

Role of Computers in Bioinformatics by Using Different Biological Datasets

Chukka Santhaiah¹, Dr.A. Rama Mohan Reddy²

¹(Research Scholar, Department of C.S.E, Sri Venkateswara University, Tirupati, A.P, India)

²(Professor, Department of C.S.E, Sri Venkateswara University, Tirupati, A.P, India)

Abstract: The amount and variety of data in natural sciences increases rapidly. Data abstraction, data manipulation and pattern discovery techniques are of great need in order to deal with such large quantities. Integration between different sources of data is also of major interest, as complex relations may arise. Biology is a good example of a field that provides extensive, highly variable and multi-sources data. The scope of these investigations has now expanded greatly owing to the development of high throughput sequencing techniques and novel statistical and computational methods. These methods require easy-to-use computer programs. We have developed a new software package, molecular genetics analysis to integrate their diverse data and tools under Common Graphical User Interfaces (GUIs) by using BLAST, FASTA data searches. One of the most exciting things about being involved in computer programming and biology is that both fields are rich in new techniques and results. The illumina is sequence analyzer in Next Generation Sequence Technology.

Keywords: DNA, RNA, BLAST, NCBI, EMBL, Biological datasets.

I. INTRODUCTION

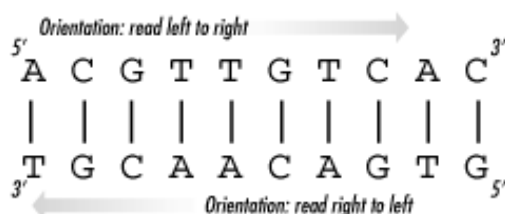
Now a day's Biological data is proliferating rapidly. With the advent of the World Wide Web and fast Internet connections, the data contained in these databases and a great many special-purpose programs can be accessed quickly, easily, and cheaply from any location in the world. As a consequence, computer-based tools now play an increasingly critical role in the advancement of biological research. *Bioinformatics*, a rapidly evolving discipline, is the application of computational tools and techniques to the management and analysis of biological data. Finally shows the enabling creation and analysis of group of sequences, large data sequences and adding new modules for visual representation of input data and output results on the Microsoft windows platform Bioinformatics is the field of science in which biology, computer science, and information technology merge to form a single discipline. It was Paulien Hogeweg who invented the term Bioinformatics in 1979 to study the processes of information technology into biological systems. The science of bioinformatics actually develops algorithms and biological software of computer to analyze and record the data related to biology for example the data of genes, proteins, drug ingredients and metabolic pathways. The term bioinformatics is relatively new, and as defined here, it encroaches on such terms as "computational biology" and others. The use of computers in biology research predates the term bioinformatics by many years. As biological data is always in raw form and there is a need of certain storage house in which the data can be stored, organized and manipulated. Biological software and databases provide the scientists this opportunity so that the data can be extracted from these database easily and can be used by the scientists.

Of course, biology is an old science, but many of the most interesting directions in biological research are based on recent techniques and ideas. The modern science of genetics, which has earned a prominent place in modern biology, is just about 100 years old, dating from the widespread acknowledgement of Mendel's work. The elucidation of the structure of deoxyribonucleic acid (DNA) and the first protein structure are about 50 years old, and the polymerase chain reaction (PCR) technique of cloning DNA is almost 20 years old. The last decade saw the launching and completion of the Human Genome Project that revealed the totality of human genes and much more. Today, we're in a golden age of biological research.

II. The Organization Of Dna

The study employed the ex post facto and descriptive comparative strategies, the data were collected using a researcher devised questionnaires with items on educational services. A minimum sample size of 775 respondents was determined using the Sloven's formula. A purposive sampling was utilized to select the respondents of the study. This involved finding any CBS in a lecture room, computer laboratory, resting place at a university campus and requesting them to participate in the study. The request involved self-introduction and soliciting students' cooperation by explaining the objective of the study. Any CBS who accepted was selected to fill in the questionnaire. Cronbach's alpha coefficient test revealed that the questionnaire was reliable at ($\alpha = 0.947$). The data were analyzed using summary statistics, such as means and ranks, while the null hypothesis was tested using t-test and analysis of variance (ANOVA).

DNA is a polymer composed of four molecules, usually called bases or nucleotides. Their names and one-letter abbreviations are adenine (A), cytosine (C), guanine (G), and thymine (T). The bases joined end to end to form a single strand of DNA. In the cell, DNA usually appears in a double-stranded form, with two strands wrapped around each other in the famous double helix shape. The two strands of the double helix have matching bases, known as the base pairs. An A on one strand is always opposite a T. on the other strand, and a G is always paired with a C.



There is also an orientation to the strands. One end of a nucleotide is called the 5' (five prime) end, and the other is called the 3' (three prime) end. When nucleotides join to make a single strand of DNA, they always connect the 5' end of one to the 3' end of the other. Furthermore, when the cell uses the DNA, as in translating it to RNA, it does so base by base from the 5' to the 3' direction. So, when DNA is written, it's done so left to right on the page, corresponding to the 5' to 3' orientation of the bases. An encoded gene can appear on either strand, so it's important to look at both strands when searching or analyzing DNA. When two strands are joined in a double helix, the two strands have opposite orientations. That is, the 5' to 3' orientation of one strand runs in an opposite direction as the 5' to 3' orientation of the other strand. So at each end of the double helix, one strand has a 3' end; the other has a 5' end. Because the base pairs are always matched A-T and C-G.

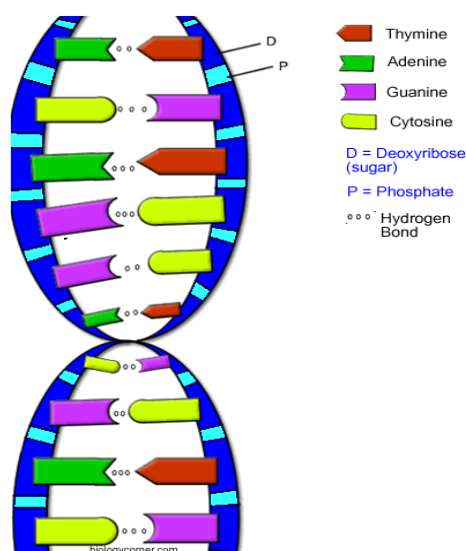


Figure 1-1. Two strands of DNA

III. Sequence Analysis

The global youth population defines young people between 15 and 24 years of age but National Youth Policy of Pakistan-2006 defines youth as all Pakistani males and females between the ages within the range of 15-29. Youth is defined as “Period of transformation from family dependent childhood to independent adulthood and integration in the society as a useful citizen”. Ferrous desires, ingenuity and perseverance are the characteristics of this age. It is also under the dominant pressure of aloofness, unpredictability and terror. As young men are influenced by various factors, they need social economic and practical co-operation. Making general policies or programs for young people aged 15-29 does not take into account the distinctive needs, experience, and attitudes defined by the younger age cohort (Adolescents aged 15-19) compared to elder age cohort (youth 20-24) and older age cohort (Adult youth 25-29). Table 1 reveals that almost half (49.5%) of the respondents were in the age group of 15-19 years or in other words they were adolescent, 32.5% respondents were in the age group of 20-24 years, while 18.0% respondents were in the age group of 25-29 years. The mean age of the respondents was 23.84 years with standard deviation 3.66 years.

Illumina have introduce sequencing systems contribution even higher throughput than the systems provided by 454, generates billions of bases in a single run. This novel method all rely on analogous, cyclic examination of sequences from spatially separated clonal amplicons. Although with shorter read lengths and a

slower sequence extraction from individual features as compared to the Sanger method, the parallelized process offers a much higher total throughput and reduces cost significantly by generating thousands of bases per second.

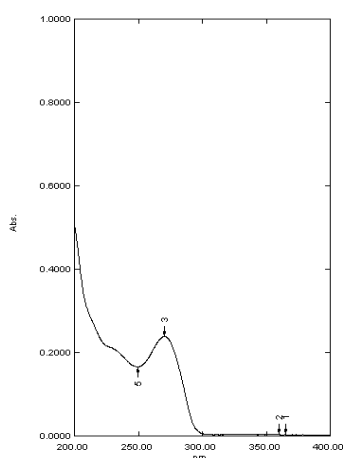


Figure 1: Sources of Planting Materials Used by Cassava Farmers in Osun State



Figure 2: Illumina for sequence analyser.

Table 1: Relationship between Depression Anxiety, Self Esteem and Life Satisfaction (N=100) among Physically Disable Individuals

Variable	BDI	BAI
BDI	-	.218
BAI	.218	-
RS	.205*	.097
LS	-.027	-.107

*Correlation is significant at the 0.05 level

Table 2: Results of Assay

Label Claim	*Amount Found (%)	
	UV-Spectrophotometric \pm SD	First Order Derivative Spectrophotometry \pm SD
Lamivudine (300mg/tablet)	100.75 \pm 0.528	100.04 \pm 0.349

*mean of six determinations

IV. Conclusions

This study showed that knowledge factor influence the use of nursing process more than other variables. One of the biggest problems currently facing the nursing profession is that of implementing the nursing process as lamented by Milne (1985) which the reporter believed that it can be influenced by the variables such as knowledge, profession, attitude, institution. Institutional factor ranks the second highest predictive value in the use of nursing process but currently, many institutions do not use nursing process for the care of their clients. for the negative attitude of nurses which is the least ranked in the use of nursing process. The illumina is used in next generation sequence technology.

Acknowledgements

The author thanks to University Grant Commission for financial support to implement the project.

References

- [1]. Arua, Uche and Chinaka, George I. (2011). Use of Library Resources by Staff and Students of Secondary Schools in Umuahia North Local Government Area of Abia State. *Library Philosophy and Practice*. 1-7.
- [2]. Edem, Nkoyo, Ani, Okon and Ocheibi, Jonathan A. (2009). Students' perceived effectiveness in the use of library resources in Nigerian universities. *Educational Research and Review*, 4(6), 322-326.
- [3]. http://en.wikipedia.org/wiki/Indian_Institute_of_Management_Rohtak
- [4]. Oyesiku FA, Oduwole AA (2004). Use of an Academic Library: A Survey on the Olabisi Onabanjo University Libraries. *Lagos J. Libr. Info. Sci.*, 2(2): 96-101.
- [5]. Ugah D (2007). Evaluating the Use of University Libraries in Nigeria: A Case Study of Michael Okpara University of Agriculture, Umudike. *Library Philosophy and Practice* 1-8.
- [6].

APPENDICES

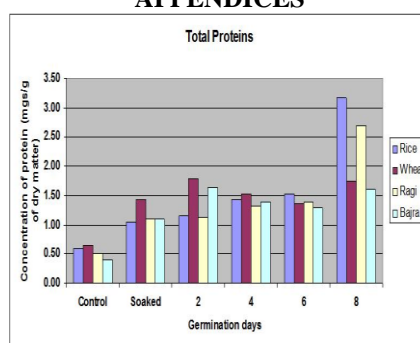


Figure 1: Total Proteins of Dry and Germination Seeds of Rice, Wheat, Ragi and Bajra

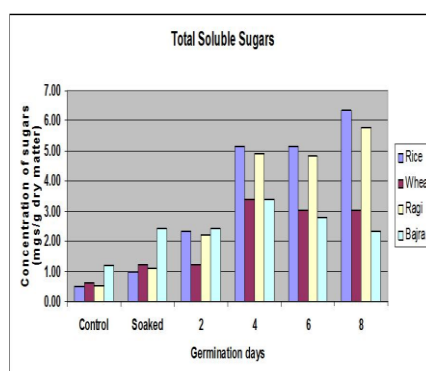


Figure 2: Total Soluble Sugars of Dry and Germination Seeds of Rice, Wheat, Ragi and Bajra