National Centre for Bioinformatics

Quaid-i-Azam University Islamabad

http://ncb.qau.edu.pk

Bioinformatics derives knowledge from computer analysis of biological data. These can consist of the information stored in the genetic code, but also experimental results from various sources, patient statistics, and scientific literature. Research in bioinformatics includes method development for storage, retrieval, and analysis of the data. Bioinformatics is a rapidly developing branch of biology and is highly interdisciplinary, using techniques and concepts from informatics, statistics, mathematics, chemistry, biochemistry, physics, and molecular biology. It has many practical applications in different areas of biology and medicine. The use of computational methods in biomedicine deals with the analysis, storage, manipulation and interpretation of macromolecules such as DNA, RNA and proteins. Conversely, wet lab analysis of computationally predicted functionally relevant motifs/segments of macromolecules further enhances our understanding of complex mechanisms occurring in cells.

On the basis of significance of this discipline, NCB was established as a faculty affiliated research Centre in 2008 through funding by Higher Education Commission. It is housed in 53000 square feet building comprising multiple research and computational labs. Besides, it integrates Department of Computer Sciences, other Departments of Biological Sciences and Department of Mathematics.

NCB is offering M.Phil and Ph.D programs in the area of Bioinformatics.

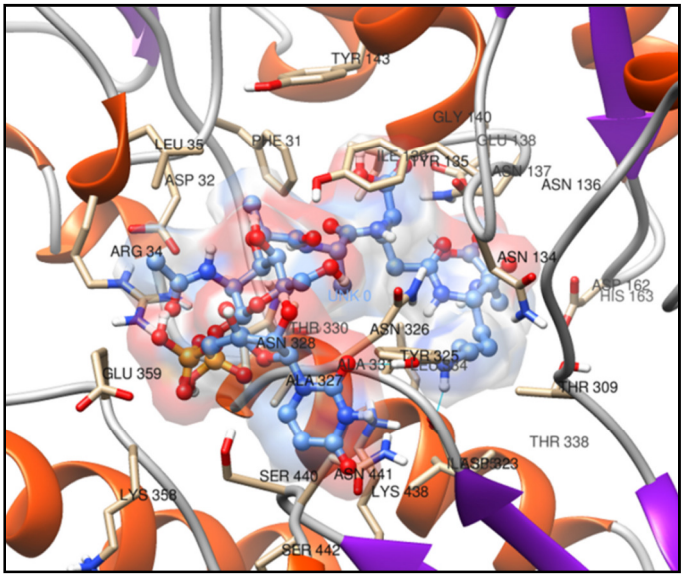
Major objectives

- To promote advance studies and research in Bioinformatics and allied fields such as Genomics and Proteomics.
- To Train young scientists for research and help in developing a high-profile scientific and technical talent in Universities, R&D organizations and industries.
- To undertake joint investigation and collaborative research in multiple areas of Bioinformatics.
- To promote cooperation in inter-disciplinary relationship with teaching and research institutions in Pakistan.
- To arrange conferences, seminars and refresher courses for researchers from Pakistan, OIC member states and from other developing countries in the region.

MOLECULAR DYNAMICS (MD)

The focus of the research projects carried out by Computational Biology Laboratory is drug designing which incorporates the comparative homology modeling, docking and molecular dynamic simulation studies of clinically relevant proteins. The organism source for the studied biomolecules ranges from infectious agents such as bacteria and viruses to human proteins that are responsible for diseased state. Owing to the problem of multi-drug resistance and its effect on the health-care system of Pakistan and other countries around the globe, genome based subtractive channel analysis is a recent addition to the research strategies.

During 2012-13, various aspects of medicinally significant proteins and enzymes were elucidated through the collective efforts of lab members. The proteins/enzymes from the work published in this duration include alpha-glucosidase, methyltransferase, glucosaminyltransferase, HCV NS5B-3a RNA polymerase and human paraoxonase 1. The dimensions of the recent research work encompass comparative modeling, molecular docking, molecular dynamics simulation, glycation-induced changes in the enzyme, structural and mechanistic elucidations of potent novel inhibitors and comparison of widely used molecular docking tools. Designing of advance tools for cutting-edge analysis of MD trajectories is yet another aspect that is being actively pursued. Research grants for different projects were approved by HEC. Acquisition of high performance computing cluster is a recent accomplishment in terms of up-gradation of the existing computational facilities at the lab. International collaborations are well established and two PhD students have been approved for training under IRSIP program.



There is an ample aptitude of linking the research with industry:

- Research pertaining to drug designing holds promising applications in pharmaceutical industries.
- Provision of structural insights into biomolecules such as active site recognition, protein folds assessment through *in silico* studies complements biotechnological research.
- In collaboration with software development industries, we can design algorithms specific to biological research including advanced molecular dynamics analysis tools.
- The tools currently being developed by the lab for advanced MD analysis can provide higher level of theoretical computation and can be linked with software industry.

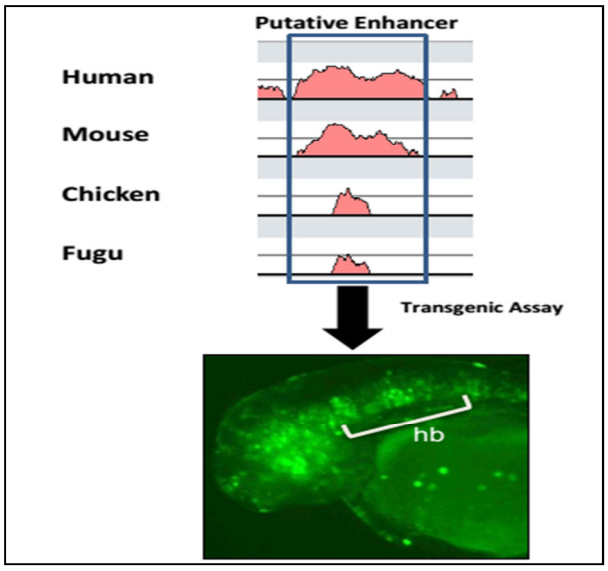
EVOLUTIONARY AND COMPARATIVE GENOMICS

Our Group is long fascinated in several areas of research landscapes within the scope of Evolutionary and Functional Genomics, Ancient DNA, Molecular Anthropology, Evolutionary Developmental Biology & Genome Bioinformatics. Our computational methodological work is tight in with experimental research and external collaborations, where we study the variability of molecular traits in different systems, including Zebrafish and transgenic mice models. We unite graduate research students from various backgrounds (Natural & Computer Sciences) that are completely being trained to make their mark on the world with Bioinformatics.

RESEARCH INTERESTS

Genomic analysis of human ancient enhancer elements

Mechanism governing the gene expression play a pivotal role in the morphological and anatomical variations across vertebrate lineages. Precise spatio-temporal gene expression monitored by cis-acting regulatory repertoire appears to be designed to perform its modular functions in association with target genes in tissue-specific cell lines. Currently, we're employing the combination of computational and functional assays to segregate the evolutionarily older and new components of vertebrate ancient cis-regulatory modules. For this purpose we selected the *in vivo* model Zebrafish to test the functionality of human-fish conserved putative cis-regulatory modules and to dissect their evolutionary trajectories.



Plot of non-coding sequence conservation between mammals and fish. Transient GFP expression in the primary neurons of hindbrain after the injection of a Human a CNE (Conserved Non Element) into zebrafish embryo by our group.

Genomic basis of human biological uniqueness

Human is markedly different from other non-human primates by its inimitable morphological, anatomical, physiological and behavioral eccentrics, including large brain, reproductive system, higher cognitive prowess and practical hairlessness. For genetic perspective these peculiarities emerge by entail concomitant alteration in genes and non-coding regulatory elements; however the exact genetic basis of these peculiarities is remains surreptitious. The decoding of human, archaic humans and other primates genomes was heralded as an opportunity for our lab to truly comprehend how and what genetic underpinnings resulted in the evolution of human unique eccentrics. The stunning revelation of our group has recently analyzed the genetic basis of human hairlessness and has pinpointed the genetic basis of reducing hair cover during the recent history of human evolution.

Evolution of vertebrate gene families and genomes

Another enthralling area for our lab is to explore those events that diversified the human gene sets, and shaped vertebrate genome architecture deep in history (>450 Mya). Unrevealing these events is a key to unfold genomic basis of major morphological transitions that vertebrates accomplished during their history. Currently our group uses the combination of phylogenetic approaches and inter-genomic/intra-genomic synteny comparisons to shed important insight into ancient vertebrate genome shaping events.

Industrial medicinal research implications:

- Using positive selection to identify the pathogenic mechanisms of various infections (i.e. HIV, HCV, HBV, etc.) in humans.
- Human evolutionary genetics has its implications to pin down the differences between human genomes and form the basis of anthropological, forensic research.
- Our research has bright prospective in highlighting the geographical distribution of genetic variations to pin down the clinical patterns of disease variants.
- Tracking epidemics in human populations.
- Prenatal diagnosis and Cytogenetic analysis through Fluorescent *in situ* hybridization (FISH).

FUNCTIONAL INFORMATICS (FI)

Integrates various Bioinformatics approaches with informational systems which provide vast improvements to the overall development and research in the areas of drug designing, target isolation, enzyme-inhibitor screening and protein interactions involved in signaling pathways (Figure 1). The general schema of technology integration in Biopharmaceutical endeavor is employed in computational drug development through designing of multiple structural and applied proteomics tools. For example, MotViz tool is used for the prediction of structural motifs and identification of conserved sequences to understand the functional implications of proteins. FI lab is well-equipped in advanced methodologies including 2D gel electrophoresis, BN-PAGE, and Chromatin immunoprecipitation assays which aid in finding novel protein complexes.

The areas of structural and functional proteomics are well-established in FI lab, which may boost up Pharmaceutical industry and biotechnology. On the basis of these distinctions, structure and ligand-based drug designing is carried out by molecular modeling, docking, Virtual screening (VS) and high-throughput screening (HTS) assays to characterize macromolecular binding sites (Figure 2). HTS is the most common experimental method used to identify lead compounds.

By developing close interfaces between computational and experimental screening, new target protein identification processes will be streamline, which could be used for the synthesis of new pharmaceuticals. FI lab enumerates the predicted drug response on mice models for cancer and infertility disorder to addresses their prognostic and therapeutic utilities which would prove to be a solid platform for preclinical target validation and experimental therapeutics.

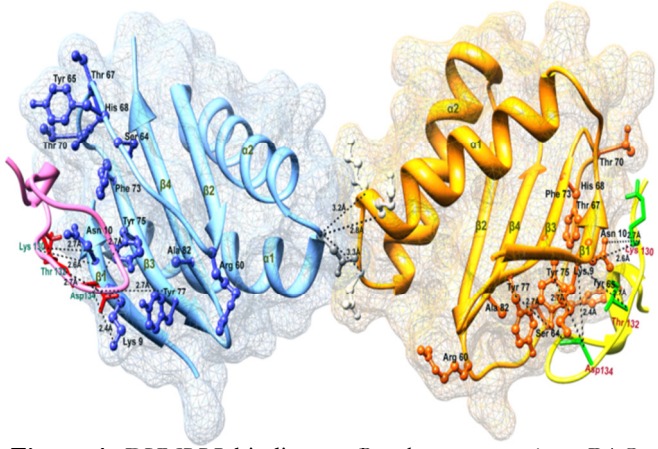


Figure 1: DYNLLI binding to *Pseudomonas aeruginosa* PAO peptide (adapted from Kausar et al., 2013).

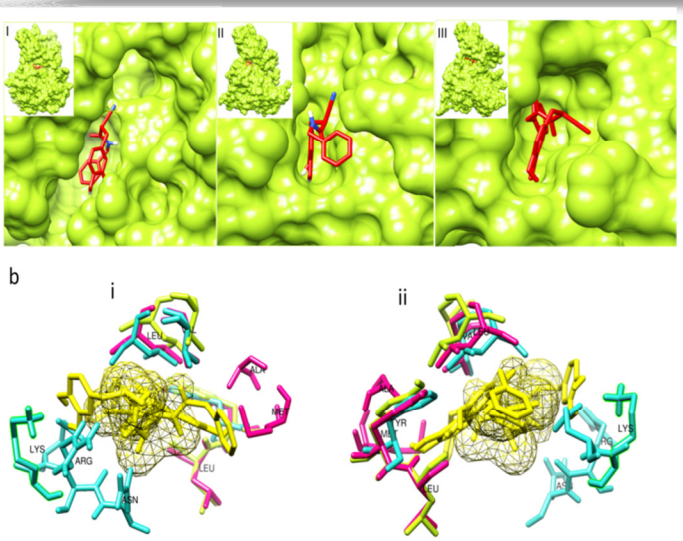


Figure 2: Binding analysis of inhibitor to multiple target proteins.

Industrial Implications

- Designing of novel tools and algorithms will help in debt generation in the field of software industry.
- Exploration of novel drugs, screening of previously known drugs and their comparative analysis for understudying mechanism would strengthen the areas of drug designing to boost up health industry.
- Generation of knockout and transgenic mice models can be commercialized for earning.

STATISTICAL BIOINFORMATICS

Statistical methods are crucial in many bioinformatics problems and applications. Especially the analysis of large scale omics data requires not only massive computation and efficient algorithms, but also relies on solid statistical methodology. Statistical Bioinformatics Lab is enthusiastically involved in developing statistical methods to understand the genetic variation such as common-variant and rare-variant. Detection of rare variant is one of the challenges that SB lab is dealing with. The boom of next generation sequencing (NGS) technology and its applications to a wide range of biomedical fields has brought about many computational and statistical challenges. SB lab is also dealing with these challenges such as analyzing RNA-seq, Chip-Seq data. Estimation of missing value is one of the core research interests of SB lab and it is not limited to, in the area of Applied Spatial Statistics, Analysis of GWAS and Microarray Data and their Meta-Analysis.

- Innovative mathematical modeling to understand genetic variation of viral diseases
- Statistical challenges in Next Generation Sequencing (NGS) Data Analysis
- To develop robust approach for estimating missing values
- Applied Spatial Statistics

LIST OF SELECTED PUBLICATIONS

1. Azam, S.S. and Abbasi, S.W. (2013) **Molecular docking studies for the identification of novel melatoninergic inhibitors for acetylserotonin-O-methyltransferase using different docking routines**, *Theoretical Biology and Medical Modelling*, 10, 675-684. **IF: 1.46**.
2. Azam, S.S., Abbasi, S.W., and Batool, M. (2013) **Structure modeling and docking study of HCV NS5B-3A RNA polymerase for the identification of potent inhibitors**, *Medicinal Chemistry Research*, 22, 1-10. **IF: 1.61**
3. Saleem, A., Azam, S.S., and Zarina, S. (2012) **Docking and Molecular Dynamics Simulation Studies on Glycation Induced Conformational Changes of Human Paraoxonase 1**, *European Biophysics Journal*, 41, 241-248. **IF: 2.27**
4. Azam, S.S., Lim, L.H.V., Hofer, T.S., Randolf, B.R., and Rode, B.M. (2010) **Hydrated Germanium (II): Irregular Structural and Dynamical Properties Revealed by a Quantum Mechanical Charge Field Molecular Dynamics Study**, *Journal of Computational Chemistry*, 31, 278-285. **IF: 3.83**
5. Azam, S.S., Hofer, T.S., Bhattacharjee, A., Lim, L.H.V., Pribil, A.B., Randolf, B.R., and Rode, B.M. (2009) **Beryllium(II)-- the Strongest Structure Forming Ion in Water? A QMCF MD Simulation Study**, *The Journal of Physical Chemistry B*, 113, 9289-9295. **IF: 3.6**
6. Abbasi, A.A., Minhas, R., Schmidt, A., Koch, S., and Grzeschik, K.H. (2013) **Cis-regulatory underpinnings of human GLI3 expression in embryonic craniofacial structures and internal organs**,. *Development, growth & differentiation*, 55, 699-709. **IF: 2.4**
7. Parveen, N., Masood, A., Iftikhar, N., Minhas, B., Minhas, R., Nawaz, U., and Abbasi, A.A. (2013) **Comparative genomics using teleost fish helps to systematically identify target gene bodies of functionally defined human enhancers**, *BMC Genomics*, 14, 122. **IF: 4.4**
8. Asrar, Z., Haq, F., and Abbasi, A.A. (2013) **Fourfold paralogy regions on human HOX-bearing chromosomes: role of ancient segmental duplications in the evolution of vertebrate genome**, *Molecular phylogenetics and evolution*, 66, 737-747. **IF: 4**
9. Abbasi, A.A. (2011) **Molecular evolution of HR, a gene that regulates the postnatal cycle of the hair follicle**, *Scientific reports*, 1. **IF: 2.93**
10. Abbasi, A.A. (2010) **Piecemeal or big bangs: correlating the vertebrate evolution with proposed models of gene expansion events.** *Nature reviews Genetics*, 11, 166. **IF: 41.06**
11. Faisal, M., Futschik, A., and Hussain, I. (2013). **A New Approach to Choose Acceptance Cutoff for Approximate Bayesian Computation**, *Journal of Applied Statistics*; 40, 862-869. **IF: 0.449**
12. Hussain I., Kazianka H., Pilz J., and Faisal M. (2013). **Spatio-Temporal Modeling of Particulate Matter Concentrations Including Covariates**, *Sci.Int.(Lahore)*, 25, ,15-21.
13. Hussain, I., Spoeck, G., Pilz J., Faisal, M., and Yu H. (2012). **Spatio-Temporal Interpolation of Precipitation including Covariates: During Monsoon Periods in Pakistan**, Pakistan Journal of Statistics,Vol. 28, No. 3, pp:351-365. **IF: 0.252**

14. Kausar, S., Asif, M., Bibi, N., and Rashid, S. (2013) **Comparative molecular docking analysis of cytoplasmic dynein light chain DYNLL1 with Pilin to explore the molecular mechanism of pathogenesis caused by Pseudomonas aeruginosa PAO**, *PLoS ONE*, 8, e76730**. IF: 3.73**
15. Bibi, N., Parveen, Z., and Rashid, S. (2013) **Identification of Potential Plk1 Targets in a Cell-Cycle Specific Proteome through Structural Dynamics of Kinase and Polo Box-Mediated Interactions**, *PLoS ONE*, 8, e70843. **IF: 3.73**
16. Rashid, S., Parveen, Z., Ferdous, S. and Bibi, N. (2013) **Mutually exclusive binding of APPL(PH) to BAR domain and Reptin regulates β-catenin dependent transcriptional events**, *Computational Biology and Chemistry*; 47C, 48-55. **IF: 1.793**
17. Ain, Q., Seemab, U., Rashid, S., Nawaz, M.S., and Kamal, M.A. (2013) **Prediction of Structure of Human WNT-CRD (FZD) Complex for Computational Drug Repurposing.** *PLoS ONE*; 8, e54630. **IF: 3.73**
18. Rashid, S., Pilecka, I., Torun, A., Olchowik, M., Bielinska, B., and Miaczynska, M: (2009) **Endosomal adaptor proteins APPL1 and APPL2 are novel activators of beta-catenin/TCF-mediated transcription**, *J Biol Chem*; 284, 18115-18128. **IF: 4.651**
19. Batool, S., Ferdous, S., Kamal, M.A., Iftikhar, H., and Rashid, S. (2013) **In silico Screening for Identification of Novel Aurora Kinase Inhibitors by Molecular Docking, Dynamics Simulations and Ligand-Based Hypothesis Approaches**. *Enzyme Engineering*; 2, 2.

National Centre for Bioinformatics
Quaid-i-Azam University Islamabad
45320, Pakistan
ncb.qau.edu.pk
Email:bif@qau.edu.pk
Tel: +91-51-90644066