts are leading to highly impactful translational results as seen in our publications on the interactomes of genes associated with diseases. I have ongoing collaboration with the Co-I Vishwajit Nimgaonkar (resulting in 2 papers and 2 clinical trials). I have recently started a collaboration with Co-I Donald DeFranco. He is planning to carry out experimental validations of new PPIs we put together for their study. In the coming months Co-Is and I will start collaborations to validate some new PPIs in their labs. By bringing together this team of Co-Is, consultants and other significant contributors, and with highly qualified doctoral students whom I can recruit from programs at Pitt and Carnegie Mellon University, I am confident that I can carry out the work proposed here. (See C.2)

**Mentoring students in multidisciplinary STEM fields:** As is common in computer science field, I work predominantly with graduate students as opposed to postdoctoral scholars and technicians. I have trained numerous graduate and undergraduate biomedical science students in computer science and bioinformatics, and vice versa. (See C.5)

**Relevant Publications:**

1. Ganapathiraju M.K., Thahir M., Handen A., Sarkar S., Sweet R.A., Nimgaonkar V.L., Loscher C., Bauer E. & Chaparala S. (2016).
*Schizophrenia interactome with 504 novel protein-protein interactions*.
**NPJ Schizophrenia**. 2:16012. PMCID: PMC4898894.
2. Karunakaran, K.B., Chaparala, S. and Ganapathiraju, M.K., 2019.
*Potentially repurposable drugs for schizophrenia identified from its interactome*.
**Scientific Reports**, 9(1), pp.1-14. PMCID: PMC6722087.
3. Mohamed, T.P., Carbonell, J.G. and Ganapathiraju, M.K., 2010.
*Active learning for human protein-protein interaction prediction*.
**BMC Bioinformatics**, 11(1), p.S57. PMCID: PMC3009530
4. Li Y., Klena N., Gabriel G., Liu X., Lemke K., Chen Y., Kim A., Chatterjee B., Damerla R., Chang C., Yagi H., Shan Y., Agustin J.S., Thahir M., Anderton S., Lawhead C., Vescovi A., Weeks D., Pratt H., Morgan J., Haynes L., Smith C., Eppig J., Reinholdt L., Leatherbury L., Tobita K., Francis R., Ganapathiraju M.K., Pazour G. & Lo C.W. (2015).
*Global genetic analysis in mice unveils central role for cilia in congenital heart disease*.**Nature**, 521(7553):520-524. PMCID: PMC4617540.
5. **Positions and Honors**

**Positions and Employment**

|  |  |
| --- | --- |
| 1998 - 2000 | Project Associate, Indian Institute of Science |
| 2000 | Visiting Scholar, Carnegie Mellon University, Pittsburgh, PA |
| 2001 - 2007 | Multimedia Computer Information Specialist, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA |
| 2004 | Mentor, International Institute of Information Technology, Hyderabad, India |
| 2008 - 2016 | Assistant Professor, Intelligent Systems Program, University of Pittsburgh |
| 2008 -  | Member, Molecular & Cellular Cancer Biology Program, University of Pittsburgh Cancer Institute Pittsburgh |
| 2008 -  | Faculty, University of Pittsburgh, Joint CMU-Pitt PhD Program |
| 2008 - 2016 | Assistant Professor, Department of Biomedical Informatics, University of Pittsburgh |
| 2012 -  | Advisory Board Member, Biotechnology Innovation and Computing Program, Carnegie Mellon University, Pittsburgh, PA |
| 2014 -  | Faculty, Integrative Systems Biology Program, University of Pittsburgh |
| 2016 -  | Associate Professor, Department of Biomedical Informatics, University of Pittsburgh |
| 2016 -  | Associate Professor, Intelligent Systems Program, University of Pittsburgh |

**Other Experience and Professional Memberships**

|  |  |
| --- | --- |
| 2007 -  | Member, American Medical Informatics Association (AMIA) |
| 2008 -  | Member, Admissions Committee The Intelligent Systems Program (ISP)  |
| 2008 -  | Member, Psychoneuroimmunology Research Society (PNIRS) |
| 2008 -  | Member, Admissions Committee Department of Biomedical Informatics, Pitt |
| 2008 -  | Member, Curriculum Committee Department of Biomedical Informatics, Pitt |
| 2009 - 2013 | Member, Diversity Liason |
| 2013 | Reviewer of EUREKA grants: Special Emphasis Panel ZMH1 ERB-M(07), NIMH |
| 2013 | Reviewer of BRAINS grants: Special Emphasis Panel: ZMH1 ERB-L(04), NIMH |
| 2014 | Reviewer of BRAIN Initiative grants: Special Emphasis Panel ZMH1-ERB-M(06), NIMH |
| 2014 | Reviewer of EUREKA grants: Special Emphasis Panel ZMH1 ERB-R (02), NIMH |
| 2014 -  | Member, Curriculum Committee, Integrative Systems Biology Program, Pitt |
| 2016 - 2019 | Editorial Board Member, IEEE Consumer Electronics Magazine |
| 2017 | Reviewer for EBIT Study Section Proposals, NIH |
| 2017 | Reviewer of EUREKA grants: Special Emphasis Panel ZMH1 ERB-R (02), NIMH |

**Honors**

|  |  |
| --- | --- |
| 2011  | BRAINS 2011 from NIMH: Biobehavioral Research Awards for Innovative New Scientists from National Institute of Mental Health |

**C. Contributions to Science**1. Discovery of novel protein-protein interactions in the human interactome

I have been focused on discovering protein-protein interactions (PPIs) using computational methods. While I have a Ph.D. in Computer Science from the top-ranking Carnegie Mellon University School of Computer Science, I chose to join School of Medicine as faculty, because although the methods are computational, my interest is in making impact on biology and medicine. So on one hand I develop these methods with computationally strong group of trainees in my group while on the other hand I discuss extensively with biologists from school of medicine so that they can accelerate their research using our predictions. When results are presented to a specific biologist, it is no good if the algorithm has good performance on test data; when the biologist invests resources based solely on our predictions, we need to be sure that algorithm performs accurately for proteins of interest to *that specific* biologist. So, I have been developing methods to make sure that our results are accurate even if only few (i.e. achieve high precision even if recall is low). Thus, we identified thousands of novel PPIs that we deemed accurate based on several computational and experimental evaluations. Our specific activities include development of methods for active learning, machine learning, and web application development – to discover and disseminate protein-protein interactions.
***See Relevant Publications #1 and 3, and:***

1. Ananthasubramanian S., Carbonell J.G. & Ganapathiraju M.K. (2013). Active learning with one-class and reluctant oracles in human protein-protein interaction prediction. WCSB 2013; Finland.
2. Ganapathiraju M.K. & Orii N. (2013). Research prioritization through prediction of future impact on biomedical science: a position paper on inference-analytics*.* GigaScience*,* 2(1).
3. Thahir M., Sharma T. & Ganapathiraju M.K. (2012). An efficient heuristic method for active feature acquisition and its application to protein-protein interaction prediction. BMC Proceedings, 6:S2. PMCID: PMC3504800.
4. Embar V., Handen A. & Ganapathiraju M.K. (2016). Is the average shortest path length of gene set a reflection of their biological relatedness? Journal of Bioinformatics and Computational Biology14:6, SI.PMCID: PMC5726383.

2. Translation of novel protein-protein interactions.

We collaborated with biologists to validate some of the novel PPIs (of thousands). All 16 of 16 experimentally tested novel PPIs have been validated to be true predictions. I am taking great interest in bringing these predicted PPIs to biologists through face-to-face meetings (which is how I have long standing collaborations with Dr. Vishwajit Nimganokar and Dr. Cecilia Lo and others), presenting posters at biology-focus meetings (university based retreats, annual meetings of research societies, etc.), as well as by developing online tools for the dissemination of PPIs. For these efforts, I have been awarded administrative supplement by NIMH to develop a website of PPIs with a search engine that would allow biologists to ask intuitive and biologically relevant questions. This website (called Wiki-Pi) was launched in November 2013 and it gets accessed by 600 to 2,500 unique visitors each month. We also developed another website (‘LENS: lens for enrichment and network analysis of human genes’) that allows biologists to get an interactome network-based analysis of their gene sets, with the click of a button. Thus, I pay attention to the needs of biologists and promptly develop mechanism to disseminate the information in the right way. It is these strong interactions with biologists that resulted in high impact publications, high-visibility website(s), and two clinical trials.
***See Relevant Publications #1, 3 and 4, and:***

1. Zhu J., Zhang Y., Ghosh A., Cuevas R.A., Forero A., Dhar J., Ibsen M.S., Schmid-Burgk J.L., Schmidt T., Ganapathiraju M.K., Fujita T., Hartmann R., Barik S., Hornung V., Coyne C.B. & Sarkar S.N. (2014). *Antiviral activity of human OASL protein is mediated by enhancing signaling of the RIG-I RNA sensor.* **Immunity**, 40 (6):936-48. PMCID: PMC4101812.
2. Liu X., Yagi H., Saeed S., Bais A.S., Gabriel G.C., Chen .Z, Peterson K.A., Li Y., Schwartz M.C., Reynolds W.T., Saydmohammed M., Gibbs B., Wu Y., Devine W., Chatterjee B., Klena N.T., Kostka D., de Mesy Bentley K.L., Ganapathiraju M.K., Dexheimer .P, Leatherbury L., Khalifa O., Bhagat A., Zahid M., Pu W., Watkins S., Grossfeld P., Murray S.A., Porter G.A. Jr., Tsang M., Martin L.J., Benson D.W., Aronow B.J. & Lo C.W. (2017).
*The complex genetics of hypoplastic left heart syndrome*.
**Nature Genetics**, 49(7):1152-1159. PMCID: PMC5737968.
3. Malavia T.A., Chaparala S., Wood J., Chowdari K., Prasad K.M., McClain L., Ganapathiraju M.K. & Nimgaonkar V.L. (2017).
Generating testable hypotheses for schizophrenia and rheumatoid arthritis pathogenesis by integrating epidemiological, genomic and protein interaction data.
**NPJ Schizophrenia**, 3:11. PMCID: PMC5441529.
4. Ganapathiraju, M.K., Karunakaran, K.B. and Correa-Menéndez, J., 2016.
*Predicted protein interactions of IFITMs may shed light on mechanisms of Zika virus-induced microcephaly and host invasion*. **F1000Research**, 5. PMCID: PMC5747333

3. Open source and web-based tools

We developed Wiki-Pi and LENS web applications to disseminate the protein-protein interactions to biologists. The web applications are designed with biologists in mind to make several annotations of the proteins available at the click of a button. For example, searching for any protein-protein interaction shows not only the information about the interaction but several functional annotations of the two proteins side-by-side so that the biologist may interpret the information conveniently. Similarly, in LENS, given a set of genes, its interactome, and biological annotations of all these genes and their interactome, including pathways, diseases, drugs, etc automatically annotated and statistically analyzed, images automatically generated and makes the information downloadable in formats compatible with other software. Previously, I developed the Biological Language Modeling Toolkit, which we have made available in open source and have also used it ourselves extensively in several sequence analysis applications. I also developed the web-application to TM helix prediction algorithm (this web-application had to be retired when I moved after Ph.D. and my advisor ran out of funds to maintain it). I also developed TelMor Telugu language morphological generator and designed and supervised the development of Om transliteration software. Thus, for the resources that I developed, my goal has always been to make it available to users with ease-of-use features.

1. Orii N. & Ganapathiraju MK. (2012).
*Wiki-pi: a web-server of annotated human protein-protein interactions to aid in discovery of protein function*. **PloS One**, 7(11):e49029. PMCID: PMC3509123.
2. Handen A. & Ganapathiraju M.K. (2015).
*LENS: Web-based Lens for Enrichment and Network Studies of Human Proteins*.
**BMC Medical Genomics**, 8 Suppl 4:S2. PMCID: PMC4682415.
3. Ganapathiraju M.K.*,* Mitchell A.D., Thahir M., Motwani K. & Ananthasubramanian S*.* (2012). *Suite of tools for statistical N-gram language modeling for pattern mining in whole genome sequences*. **Journal of Bioinformatics and Computational Biology**,10(6):1250016.
4. Ganapathiraju M.K., Jursa C.J., Karimi H.A. & Klein-Seetharaman J. (2017).
*TMpro Web Server and Service: Transmembrane Helix Prediction Through Amino Acid Property Analysis*. **Bioinformatics** (Oxford, England), 23 (20):2795-6.

4. Biological Language Modeling using “N-gram” analysis

For my Ph.D. research, I had worked on a novel concept of biological language modeling. While the original idea of applying language processing algorithms to genome and proteome sequences to answer biologically relevant questions was proposed by my Ph.D. advisors Prof. Raj Reddy and Prof. Judith Klein-Seetharaman, I had developed original ideas on using n-gram analysis and latent semantic analysis of protein/proteome sequences. Application using latent semantic analysis for transmembrane helix prediction reduced the error rate of TM helix prediction by 50% by taking then-saturated accuracy from 90% to 95%, and the n-gram analysis work is highly-cited as seminal work that was adopted by the bioinformatics community into algorithm development for various different protein sequence analysis applications.

1. Ganapathiraju M.K., Weisser D., Klein-Seetharaman J., Rosenfeld R., Carbonell J.G. & Reddy R. (2002). *Comparative N-Gram Analysis of Whole-Genome Sequences*.
**HLT2001: Human Language Technologies Conference**, San Diego, CA. (Full-length peer-reviewed)
2. Ganapathiraju M.K., Klein-Seetharaman J., Balakrishnan N. & Reddy R. (2004).
*Characterization of Protein Secondary Structure using Latent Semantic Analysis*.
**IEEE Signal Processing Magazine**, 21 (3):78-87.
3. Ganapathiraju M.K., Balakrishnan N., Reddy R. & Klein-Seetharaman J. (2008).
*Transmembrane helix prediction using amino acid property features and latent semantic analysis*.
**BMC Bioinformatics**, 9:S4. PMCID: PMC2259405.
4. Giorgio E., Rolyan H., Kropp L., Chakka A.B., Yatsenko S., Gregorio E.D., Lacerenza D., Vaula G., Talarico F., Mandich P., Toro C., Pierre E.E., Labauge P., Capellari S., Cortelli P., Vairo F.P., Miguel D., Stubbolo D., Marques L.C., Gahl W., Boespflug-Tanguy O., Melberg A., Hassin-Baer S., Cohen O.S., Pjontek R., Grau A., Klopstock T., Fogel B., Meijer I., Rouleau G., Bouchard J.P., Ganapathiraju M.K., Vanderver A., Dahl N., Hobson G., Brusco A., Brussino A. & Padiath Q.S. (2013).
*Analysis of LMNB1 Duplications in Autosomal Dominant Leukodystrophy Provides Insights into Duplication Mechanisms and Allele-Specific Expression*.**Human Mutation**,34(8): 1160-1171. PMCID: PMC3714349.

5. Training Undergraduate and/or Underrepresented Minorities Students to Research in STEM fields

I have trained several graduate and undergraduate students in Science, Technology, Engineering and Mathematics (STEM) fields, often giving them training in cross disciplinary fields from their major field of study: For example, Asia D. Mitchell who was a genetics major with no prior experience in informatics in undergrad studies, went on to win a Best Session in Bioinformatics award for the work that she carried out with me; obtained a PhD in genetics with bioinformatics specialization and is now a postdoctoral scholar in this field. Similarly, Helen Li is another biology undergraduate student who gained experience in multidisciplinary field with me and went on to get MS in human computer interaction studies and joined Bioinformatics company GenoSpace. Another stellar example is Kalyani Bindu who joined me during her undergraduate education and got highly motivated in research involving biology and bioinformatics, and worked with me in post-baccalaureate research and wrote 4+ high impact papers published in Scientific Reports, F1000 Research, etc. There are several more examples, and each of these students has even produced high quality research products, published in prestigious journals.

***Examples of these are: #2 in Relevant Publications, C.2.c, C.2.d, C.3.c, and:***

1. Chattopadhyay A, Ganapathiraju MK. *Demonstration Study: A Protocol to Combine Online Tools and Databases for Identifying Potentially Repurposable Drugs*.
**Data**. 2017 Jun;2(2):15.
2. Kuppuswamy U., Ananthasubramanian S., Wang Y., Balakrishnan N. & Ganapathiraju MK. (2014). *Predicting gene ontology annotations of orphan GWAS genes using protein-protein interactions*.
Algorithms for Molecular Biology, 9(1):10. PMCID: PMC4124845.
3. Ananthasubramanian S., Metri R., Khetan A., Gupta A., Handen A., Chandra N. & Ganapathiraju M.K. (2012). *Mycobacterium tuberculosis and Clostridium difficille interactomes: demonstration of rapid development of computational system for bacterial interactome prediction*.Microbial Informatics and Experimentation, 2:4. PMCID: PMC3353838.
4. Osmanbeyoglu H.U., WehnerJ.A., Carbonell J.G. & Ganapathiraju MK. (2010). *Active machine learning for transmembrane helix prediction*.
**BMC Bioinformatics**, 11 Suppl 1:S58. PMCID: PMC3009531.

**Google Scholar:**
<https://scholar.google.com/citations?user=madhavicmu>

##### **D. Additional Information: Research Support and/or Scholastic Performance**

**Ongoing Research**

**Grant ID** **612571 Ganapathiraju (PI) 01/01/19 – 12/31/20
Mesothelioma Applied Research Foundation** Role: MPI

Title Genetic Variants, Pathways and Drug Targets of Peritoneal Mesothelioma

To carry out interactome analysis of genes associated with malignant peritoneal mesothelioma