

# Application of Bioinformatics in the Field of Cancer Research

Dr. (Mrs.) Padma R. Chavan

---

Biological data include extensive information regarding genomic sequences of different species, changes due to evolution, and changes in their protein sequences, which in turn contributes to variations in their specificity and reactivity. Such a massive data cannot be handled with ease. This requires systematic sieving of the data to categorize and catalogue them. Based on this need arose the field of Bioinformatics.

Bioinformatics is a discipline, which encompasses branches like biology, computer science, IT & mathematics. It is science of managing & analysing vast biological data using advanced computing techniques.

With the advent of Bioinformatics the basic concept of Central Dogma has changed from traditional: Gene----- mRNA----- Protein to newer version of Genome----- Transcriptome----- Proteome which has direct application to bench research work.

The primary goal of bioinformatics is to increase our understanding of biological processes. Major research efforts in this field include sequence alignment, gene finding, genome assembly, protein structure alignment, protein structure prediction, prediction of gene expression and protein-protein interactions, and the modeling of evolution.

## Application of Bioinformatics tool in molecular biology research

Specifically, this includes: (1) Searching for homologous sequences for DNA hybridization techniques (DOT BLOT, FISH) (2) For identification of restriction sites for recombinant DNA techniques([http://tools-neb.com/NEB\\_cutter2/](http://tools-neb.com/NEB_cutter2/)) (3) For identification of splice sites for transcriptome studies (4) to understand ORF(<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>), coding regions for design of primers (5) to study promoter regions for binding studies (6) to locate flanking regions for promoter area (7) to design primers for PCR , chimeric DNA(8) In gene cloning techniques - to compare the sequences of cloned gene with existing database (Gene bank) to identify which family it belong to (9) to locate the gene on the chromosome (NCBI - Mapviewer)

## Medical applications of Bioinformatics:

(1) In understanding genetic diseases like cystic fibrosis, sickle cell anemia (SNPs) (2) for gene therapy (3) In pharmacogenomics to personalize drugs for better bioavailability, to design new and better drugs, to develop better drug delivery system .

## Application of Bioinformatics in cancer research:

(1)To study change in tissue specific protein expression in cancerous and non cancerous tissue,(2) change in immune response of the proteins in cancerous conditions,(3) to conduct translational research in different organ confined carcinoma: translating results from the laboratory bench to bed side delivery of patient care & communicating the lessons learned back to the bench,(4) to develop an ANN (artificial neural network) to distinguish among members of a family of childhood tumors that include neuroblastoma rhabdomyosarcoma, non Hodgkins lymphoma etc,(5) to develop and implement algorithms that help in differential diagnosis of organ confined tumor.

In the post-genome era, efforts are focused on biomarker discovery and the early diagnosis of cancer through the application of various omics technologies - transcriptomics, proteomics, metabonomics, peptidomics, glycomics, phosphoproteomics or lipidomics - on tissue samples and body fluids. No matter which omics technology is used in biomarker development, bioinformatics tools are required to extract the diagnostic or prognostic information from the complex data.

Based on pattern recognition technologies, discriminatory patterns (a panel of gene, protein or peptide patterns) can be identified for the diagnosis of persons with and without cancer.

Benchtop oncology supplies Desktop oncology with large amounts of omics data produced by high-throughput technology. Desktop oncology establishes knowledge on cancer-related biomarkers, such as predisposition markers, diagnostic markers, prognostic markers, and therapeutic markers, by using bioinformatics and human intelligence of experts for data mining and text mining.

Biomarker monitoring contributes to therapeutic optional choice and drug dosage determination for cancer patients. Knowledge on biomarkers is feed forwarded from desktop to bedside in the translational research, and then biomarker monitoring is feedbacked from bedside to desktop in the reverse translational research.

Bedside oncology applies the knowledge established by Desktop oncology to determine therapeutics for cancer patients. Antibody drugs, small molecule inhibitors, conventional cytotoxic drugs, and anti-hormonal drugs are used for cancer chemotherapy.

Desktop oncology is indispensable for cancer research in the post-genome era. Combination of genetic screening for cancer predisposition in the general population and precise selection of therapeutic options during cancer management could contribute to the realization of personalized prevention and to dramatically improve the prognosis of cancer patients in the future.

Bioinformatics tools also assist in extraction of SNP data from public resources, approaches for SNP discovery by re-sequencing, software tools for haplotype inference and optimal SNP selection for genotyping, platform approaches for SNP genotyping, and the analyses of these datasets for genotype-phenotype study.

### **Single nucleotide polymorphism in androgen regulated gene in prostate cancer**

Prostate cancer is one of the most common diseases among men over 50 years of age. Androgen plays a major role in proliferation and carcinogenesis of prostatic glandular cells. The principal metabolite is dihydrotestosterone produced by 5 $\alpha$  reductase from testosterone within prostate epithelial cells. The dihydrotestosterone forms a complex with androgen receptor which binds with the androgen responsive elements in promoter region of the target gene. One such candidate gene is Prostate specific antigen which codes for a serine protease glycoprotein, the PSA, expressed at high levels in prostate glandular epithelium. PSA is widely used as a marker of prostate cancer occurrence, progression and response to treatment. Recent studies have indicated possible association between polymorphism in PSA gene at -158 G/A and prostate cancer, especially in the Caucasian population. The aim of the present study was to evaluate the presence of SNP (G/A) in the promoter region of PSA gene and to explore its association with clinical findings in Indian men with prostate cancer. The study included 30 cases of prostate cancer patients which were clinically, biochemically assessed and histologically confirmed and 48 age as well as socio economically matched normal healthy controls.

The polymorphism was determined by PCR - RFLP method using DNA from peripheral blood. A standard nucleotide sequence of human PSA gene was obtained from NCBI database ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). The polymorphism to be studied is reported at the position -158 G/A from transcription start site. To analyze the Single nucleotide polymorphism, polymerase chain reaction was standardized with the primers designed using Primer 3 software. The precision of primers designed was assessed using the NCBI BLAST software. A 300 bp PSA gene fragment was amplified using the standard PCR technique. The restriction enzyme required for the digestion of PCR product at a particular site was predicted using NEB Cutter V 2.0 software. The NheI restriction enzyme used, digested the PCR product (G CTAGCTC) into three different patterns 1) Genotype AA of 300 bp fragment with no restriction site on either of the allele 2) Genotype AG of 150 and 300 bp with restriction site only on one allele 3) Genotype GG OF 150 bp with restriction sites on both the alleles. The results obtained were further analyzed for the association studies in prostate cancer.