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: Disha Sharma

eRA COMMONS USER NAME (credential, e.g., agency login): DISHASHARMA

POSITION TITLE: Postdoctoral Fellow, Cardiovascular Medicine, Stanford University

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

|  |  |  |  |
| --- | --- | --- | --- |
| INSTITUTION AND LOCATION | DEGREE*(if applicable)* | Completion DateMM/YYYY | FIELD OF STUDY |
|  |  |  |  |
| Banasthali University, Rajasthan, India | B.Sc. | 06/2010 | Biotechnology |
| Indian Institute of Technology, Roorkee, India | M.Sc. | 06/2012 | Biotechnology |
| CSIR-Institute of Genomics and Integrative Biology | Ph.D. | 03/2020 |  |

**A. Personal Statement**

I am currently a Postdoctoral Fellow with Dr. Thomas Quertermous at Stanford University. I have joined the lab with more than 7 years of research experience in the field of computational biology wherein I have worked with multi-omics data for multiple diseases to get a deeper understanding of the disease identification and progression.

 My background in engineering and bioinformatics provide an excellent background for the studies proposed in this application, which proposes to investigate the genetics and genomics of smooth muscle cell biology in the context of vascular disease. I first pursued a Bachelor's in Biotechnology program at one of the premier institutes in India, Banasthali Vidyapeeth and received my degree in 2007. After qualifying with the IIT-JAM exam in 2010, I joined the Master’s in Science (Biotechnology) program at the prestigious Indian Institute of Technology Roorkee in a program of engineering and technology. After my Master's, I joined Dr. Vinod Scaria’s lab at CSIR-IGIB as a Project Fellow. During the tenure as Project fellow from 2012-2014, I had the opportunity to work with different transcriptomics data from model organisms including zebrafish, rat and human cell lines to understand the role of long non-coding RNAs and miRNAs. I also worked on clinical datasets of autoimmune disorders. With one and half years of research experience and a UGC fellowship awarded through the NET-JRF examination, I continued working with Dr. Vinod Scaria to pursue my PhD. My research interest for the degree focused on the identification and characterization of circular RNAs, and this work has now been published in multiple manuscripts listed below. Over the years at CSIR-IGIB, I have had the chance to work on interesting ideas with multiple collaborating groups. One of them was Dr. Sridhar Sivasubbu, with whom I worked to understand the transcript-level interactions between mitochondria and the nucleus, using zebrafish as a model organism.

 In view of my interest in the translational aspects of biology, I obtained the opportunity to work as part of the GUaRDIAN Consortium with Dr. Vinod Scaria and Dr. Sridhar Sivasubbu at CSIR-IGIB. This pioneering project is the largest network of researchers and clinicians in India pursuing sequencing patient DNAs to identify rare SNVs and structural variants responsible for muscular dystrophy in these patients. In the interest of advancing genomics in clinical and healthcare settings, I was selected as Intel Fellow 2019 to work for the Intel-IGIB collaboration focussing on “Accelerating Clinical Analysis and Interpretation of Genomic Data through advanced tools/libraries”. Our project was selected among top 3 from 50 premier research institutes and I was awarded the Intel-India Fellowship for a year to pursue this project. I was also part of the core team of IndiGen (Genomes for Public Health in India). With the spread of COVID-19 around the world, our group contributed by sequencing and analysing COVID19 genomes to get a better understanding of the disease and I had the opportunity to be part of the core team to analyse the viral sequencing datasets and viral assembly.

 I am extremely pleased to have joined the Quertermous lab at Stanford where I can apply my command line Unix skills and basic understanding of genetics and genomics to the study of the molecular mechanisms of cardiovascular disease. Work that I am pursuing in this laboratory, and proposed in this application, are directly in line with my personal aspiration to start an independent career in the field of scientific research to work on projects with high translational value and of interest to the public health.

**ORCID iD : 0000-0001-9486-2709**

**B. Positions and Honors
Positions**

2019-2020 Research Fellow at CSIR-Institute of Genomics and Integrative Biology

2016-2019 Senior Research Fellow at CSIR-Institute of Genomics and Integrative Biology

2014-2016 Junior Research Fellow at CSIR-Institute of Genomics and Integrative Biology

2012-2014 Project Fellow at CSIR-Institute of Genomics and Integrative Biology

**Honors**

2019-2020 Intel-India Fellowship

2014-2019 UGC NET-JRF Fellowship

**C. Contributions to Science
1. Identification and Characterization of Circular RNAs.**

I have worked on three different objectives for my Ph.D. that includes:

**1a. Genome-wide transcriptome map of circular RNAs in zebrafish and rat.**

*Danio rerio*, zebrafish, has been a popular model system to understand human diseases. It has been estimated that over 70% of human genes have an orthologue in zebrafish. Apart from the advantages of being small and having transparent embryos in early development, the availability of efficient tools to manipulate the genome and the availability of genome sequences have been well documented. We used a whole genome RNA-seq approach along with computational algorithms to discover 3,428 novel circular RNAs in adult zebrafish. Further in-depth analysis revealed that the circular RNAs were largely derived from known protein-coding loci. A small subset was also derived from lncRNA loci. A significant number of the circular RNAs showed tissue specificity. A subset of newly discovered circular RNAs was independently validated experimentally, and their expression analysis revealed strong concordance with the tissue specificity observed. Our analysis revealed that the human genes orthologous to zebrafish genes producing circular RNAs indeed also generated circular RNAs, indicating an evolutionarily conserved molecular mechanism between the two species. We produced a comprehensive genome-wide map of tissue-specific circular RNAs in zebrafish and uncovered a hitherto unknown repertoire of circular RNAs. We also identified circular RNAs in 9 developmental stages from zebrafish embryos to 5 dpf.

 Rat is one of the most widely used model organisms in medical research and an excellent model for cardiovascular disease, for stroke and hypertension and there are a variety of rat genetic models. The physiology of rat is very similar to humans, and shares a number of gene by environment interactions for a number of complex diseases. We identified the circular RNA map in this species for different tissues and developmental stages.

[**Sharma D**, Sehgal P, Mathew S, Vellarikkal SK, Singh AR, Kapoor S, et al. A genome-wide map of circular RNAs in adult zebrafish. Sci Rep. 2019;9: 3432.](http://paperpile.com/b/nC0pnf/0lX8)PMCID: PMC6401160

[**Sharma D**, Sehgal P, Hariprakash J, Sivasubbu S, Scaria V. Methods for Annotation and Validation of Circular RNAs from RNAseq Data. Methods Mol Biol. 2019;1912: 55–76.](http://paperpile.com/b/nC0pnf/SAns)PMID: 30635890

**1b. Creating a publically available resource for disease associated putative circular RNA biomarkers**

The recent years have seen a number of circular RNAs being identified in organisms from the plant and animal kingdoms, through genome-wide transcriptome analysis. Existing resources compiling circular RNAs lack a spectrum of organisms and coverage of circular RNAs identified, which makes cross-comparisons difficult. Additionally most of the resources do not list candidate circular RNAs which have been independently validated through orthogonal assays nor provide information on assays to validate them. We have created circRNome, a comprehensive compendium of circular RNAs encompassing over 1.6 million circular RNAs from 31 organisms. This includes transcriptome wide suggested divergent primers across circular RNA junctions for 17 organisms and 1118 validated circular RNAs with their primers. To the best of our knowledge, this is the most comprehensive collection of circular RNAs available to date.

 Studies have shown the role of circRNAs in a number of diseases and increasing evidence points to their potential application as biomarkers in these diseases. We have also created a comprehensive manually curated database of circular RNAs associated with different human diseases. This database is available at URL http://clingen.igib.res.in/circad. The Database lists more than 1100 circRNAs associated with 138 diseases and mapping to 95 International Statistical Classification of Diseases (ICD) codes with evidence of association linked to published literature. The database is unique in many ways. Firstly, it provides ready to use primers to work with, in order to use circRNAs as biomarkers or to perform functional studies. It additionally lists the assay and PCR primer details including those experimentally validated as a ready reference to researchers along with fold change and statistical significance.

[Rophina M, **Sharma D**, Poojary M, Scaria V. Circad: a comprehensive manually curated resource of circular RNA associated with diseases. Database . 2020;2020. doi:](http://paperpile.com/b/nC0pnf/IOpF)[10.1093/database/baaa019](http://dx.doi.org/10.1093/database/baaa019)**.**PMID: 32219412

**1c. Role of RNA binding proteins in regulating circular RNA biogenesis**

In order to understand the biogenesis and mechanism of action of circular RNAs, we have taken publically available CLIP-seq data for RNA binding proteins and analysed their binding sites with respect to circular RNA junctions. We have characterized the putative RNA binding proteins using zebrafish, a very good model organism. We shortlisted 4 proteins to undergo validation and functional study based on their peak ratio and their orthology in zebrafish.

**2. Saliva microbiome in autoimmune disorders reveals a distinct set of microbes associated with the disease**

Microbial dysbiosis has been described in many rheumatological diseases. However studies on comparison of the microbiome across different rheumatological diseases are limited. This study systematically evaluated the salivary microbiome in patients affected with three major rheumatological disorders i.e. rheumatoid arthritis (RA), spondyloarthritis (SpA) and primary Sjogren’s Syndrome (pSS) with control healthy individuals using 16S rRNA sequencing approach.

[**Sharma D**, Sandhya P, Vellarikkal SK, Surin AK, Jayarajan R, Verma A, et al. Saliva microbiome in primary Sjögren’s syndrome reveals distinct set of disease-associated microbes. Oral Dis. 2020;26: 295–301.](http://paperpile.com/b/nC0pnf/XSSY)PMID: 31514257

[Sandhya P, Danda D, **Sharma D**, Scaria V. Does the buck stop with the bugs?: an overview of microbial dysbiosis in rheumatoid arthritis. Int J Rheum Dis. 2016;19: 8–20.](http://paperpile.com/b/nC0pnf/Amuk)PMID: 26385261

**3. Genomics for Public Health**

**3a. GUaRDIAN (Genomics for Understanding Rare Disease India Alliance Network)**

GUaRDIAN is a collaborative research programme aimed at understanding the genetic basis and molecular mechanisms underlying rare genetic disorders. We used advanced sequencing technologies, extensive bioinformatics and animal models in this work. Our collaborators include a large number of clinicians and scientists. In this network, I have worked with patients affected with muscular dystrophy to identify causative mutations, copy number variation and repeat expansion.

[GUaRDIAN Consortium, Sridhar Sivasubbu, and Vinod Scaria. 2019. “Genomics of Rare Genetic Diseases-Experiences from India.” Human Genomics 14 (1): 52.](http://paperpile.com/b/unJlsT/qKkZ)PMID: 31554517

**3b. IndiGen (Genomics for Public Health in India)**

In this programme, we sequenced over 1000 genomes across different ethnicities of the Indian population and developed baseline allele frequency for variations in the different population groups.

**Other work:**

[Sabharwal A, **Sharma D**, Vellarikkal SK, Jayarajan R, Verma A, Senthivel V, et al. Organellar transcriptome sequencing reveals mitochondrial localization of nuclear encoded transcripts. Mitochondrion. 2019;46: 59–68.](http://paperpile.com/b/nC0pnf/0tnS)PMID: 29486245

[Bhattacharjee J, Das B, **Sharma D**, Sahay P, Jain K, Mishra A, et al. Autologous NeoHep Derived from Chronic Hepatitis B Virus Patients’ Blood Monocytes by Upregulation of c-MET Signaling. Stem Cells Transl Med. 2017;6: 174–186.](http://paperpile.com/b/nC0pnf/nn7K)PMCID: PMC5442753

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

**N/A**