**PROJECT 3**

**BUDGET JUSTIFICATION**

**Key Personnel**

Mark Gerstein, Ph.D. (Project Leader). Dr. Gerstein is the Albert Williams Professor of Biomedical Informatics. His lab was one of the first to perform integrated data mining on functional genomics data and to do genome-wide surveys. His tools for analyzing motions and packing are widely used. Most recently, he has designed and developed a wide array of databases and computational tools to mine genome data in humans, as well as in many other organisms. He has worked extensively in the 1000 genomes project in the SV and FIG groups. He also worked in the ENCODE pilot project and currently works extensively in the ENCODE and modENCODE production projects. He is also a PI in DOE KBase and the leader of the Data Analysis Center for the NIH exRNA consortium. In these roles Dr. Gerstein has designed and developed a wide array of databases and computational tools to mine genomic data in humans as well as in many other organisms. He will lead the overall informatics effort in the project. Dr. Gerstein’s effort on this grant will be 0.75 summer months, annually**.??**

Smita Krishnaswamy, Ph.D., (Co-Investigator). Dr. Krishnaswamy is an Assistant Professor in the Department of Genetics at Yale University, School of Medicine. She will oversee the analysis of single-cell CyTOF and sequencing data generated in the project, including the development and application of novel computational tools, statistical analysis, and interpretation and visualization of results. Dr. Krishnaswamy’s effort will be 1.2 calendar months annually.

**Non-Key Personnel**

Daifeng Wang, Ph.D., is an Associate Research Scientist in the Department of MMBB at Yale University, school of Medicine. He has a strong background in scientific computation and obtained his Ph.D. in The University of Texas at Austin. His background is data modeling, mining and analysis. He used to work on dynamic pattern recognition of large-scale gene expression data. After joining the Gerstein lab, he has been actively working on developing novel algorithms/models to cluster multiplex networks in genomes, identify dynamic patterns in high dimensional genomic datasets, and characterize combinatorial co-operations among regulatory factors in gene regulatory networks. He is also interested in analyzing dynamics of social networks in large scientific projects. Dr. Wang will work on the deconvolution of cell-type signatures and the logical modeling of regulatory pathways (Aim 3). Dr. Wang’s effort on this grant will be 12 calendar months annually.

Shaoke Lou, Ph.D., Postdoctoral Assoc. Dr. Lou has multidisciplinary background with a B.S. in Chemistry, an M.S. in Biochemistry and Molecular Biology, and a PhD in Computer Science and Engineering. He earned his PhD from the Chinese University of Hong Kong in 2012. During his PhD study, he mainly works on the development of read-mapping algorithm, next-generation sequencing data analysis, data mining and statistics modeling. He has broad research interests, and is an expert in gene regulatory networks, disease association study and genomic variation prioritization. He was involved in the investigation and data analysis in model organisms: angiogenesis activities of traditional Chinese medicine in zebrafish, regulation of chromatin structure factors (CapH2) in fly, and circadian rhythmic study in mouse. Dr. Lou will work on the annotation of novel asthma-related genetic elements by clustering and network analysis (Aim 1). Dr. Lou’s effort ton this grant will be 4.0 calendar months in Year 1, then 6.0 calendar months in Years 2-5).

To Be Named (Postdoctoral Associate). The postdoc in the Krishnaswamy lab will work on development of algorithms, programming, automating statistical analysis and development of graphical software interfaces. His or her annual effort on this grant will be 4.5 calendar months.

**Other Expenses**

Laboratory Supplies

We are budgeting an incremental amount of supplies for the individuals named above. This supplies budget will be used to cover computer supplies for them, covering such expenses as: diskettes, tapes, and other miscellaneous computer parts (e.g. replacing worn out surge suppressors), software upgrades, web hosting and "cloud computing fees, and reprint charges. These items are needed to complete the proposed research and will solely benefit this project.

The supplies budget for each year is:

Year 1: $10,915

Year 2: $25,273

Year 3: $20,631

Year 4: $15,850

Year 5: $10,926

Travel

Here we are requesting incremental funds for each of the FTEs for airfare, lodging and meal expenses to attend scientific meetings annually that benefit the project. In particular, the travel will include at least 1 trip per year to a scientific meeting related to asthma, such as the American Academy of Allergy, Asthma and Immunology annual meeting.

Equipment

In Year 1, $20,000 is budgeted for a Dell Poweredge R815 server with 160GB of memory and four AMD Opteron processors. This will be used for processing our data and digital visualization. This is needed to complete the proposed research and will solely benefit this project.