

“G’s-Remedies 1.0 (Directus): Making Analysis Simple” a Unique Amalgamation of Genetic Disorder Database with Sequence Analysis Tools

Gaurav Kumar Srivastava, Santosh Kumar

Abstract: Paper 'G'S-REMEDIES 1.0 (DIRECTUS)' is a specially designed web-based software. That software fulfills the elementary requests of a Bioinformatician during the study and analysis of principal sequences. PERL, PERL-CGI, and HTML are used for the designing and development of the DIRECTUS. This special tool can be helpful for computing various parameters that are helpful in ANALYSIS of NUCLEOTIDE and AMINO ACID RESIDUES, and the READING FRAMES. The G'S-REMEDIES 1.0 (DIRECTUS), tool emanates with the personal unique database on the GENETIC DISORDERS. This unique database gives full details almost any specific genetic disorder and it also gives information about mutation type and genetic disorder affected chromosome. The latest version provides more than fifteen analysis tools and also gives a complete image of more than fifty genetic disorders, and next versions of the tool will come with databases of viral infection databases such as eye disorders databases, and also improve to PRIMER designing tools, etc. G'S-REMEDIES 1.0 (DIRECTUS) is freely available for academic use.

Keywords: 'G'S-REMEDIES 1.0 (DIRECTUS), Directus, Genetic Disorder Database, Sequence Analysis tools.

I. INTRODUCTION

In current era of technology and sciences, bioinformaticians have developed various databases to provide details about genetic databases and their overall picture of occurrence; however these databases are area restricted. Some of these databases include, In India IGDD (Indian Genetic Disease Database) is the first patient based "genetic disease" database. This is database collection of 52 diseases also with detail information on the 5760 persons who is transports the mutant alleles of causal genes. Info about the type of mutation, locus heterogeneity, biochemical and clinical data, diseases based on geographical location and common mutations are also present in the literature [17]. In this database can be searched based on the disease of interest, type of mutation, causal gene, and geographic location of the patients or any carriers. It is information tool is developed and maintained at IICB (Indian Institute of Chemical Biology) and this is the unit of Council of Scientific and Industrial Research.

Revised Manuscript Received on February 25, 2020.

Er. Gaurav Kumar Srivastava, Research Scholar, Maharshi University of Information technology, Lucknow India.

Dr Santosh Kumar, Associate Professor and Head, School of computer science, Maharshi university of information technology Lucknow India.

Others include UK Genetic disorders databases (which provides information about genetic disorders in UK) the main aim of UK Genetic disorders databases is a improve the lives of persons and their families who is affected by genetic disorders. Frequency of Inherited Disorders Database, OMIM (Online Mendelian Inheritance in Man) knowledge database of human genes and genetic disorders [1] is a complete, confident collection of genetic phenotypes and human genes. OMIM is freely available for every one and it is daily updated. The OMIM is emphases on the link between phenotype and genotype. The HGMD® (Human Gene Mutation Database) [12] represents an effort to organize known (published) gene lesions responsible for human inherited disease [13] and is this is maintained at Cardiff by P.D. Stenson, D.N. Cooper, E.V. Ball, K. Shaw, A.D. Phillips, and M.E. Mort.).

After the Human Genome Project, a wide range of genomic and proteomics data is available worldwide [16]. For numerous genes, something is known about their molecular and biological function such as Gene Ontology Consortium [2], pathway membership [10], physical chromosomal location, level of polymorphism, RNAi phenotypes, disease phenotypes, and rate of molecular evolution [19]. The information about genomic should be presented for analysis within a combined structure where group of genes from any experiment can be questioned simply [7].

II. METHODOLOGY OF G'S-REMEDIES 1.0 (DIRECTUS)

There are many tools available such as IGDD, IICB, OMIM etc. some are only database some are only tool, G'S-REMEDIES 1.0 (DIRECTUS) provides the serve both as databases as well as data analysis tools, to save both time as well as computational memory. G'S-REMEDIES 1.0 (DIRECTUS) hence is one of a kind of its innovative software which provides server for both as a platform for genetic disorder database which covers more than fifty genetic disorders, and also an online sequential tool for primary nucleotide and amino acid sequences like six reading frames, composition counts, melting temperature, transcription, translation, extinction coefficient, half-life and gravy value. The software has been made on a user-friendly interactive CGI interface for the database and analysis tools.

III. G'S-REMEDIES 1.0 (DIRECTUS)

G'S-REMEDIES 1.0 (DIRECTUS) is known to contain a number of analysis tools, which play a crucial role during primary sequence study and also a platform for genetic disorders databases.

“G’s-Remedies 1.0 (Directus): Making Analysis Simple” a Unique Amalgamation of Genetic Disorder Database with Sequence Analysis Tools

G’S-REMEDIES 1.0 (DIRECTUS), once launched on the website, the current version will be followed by regular updates in the existing genetic disorder database. Also, the sequential tools will be upgraded to more analysis tools pertaining to PRIMER DESIGNING, RESTRICTION ENZYMES ANALYSIS and so on. Also, the current databases will be promoted to the Disorders Related to Blood, Disorders Related to Bone and Joints, Disorders Related to Growth Eye, Gastro-Intestinal Disorders, and Hearing Disorders etc. Henceforth, G’S-REMEDIES 1.0 (DIRECTUS) as a database keeps track of mutations in the genes, chromosomes and proteins etc. being affected due to the disorders. The special database will also useful to the Geneticists, Physicians and other researcher and scientist in India and abroad related to genetic disorders to retrieve and take the data for helpful to the human. Also, considering the sequential applications which have been implemented on

sequences, G’S-REMEDIES 1.0 (DIRECTUS) has a bright future aspect in primer designing and advanced genomics and proteomics assays.

The need of the hour demands common servers which can serve both as databases as well as data analysis tools, to save both time as well as computational memory. G’S-REMEDIES 1.0 (DIRECTUS) hence is one of a kind of its innovative software which provides server for both as a platform for GENETIC DISORDER DATABASE which covers more than fifty genetic disorders, and also an online sequential tool for primary nucleotide and amino acid sequences like SIX READING FRAMES, COMPOSITION COUNTS, MELTING TEMPERATURE, TRANSCRIPTION, TRANSLATION, EXTINCTION COEFFICIENT, HALF LIFE AND GRAVY value. The software has been made on a user-friendly interactive CGI interface for the database and analysis tools.

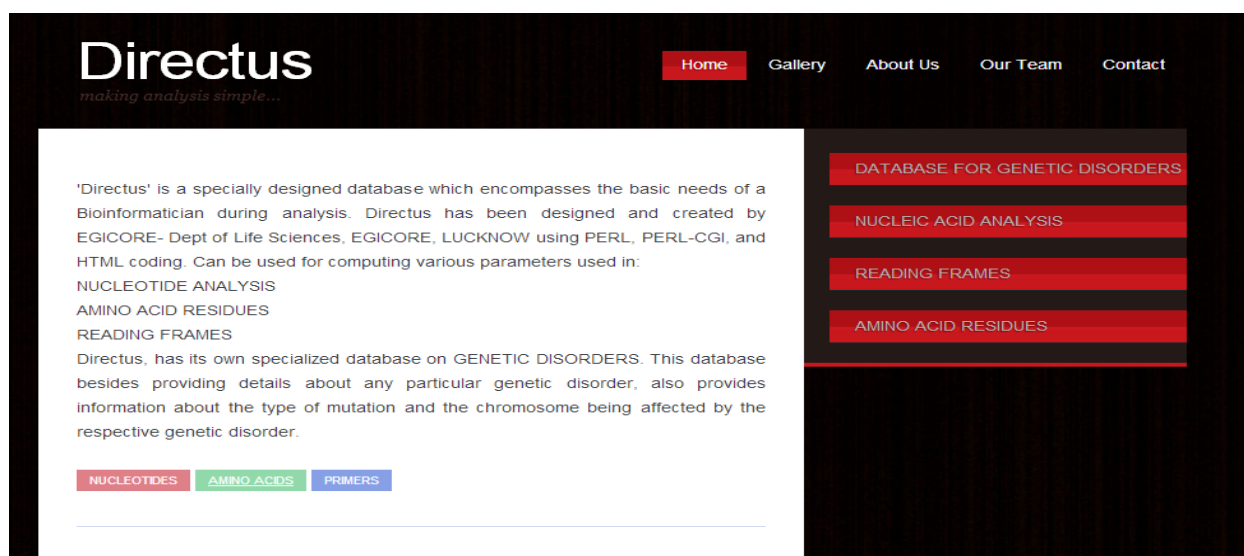


Figure 1: The home page of 'G’S-REMEDIES 1.0 (DIRECTUS)

IV. GENETIC DISORDER DATABASE

G’S-REMEDIES 1.0 (DIRECTUS) comes with its specialized database covering several genetic disorders and also the kind of mutations causing the disorders, the chromosome number on which the mutation is affecting and also an overview of the disorder accompanied by images of patient cases suffering from the respective genetic orders. In the field of bioinformatics, a Genetic disorder Database is a schematized pool of data, it provides typically structured to the model of features of the reality. This type of database is a way to understand the fundamental technique of composite diseases, by considerate numerous compound interactions by the relationship between phenotype-genotype and gene-disease technique.

The need of the hour demands common servers which can serve both as databases as well as data analysis tools, to save both time as well as computational memory. G’S-REMEDIES 1.0 (DIRECTUS) hence is one of a kind of its innovative software which provides server for both as a platform for GENETIC DISORDER DATABASE which covers more than fifty genetic disorders, and also an online sequential tool for primary nucleotide and amino acid sequences like SIX READING FRAMES, COMPOSITION COUNTS,

MELTING TEMPERATURE, TRANSCRIPTION, TRANSLATION, EXTINCTION COEFFICIENT, HALF LIFE AND GRAVY value. The software has been made on a user-friendly interactive CGI interface for the database and analysis tools

V. NUCLEIC ACID ANALYSIS TOOLS

G’S-REMEDIES 1.0 (DIRECTUS) can be used for calculating the nucleotide composition of a DNA sequence. A user can obtain reverse complement (biofinx), transcribed sequence [5] and translated sequence GC content [18] on inputting a DNA sequence. The melting temperature [14] and molecular weight [6] for DNA sequence can also be calculated using this software.

VI. SIX READING FRAMES

The six reading frames can also be translated for a user input DNA sequence in both codons [15] form as well as its translated amino acid residues sequences [4].

VII. AMINO ACID ANALYSIS

G'S-REMEDIES 1.0 (DIRECTUS), is enabled with analysis tools for amino acid residues which can calculate composition of user-input protein sequence [11]. Primary sequence information like half-life [3] for proteins, extinction coefficient [9] polarity percentage and molecular weight [8] can also be obtained by a user of G'S-REMEDIES 1.0 (DIRECTUS) by inputting a protein amino acid sequence. Discussion

VIII. RESULT AND DISCUSSION

In 'G'S-REMEDIES 1.0 (DIRECTUS), is includes genetic disorders Database, six reading frames, and amino acid analyzer. The 'G'S-REMEDIES 1.0 (DIRECTUS) provides to user interface to access and it give complete study of diseases based on genetic disorder. The multiple reading frames can provide the translation for a user input DNA sequence in both codons form as well as its translated amino acid residues sequences. The amino acid analyzer is a tool for amino acid residues which calculates composition of user-input protein sequence.

IX. CONCLUSION

'G'S-REMEDIES 1.0 (DIRECTUS)' is an online web-based application. That tool fulfills the elementary requirements of a Bioinformatician during the analysis of primary sequences. 'G'S-REMEDIES 1.0 (DIRECTUS) would provide a key platform for clinicians, epidemiologists, geneticists and genetic counselors to contact a central genomic data-source for the people. PERL, PERL-CGI, and HTML are used for the designing and development of the DIRECTUS. This exclusive tool can be benefit for analyzing the computing numerous parameters that are helpful in AMINO ACID RESIDUES, NUCLEOTIDE ANALYSIS, and the READING FRAMES. The G'S-REMEDIES 1.0 (DIRECTUS), tool comes with the own unique database on GENETIC DISORDERS.

ACKNOWLEDGMENT

The authors of G'S-REMEDIES 1.0 (DIRECTUS) would like to express their gratitude towards the faculty of Sam Higginbottom Institute of Agriculture, Technology and Sciences, Allahabad and Bio-EGICORE, Lucknow for their valuable comments, reviews and technical support, which played a crucial role in completion of the project.

REFERENCES

1. Ada Hamosh, Alan F. Scott, Joanna S. Amberger, Carol A. Bocchini and Victor A. McKusick nucleic Acids Research, Vol. 33, 2005. doi:10.1093/nar/gki033
2. Ashburner M1, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, Nature genetics, 2000
3. Bachmair, A., Finley, D. and Varshavsky, A. (1986) In vivo half-life of a protein is a function of its amino-terminal residue. Science 234, 179-186. [PubMed: 3018930]
4. Badger JH, Olsen GJ "CRITICA: Coding Region Identification Tool Invoking Comparative Analysis". Mol Biol Evol 16 (4): 512-24. 1999. PMID 10331277.
5. Berg J, Tymoczko JL, Stryer L (2006). Biochemistry (6th edition). San Francisco: W. H. Freeman. ISBN 0-7167-8724-5

6. Casey, J. and Davidson, N. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide, Nucleic Acids Research, 1977.
7. Cristian I. Castillo-Davis and Daniel L. Hartl GeneMerge—post-genomic analysis, data Mining and hypothesis testing Bioinformatics Applications Note Vol. 19 no. 7, pages 891–892, 2003
8. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005).
9. Gill, S.C. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. 182:319-326(1989). [PubMed: 2610349]
10. Kanehisa, M., Goto, S., Kawashima, S. and Nakaya, A. The KEGG databases at Genome Net. Nucleic Acids Res., 30, 42–46, 2002.
11. Kyte, J. and Doolittle, R.F. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. 157, 105-132., 1982.
12. Peter D Stenson, Matthew Mort, Edward V Ball, Katy Howells, Andrew D Phillips, Nick ST Thomas and David N Cooper, The Human Gene Mutation Database: 2008 update, Genome Medicine 2009
13. Turanov AA, Lobanov AV, Fomenko DE, Morrison HG, Sogin ML, Klobutcher LA, Hatfield DL, Gladyshev VN (January 2009). "Genetic code supports targeted insertion of two amino acids by one codon". Science 323 (5911): 259–61.
14. Rychlik, W. and Rhoads, R.E. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA, Nucleic Acids Res. 1989
15. Rainey S, Repka J. "Quantitative sequence and open reading frame analysis based on codon bias". Systemics, Cybernetics and Informatics 4 (1): 65–72, 2000
16. Sanchari Pradhan , Mainak Sengupta , Anirban Dutta , Kausik Bhattacharyya , Sumit K. Bag , Chitra Duttal and Kunal Ray, Indian genetic disease database, Nucleic Acids Research, Vol. 39, 2011
17. Stevenson AC, Kerr CB. On the distribution of frequencies of mutation to genes determining harmful traits in man. Mutation Res. 4: 339-352, 1967.
18. Yakovchuk P, Protozanova E, Frank-Kamenetskii MD "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix". Nucleic Acids Res. 34 (2): 564–74, 2006. doi:10.1093/nar/gkj454. PMC 1360284. PMID 16449200
19. You WC, Blot WJ, Chang YS, Ershow A, Yang ZT, An Q, Henderson BE, Fraumeni JF Jr and Wang TG, Allium Vegetables and Reduced Risk of Stomach . Cancer. J. Natl. Cancer Instit. , Vol. 81, pp. 162-164, 1989.

AUTHORS PROFILE



Er. Gaurav Kumar Srivastava, He has completed B. Tech in Computer Science & Engineering from Dr. A.P.J. Abdul Kalam Technical University, Lucknow, completed M. Tech in Computer Science & Engineering from Babu Banarasi Das University, Lucknow. Currently he is a research scholar of Maharshi University of Information technology, Lucknow India. He has published various research article in International peer reviewed Journals/ Conferences.



Dr Santosh Kumar, He has Associate Professor and Head of the School of computer science, in Maharshi university of information technology Lucknow INDIA. He has completed Master Degree in computer applications in year 2007 and Ph. D. Degree from Babasaheb Bhimrao Ambedkar University Lucknow (A Central University) India in year 2013. He has 7 years of teaching experience and 5 years of research experience in the field of software engineering & Fuzzy database. He has published more than 42 international and national research papers and guided more than 5 Ph.D. Scholars.