Table of Contents

[*Projects* 2](#_Toc41557976)

[*1.* *Pre-PhD projects:* 2](#_Toc41557977)

[**1.1.** **B.Sc. (Honours with Biotechnology)** 2](#_Toc41557978)

[***1.1.1.*** ***Drug designing: evaluation of iron chelation through blood in phytic acid treated thalassemia patients*** 2](#_Toc41557979)

[***1.1.2.*** ***Study of ecological diversity in the Sundarbans, especially on kingfishers*** 4](#_Toc41557980)

[**1.2.** **Master’s project (M.Tech. one-year project)** 5](#_Toc41557981)

[***1.2.1.*** ***Identification and evaluation of Intron Length Polymorphism (ILP) marker in Foxtail millet (Setaria italica) and to study evolutionary relationship between species by cross-species transferability*** 5](#_Toc41557982)

[*2.* *Ph.D. projects - In-silico studies of MicroRNA Regulations in higher eu- karyotes from the perspective of Molecular Evolution* 6](#_Toc41557983)

[**2.1.** **Complex-forming proteins escape the robust regulations of miRNA in human** 7](#_Toc41557984)

[**2.2.** **Insights into the miRNA regulations in human disease genes** 8](#_Toc41557985)

[**2.3.** **Explicating the role of old miRNAs in human disease progression** 9](#_Toc41557986)

[*3.* *Post-PhD projects: DNA methylation pattern in tuberculosis* 10](#_Toc41557987)

[**3.1.** **Identification of DNA methylation patterns predisposing for an efficient response to BCG vaccination in healthy BCG-naïve subjects** 11](#_Toc41557988)

[**3.2.** **Advanced image tracking of bacterial biofilms from live cell imaging analysis using MATLAB image processing** 12](#_Toc41557989)

[**3.3.** **Machine Learning algorithm to identify bio-signatures from tuberculosis-exposed and non-exposed individual** 13](#_Toc41557990)

[**3.4.** **Identification of differential methylation patterns in tuberculosis patients, household contacts and healthy participant** 14](#_Toc41557991)

[**3.5.** **Analysis of differential methylation patterns between tuberculosis-exposed and non-exposed individuals using different illumina platforms** 15](#_Toc41557992)

[**3.6.** **Reduced representation of bisulfite sequencing (RRBS) data analysis from tuberculosis-exposed samples in different cell types** 16](#_Toc41557993)

[**3.7.** **RNA sequencing data analysis using genome-wide transcriptome data from tuberculosis-exposed samples in high endemic country** 17](#_Toc41557994)

[**3.8.** **Identification of exosomal-derived microRNAs from Mycobacterium tuberculosis infected cells** 18](#_Toc41557995)

# ***Projects***

*Lists all projects done/doing from undergraduate to postdoctoral degrees.*

# ***Pre-PhD projects:***

## ***B.Sc. (Honours with Biotechnology)***

### ***Drug designing: evaluation of iron chelation through blood in phytic acid treated thalassemia patients***

*Supervisor: Prof. Sudipa Chakraborty, (****2006-01 to 2006-05****)*

*The main objective of the project is to use rice bran as a source of phytic acid for the thalassemia patients. Thalassemia is one of genetic blood disorders characterized by decreased hemoglobin production. In India,* β-*thalassemia is more common. The autosomal recessive disorder has reported insertion of 5 bp in the* β*-globin gene. The oxygen carrying capacity of hemoglobin molecules reduces and hence iron chelation is one of treatment measures for thalassemic patients. The iron overload occurs very rapidly in the patients and the toxicity level reaches high within 2-3 weeks after blood transfusion, even some patients need to get the transfusion 2 times in a week. Therefore, the chelation of iron shortens the number of transfusions for the patients. However, the iron chelation agents for thalassemia patients are costly and it is quite a burden for the patients’ family. Researches indicated that rice bran is a good natural source for the phytic acid which is often considered as one of the well-known references as iron-chelating agents. In this project, we worked as a team of four members whose objectives are collection of patients’ data, extraction of phytic acid from rice bran, identifying the chemical compositions of the rice bran and test the production of phytic acid. As a part of the group, my job was to collect blood samples from patients, extract the phytic acid from rice bran and test the concentration of phytic acid. Handling of lot of patients’ data, extraction and purification of phytic acid from rice bran, measurement of phytic acid concentration using calorimeter and spectrophotometer were part of the project.*

***Key responsibilities:***

* *Handling of patients, aged less than one year to 34 years,*
* *Study the iron chelation technique,*
* *Sample handling and data collection,*
* *Statistical analysis of data,*
* *Colorimetry and spectrophotometry analysis*

***Key achievements***

* *learning new techniques like color estimation from samples,*
* *spectrophotometry analysis of samples*
* *collection of patients’ data*
* *biochemical analysis of phytic acid*

### ***Study of ecological diversity in the Sundarbans, especially on kingfishers***

*Supervisor: Prof. Amit Chakraborty*

***2006-01 to 2006-05***

*Sundarban is the largest mangrove delta of the world and home for more than 150 species of fish, 270 species of birds, 100 species of mammals, reptiles, amphibians, nearly 350 species of trees. The word “sundar” came from the local name of the plant species, Heritiera fomes (****Sundari)****. While the pneumatophores are the main attraction of these delta that made almost impossible for people to enter the area, the area is also the house of great Royal Bengal Tigers. The distinct ecological diversity with rich flora and fauna in the South part of West Bengal, India also inhabited by more than fifty thousand human beings.*

*Kingfisher, the brightly colored birds has a cosmopolitan distribution all over the world and more than 110 species are found. In India, nearly 19 species are observed over the country. My objective was to observe and documented the species of kingfishers found in Sundarbans. With the help of locals and forest services, the area was visited during the spring of 2006 in a group of fifty people. From official forest records and observations made in the forest area, I was successfully identified 13 species of kingfishers in the area. The collection of data, photographs, characteristic details were done with the help of local habitants and forest personnel.*

***Key responsibilities***

* *Data collection from locals and forest personnel*
* *Detail characteristic analysis of kingfishers*
* *Observe and photographed kingfishers*

***Key achievements***

* *Study of ecological niche*
* *Field work*
* *Sample collection and characteristic observation*

## ***Master’s project (M.Tech. one-year project)***

### ***Identification and evaluation of Intron Length Polymorphism (ILP) marker in Foxtail millet (Setaria italica) and to study evolutionary relationship between species by cross-species transferability***

### *Supervisor: Dr. Manoj Prasad, Scientist VI*

### ***(2009-06 to 2010-05)***

*Introns are noncoding sequences in a gene that are transcribed to precursor mRNA but spliced out during mRNA maturation and are abundant in eukaryotic genomes. The availability of codominant molecular markers and saturated genetic linkage maps have been limited in foxtail millet (Setaria italica (L.) P. Beauv.). Here, we describe the development of 98 novel intron length polymorphic (ILP) markers in foxtail millet using sequence information of the model plant rice. A total of 575 nonredundant expressed sequence tag (EST) sequences were obtained, of which 327 and 248 unique sequences were from dehydration- and salinity-stressed suppression subtractive hybridization libraries, respectively. The BLAST analysis of 98 EST sequences suggests a nearly defined function for about 64% of them, and they were grouped into 11 different functional categories. All 98 ILP primer pairs showed a high level of cross-species amplification in two millets and two nonmillets species ranging from 90% to 100%, with a mean of ∼97%. The mean observed heterozygosity and Nei’s average gene diversity 0.016 and 0.171, respectively, established the efficiency of the ILP markers for distinguishing the foxtail millet accessions. Based on 26 ILP markers, a reasonable dendrogram of 45 foxtail millet accessions was constructed, demonstrating the utility of ILP markers in germplasm characterizations and genomic relationships in millets and nonmillets species.*

***Key responsibilities***

* ***Cloning of bacterial genomes***
* ***Microbiology and molecular biology techniques, bacterial culture, cloning, PCR, agarose-gel running,***
* ***Polyacrylamide gel separation, microRNA isolation and identification, sequencing, primer designing***
* ***Looked after of plantlets in germination chamber, field***
* ***Designing of ILP markers***

***Key achievements***

* ***Different wet lab techniques***
* ***How to develop study pipelines***
* ***Publication***

# ***Ph.D. projects - In-silico studies of MicroRNA Regulations in higher eu- karyotes from the perspective of Molecular Evolution***

*Supervisor: Prof. Tapash C.Ghosh*

***(2012-01 to 2017-01)***

*The evolution of miRNA genes, identification of numerous new miRNAs and miRNA-targeted genes, expansion of miRNAs during animal evolution, functions of miRNAs – are the central questions for molecular biologists for the last decade. How these 22-28 nucleotides change the progression of cellular and physiological developments; helps to correlate with the protein complexes or in the advancement of the diseases in case of human. In this thesis, I have studied the association of miRNAs comparing with the protein complexes resided in the interaction network and analyzed their evolutionary rates from the perspective of miRNA regulation. I found that these “tiny” miRNAs weakly regulate the proteins that are present in the protein complexes and they also target the duplicated genes that are essential in human protein complexes to avoid dosage imbalance. The role of miRNAs in human diseases have been well established, so I have taken the opportunity to study which of the human diseases is targeted more likely by the miRNAs, obtaining the cancer related diseases as the most addressed disease class in human disease gene categories. In addition, I have also encountered the question by asking the role of newly emerged and old miRNAs in regulation of human disease genes. Our study suggested that the dominating role of old miRNAs in human disease progressions over new miRNAs. This study also demonstrated the early evolutionary miRNAs are retained in the organisms to do specific functions despite more number of target genes by new miRNAs.*

***Key responsibilities***

* ***Projects handling,***
* ***Pipeline developing***
* ***Manuscript preparation and publications***

***Key achievements***

* ***Linux system administration and maintenance in Workstation, Servers***
* ***Shell scripts, Perl, Python scripting languages***
* ***Molecular phylogenetic analyses***
* ***Statistical softwares handling***

## ***Complex-forming proteins escape the robust regulations of miRNA in human***

*Most of the proteins carry out their functions by participating in protein complexes. Recently, miRNAs are identified as promising post-transcriptional regulators that influence a large proportion of genes in higher eukaryotes. We aim to understand the role of miRNA in the regulation of human proteins that are present in protein complexes. Here, we show that robust regulations of miRNA are absent in human complex forming proteins. Moreover, the numbers of miRNA hits cannot direct the evolutionary fate of complex-forming proteins independently. However, the duplicated complex-forming proteins having a severe effect on organismal fitness are profoundly targeted by miRNA, probably for reducing the chances of dosage imbalance.*

***Key responsibilities***

* *Design analysis pipeline*
* *ENSEMBL, UCSC database handling*
* *Evolutionary rate calculations*
* *Analysis of gene duplicates*
* *Data analysis using Linux shell and in-house Perl scripts*
* *Statistical analysis through SPSS, GraphPad*
* *Wrote manuscript*

***Key achievements***

* *Learn Perl scripting*
* *Learn Linux terminal and shell scripting*
* *Orthologous, homologous gene data analysis*
* *Evolutionary rate calculation from gene sets*
* *Learn statistical analysis in SPSS, STATA, Tanagra*
* *API handling of databases*
* *Microarray gene expression data handling*
* *Published research article*

***Softwares, Databases and Statistics***

*Softwares: Linux OS, BLASTP, Perl, Shell*

*Databases: CORUM, TargetScan, miRwalk, BioGPS, Ensembl, OGEE,*

*Statistics: Spearman rank correlation, Mann-Whitney U test, Partial correlation*

## ***Insights into the miRNA regulations in human disease genes***

*MicroRNAs are a class of short non-coding RNAs derived from either cellular or viral transcripts that act post-transcriptionally to regulate mRNA stability and translation. In recent days, increasing numbers of miRNAs have been shown to be involved in the development and progression of a variety of diseases. We, therefore, intend to enumerate miRNA targets in several known disease classes to explore the degree of miRNA regulations on them which is unexplored till date. Here, we noticed that miRNA hits in cancer genes are remarkably higher than other diseases in human. Our observation suggests that UTRs and the transcript length of cancer related genes have a significant contribution in higher susceptibility to miRNA regulation. Moreover, gene duplication, mRNA stability, AREScores and evolutionary rate were likely to have implications for more miRNA targeting on cancer genes. Consequently, the regression analysis has confirmed that the AREScores plays most important role in detecting miRNA targets on disease genes. Interestingly, we observed that epigenetic modifications like CpG methylation and histone modification are less effective than miRNA regulations in controlling the gene expression of cancer genes. The intrinsic properties of cancer genes studied here, for higher miRNA targeting will enhance the knowledge on cancer gene regulation.*

***Key responsibilities***

* *Design study pipeline*
* *Data collection and analysis through Perl scripting*
* *Statistical analysis of data using SPSS, GraphPad, Origin*
* *Manuscript preparation*

***Key achievements***

* *Deep learning of Perl scripts*
* *Learning Python script*
* *Intensive use of shell programming in Linux*
* *Handling of clusters, servers, workstations*
* *Linux system administration*
* *Handling of Next-Generation Sequencing data*
* *Published research article*

***Softwares, Databases and Statistics***

*Softwares: Linux OS, BLASTP, Perl, Shell*

*Databases: ICGC, HGMD, GAD, TargetScan, TarBase, UCSC, Ensembl, BioGPS, AREScores, HHMD, NGSmethDB, CancerMiner*

*Statistics: Spearman rank correlation, Mann-Whitney U test, Partial correlation****,*** *Shapiro-Wilk test, McCullum proportion test*

## ***Explicating the role of old miRNAs in human disease progression***

Most of the microRNAs (miRNAs) are mainly described as conserved in the evolutionary timescales, although some miRNAs are species-specific in nature. The conserved “old” miRNAs and “newly” emerged miRNAs may have played some differential roles in human disease progression. However, modulating the level of expression of human disease genes by new and old miRNAs has not been testified yet. Here, we identified 1,190 new and 97 old miRNAs in humans and showed that these old miRNAs play more influential roles in human disease gene expression levels than the new ones. Additionally, we also performed our analyses on four other mammals to find whether a similar pattern of miRNA regulation is followed there as that showed by old miRNAs in humans. Furthermore, our analysis indicated that miRNAs targeting is dominant over other genomic regulators in controlling human disease gene expression levels. Our study demonstrated that the early evolutionary miRNAs are retained in the organisms to do some specific functions despite of the fact that a greater number of genes are being targeted by new miRNAs.

***Key responsibilities***

* *Designed the study pipelines and analyzed the data*
* *Molecular phylogenetic analysis using PHYLIP, MEGA6, UPGMA*
* *Single Nucleotide Polymorphism (SNP), gene expression levels, gene regulation analysis*
* *DNA methylation, transcription factor binding sites (TFBS), histone modifications analysis along with miRNA dataset*
* *Gene orthologous, homologous data analysis and calculation of evolutionary rates*
* *Prediction analysis of microRNA birth and death*

***Key achievements***

* *Learn phylogenetic analysis through different species*
* *In-depth scripting analysis,*
* *Learn R as a scripting language*
* *Intensive high-dimensional statistical data analysis*
* *Prepared manuscript*

***Softwares, Databases and Statistics***

*Softwares: Linux OS, BLASTP, Perl, Shell, MEGA6, Phylopat 52, COUNT,*

*Databases: miRbase, TargetScan, miRDB, miRNAMap2, MicroCosm, HGMD, GAD, GeneCards, PolymiRTS, mircancer, UCSC, TRANSFAC, HHMD, Ensembl,*

*Statistics: Spearman rank correlation, Mann-Whitney U test, Partial correlation****,*** *Shapiro-Wilk test, McCullum proportion test, Regression analysis*

# ***Post-PhD projects: DNA methylation pattern in tuberculosis, host-pathogen interactions***

*Supervisor: Prof. Maria Lerm*

***2017-06 to current***

## ***Identification of DNA methylation patterns predisposing for an efficient response to BCG vaccination in healthy BCG-naïve subjects***

*The protection against tuberculosis induced by the Bacille Calmette Guérin (BCG) vaccine is unpredictable. In our previous study, altered DNA methylation pattern in peripheral blood mononuclear cells (PBMCs) in response to BCG was observed in a subgroup of individuals, whose macrophages killed mycobacteria effectively (‘responders’). These macrophages also showed production of Interleukin-1β (IL-1β) in response to mycobacterial stimuli before vaccination. Here, we hypothesized that the propensity to respond to the BCG vaccine is reflected in the DNA methylome. We mapped the differentially methylated genes (DMGs) in PBMCs isolated from responders/non-responders at the time point before vaccination aiming to identify possible predictors of BCG responsiveness. We identified 43 DMGs and subsequent bioinformatic analyses showed that these were enriched for actin-modulating pathways, predicting differences in phagocytosis. This could be validated by experiments showing that phagocytosis of mycobacteria, which is an event preceding mycobacteria-induced IL-1β production, was strongly correlated with the DMG pattern.*

***Key responsibilities***

* *Computational pipeline development to Illumina 450K methylation data analysis*
* *Gene-gene interaction network analysis*
* *High-throughput statistical data analysis with Multiple Factorial analysis (MFA)*
* *Manuscript preparation*

***Key achievements***

* *Extensive use of R and Bioconductor packages*
* *Gene and Protein Interaction network analysis*
* *Performing factorial analysis*
* *Collaboration with different departments and faculty members*

## ***Advanced image tracking of bacterial biofilms from live cell imaging analysis using MATLAB image processing***

## ***Machine Learning algorithm to identify bio-signatures from tuberculosis-exposed and non-exposed individual***

## ***Identification of differential methylation patterns in tuberculosis patients, household contacts and healthy participant***

## ***Analysis of differential methylation patterns between tuberculosis-exposed and non-exposed individuals using different illumina platforms***

## ***Reduced representation of bisulfite sequencing (RRBS) data analysis from tuberculosis-exposed samples in different cell types***

## ***RNA sequencing data analysis using genome-wide transcriptome data from tuberculosis-exposed samples in high endemic country***

## ***Identification of exosomal-derived microRNAs from Mycobacterium tuberculosis infected cells***

# ***Collaborative projects***

## ***Exploring the major cross-talking edges of competitive endogenous RNA networks in human Chronic and Acute Myeloid Leukemia***

*Collaborator: Dr. Kamalika Sen, Bose Institute, Kolkata*

***2017–01 to 2017–05***

***Published one article together***

## ***Evaluating the role of microRNAs in psoriasis infected, non-infected and healthy controls from genome-wide methylation analysis***

*Collaborator: Dr. Deepti Verma, Linköping University, Sweden*

***2019-01 to current***

***Manuscript in preparation***

## ***Identifying the relative effects of nanomaterials in different cell types***

*Collaborators: Dr. Mohammod Azharuddin, Linköping University, Sweden*

*Dr. Hirak Patra, Cambridge University, UK*

***2020-04 to current***

***One manuscript in preparation***