Journal of Computer Science and Applications.
ISSN 2231-1270 Volume 6, Number 2 (2014), pp. 133-139

http://www.irphouse.com

Some Studies On Interleukin Database Design

N.Deepak Kumar and Dr.A.Ramamohan Reddy

Dept.Of CSE, SVUniversity, Tirupati, India,
nagiletideepakvictor@gmail.com.
Dept.Of CSE, SVUniversity, Tirupati, India,
Ramamohansvu@yahoo.com

Abstract

With interleukins (IL), a new class of potential drugs has been introduced into clinical research. These signal peptides are involved in the regulation of many physiological and pathophysiological processes. IL-1, -2, -3, -4, -6 and -11 have been tested in clinical trials. The growth promoting, growth inhibiting or immunomodulatory activities of interleukins represent the theoretical basis for large scale clinical testing, predominantly in malignant disease. Dose-dependent effects on numbers of peripheral blood cells and recovery from bone marrow failure have been demonstrated for IL-1, -3, -6 and -11. Phase III trials are in progress to determine their value for clinical practice. However, investigations on the immunomodulatory activities proved to be more difficult. This is because key mechanisms for successful treatment of malignant disease by immunomodulation are not clearly defined and the methodology for assessment of immunostimulatory effects is not well established. Besides treatment of renal cell carcinoma and malignant melanoma with IL-2, no successful trials have been reported. However, phase I clinical trials with IL-1, IL-4 and IL-6 have just been completed.
A database manages information and allows organizing data, ensuring completeness and integrity, and transforming the data from one form to another. It make search through the data efficiently to find the desired information. In the present work a database for Interleukins (proteins) have been created as the data related to Interleukins is increasing day by day, it has become difficult for researchers to manage their structure, classification and functions. There are 37 Interleukins have been discovered by the researchers and many more may be added in the future. Interleukins study is mostly useful for diagnosing and treating all the diseases of a human body.

Keywords: Interleukins, Protein, immune cells, Bioinformatics, DBMS.

Paper code: 14604- JCSA_Include_IRPH

1. INTRODUCTION

Interleukins are biologically active glycoproteins derived primarily from activated lymphocytes and macrophages. Tremendous insight into the biochemical and biological properties of interleukins has been gained with advances in recombinant DNA technology, protein purification, and cell-culture techniques. The biological properties of interleukins include induction of T-lymphocyte activation and proliferation, augmentation of neutrophil, macrophage, and T-lymphocyte cytooxicity, and promotion of B lymphocyte and multilineage bone marrow stem-cell precursor growth and differentiation. Interleukins may play a role in the pathogenesis of several important diseases. Interleukin therapy is likely to play an important role in the treatment of cancer, infectious diseases, and immunodeficiency syndromes. [12, 14].

Specified design processes are standard in the software development industry, and there are many design processes described in the software engineering literature. The details of the design process are less crucial than the use of a process. However, there are some crucial steps, such as gathering requirements. Requirements document what the database is trying to accomplish. Most databases have to make data model compromises. Databases have been used to manage and integrate large volumes of complex data in other disciplines for decades [1].

Development of a data model is another crucial step because this helps to identify potential problems in the design early on in the project, when they are still easy to correct. The most common tool used for this purpose in relational database design is the entity relationship diagram. This type of diagram represents the real-world entities about which the database will store information, and the relationships between those entities. Use the database to enforce data integrity. A database should protect the integrity or consistency of the data that it stores. The strong theoretical basis of relational DBMS provides rules of normalization, which, if followed, will ensure basic data integrity. These rules ensure that all information is stored in the smallest meaningful pieces and is stored in only one place, preventing data duplication and the concomitant potential for internal inconsistencies. A database that obeys these rules is said to be normalized. Normalization splits related data across multiple tables, requiring queries to perform operations called joins to reassemble the data.

The recent bioinformatics literature includes numerous papers about databases, but these primarily focus on the need for integration across existing databases [2, 3], report the design and use of specific databases [4–9], or argue for better large scientific databases and the systematic changes necessary to accomplish this goal [10, 11]. All this information is valuable, but does not provide much help to novice database designers**.**

2. METHODOLOGY

2.1. Organization of Database

The organization of membrane protein function database is illustrated in Fig. 1. Each entry in the database contains the following information[8]:

(i) SEQUEENCE, (ii)PROTEIN (iii) PROTEIN SEQUENCE(iv)SPECIES (v)

PROTEIN NTOLOGIES (vi) PROTEIN ATTRIBUTE INFORMATIONS, (vii)PROTEIN MEDICAL INFORMATIONS We have provided the sequence and structure information in t he form of Uniprot [9] and NCBI, OMIM, HGNC. The functional information includes relative activity of mutants with respect to wild type protein, affinity for binding, channel, drug, glycosylation, membrane insertion, cellular signaling, membrane translocation, transport etc. The database has several features such as the retrieval of data using various conditions and displaying the results. We have provided direct links to Uniprot, and literature database. In addition links are given to related structural, functional and genomic databases as well as to prediction methods.

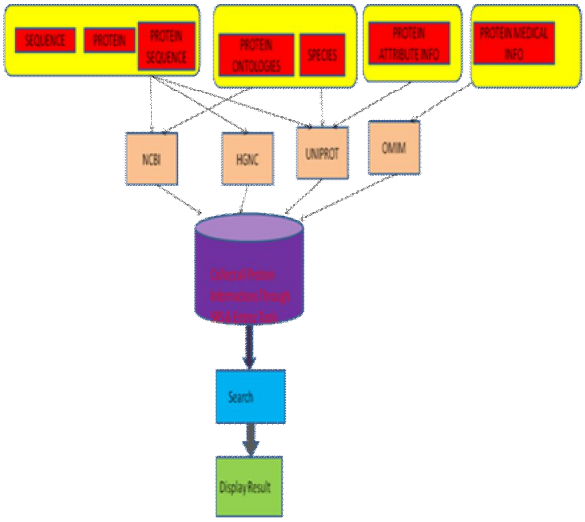

Fig I.Flow chart for organizing the Database

Factual information. Especially sequence information (nucleic acid sequences, protein sequences) is of greatest importance to biotechnology, but also map information, structure information, and property information. Commercial and financial information, especially for the biotechnology industry, play an increasing role.

Databases relevant to biotechnology can be classified

- according to subject areas:

The subject area is usually the determining point for user selection of a database. In this connection, an important factor influencing biotechnology information is the interdisciplinary of biotechnology with various sciences, application areas and related fields.

- according to the type of stored information:

Factual databases, bibliographic databases, referral databases, full-text databases.
The close link between biotechnology and information technology is particularly evident in relation to nucleic acid and protein acid sequences stored in factual

databases. In the literature, author report that the DNA sequence database will become as important as the Periodic Table of Elements. Factual databases are most important to biotechnology because these databases serve not only as an information tool but as a direct research tool. Factual databases provide a research instrument which exists at the interface between subject area and information technology, whereby the scientist increasingly assumes the role of producer and user of information, and bear witness to the increasing influence of information technology on the research process.

Depending on the type of stored information, the most importan databases in biotechnology are in the following:

(i)SEQUENCE:
a)Sequence-id
b)Protein-id
c)Sequence-Annotation
d)Sequence load date
e)Sequence-length
f)Accession Number

(ii)PROTEIN:
a)Protein-id
b)Primary-Name
c)Protein-Class Name
d)Date Created
e)Created By

(iii)PROTEIN SEQUENCE:
a)Protein-id
b)Sequence-id
c)Date Created
d)Created By
e)Comment

(iv)SPECIES:
a)Protein-id
b)Species-id (or)Taxon identifier
c)Scientific Name
d)Common Name
e)Date created
f)Created By

(v)PROTEIN ONTOLOGIES:
a)Protein-id
b)Cellular Component
c)Coding sequence Diversity
d)Molecular Function

e)Technical term

(vi)PROTEIN ATTRIBUTE INFORMATIONS:
a)Protein id
b)Species id
c)Sequence length
d)Sequence status
e)Protein Existence

(vii)PROTEIN MEDICAL INFORMATIONS:
a)Protein-id
b)Cytogenetic Location
c)HGNC Approved Symbol

3. RESULTS

The Interleukin database model consists of Protein, Protein Information, Protein OMIM information, Protein Sequence, Protein attribute informations, Sequence and Species. The entity relationship diagram is shown here for Interleukin database. In Protein, the information related to Interleukin ID, Protein primary name, Class name, date created and created by will be given. Sequence entity contain sequence id, sequence annotation, sequence load date, sequence length, accession number. Protein Sequence entity contain attributes are sequence id, protein id, date created, created By, Comment. Similarly Species entity contain the species id, protein id, scientific name, common name, Date created, created By attributes. Similarly Protein ontologies entity contain the protein id, cellular component, coding sequence diversity, Molecular function, technical term attributes. Similarly Protein Attribute entity contain Protein Gid, protein id, sequence length, species id, sequence status, protein existence attributes. Similarly Protein OMIM entity contain Protein Eid, protein id, cytogenetic location, HGNC approved symbol attributes. In each and every entity contain protein id attribute, through protein id attribute we will display the all related informations in the Database.

4. CONCLUSION

In this work an attempt performed for creation of Interleukin database for the first time by considering hypothetical information on the details of the Interleukins. As the number of discoveries on the Interleukins is increasing day by day the creation of a separate protein biological database for Interleukins along with effective robustic DBMS method is used. There are many applications of Interleukins in Biology which make them unique to understand and the information exclusively on them is very much useful to all the researchers.

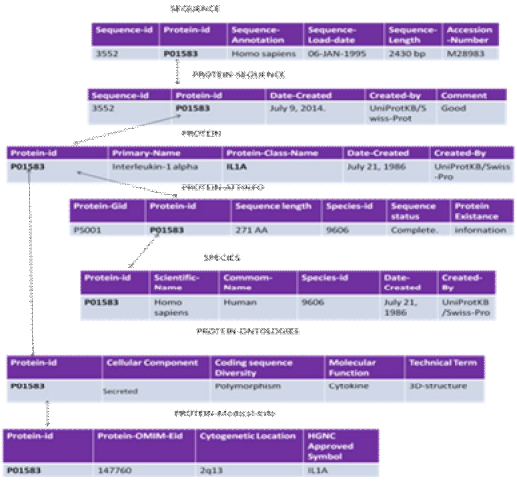

Figure II.Design Database based on above Flochart.

5. REFERENCES

Achuthsankar S Nair Computational Biology & Bioinformatics – A gentle Overview, Communications of Computer Society of India, January 2007

Aluru, Srinivas, ed. Handbook of Computational Molecular Biology. Chapman & Hall/Crc, 2006.ISBN 1-58488-406-1 (Chapman & Hall/Crc Computer and Information Science Series)

Baldi, P and Brunak, S, Bioinformatics: The Machine Learning Approach, 2nd edition. MIT Press, 2001. ISBN 0-262-02506-X

Barnes, M.R. and Gray, I.C., eds., Bioinformatics for Geneticists, first edition. Wiley, 2003. ISBN 0-470-84394-2

Peri S, et al. (2003). "Development of human protein reference database as an initial platform for approaching systems biology in humans". Genome Research 13: 2363–71.doi:10.1101/gr.1680803.

Gandhi, T.K.B. et al. Analysis of the human protein interactome and comparison with yeast, worm and fly interaction datasets. Nature Genetics. 2006. 3, 285–293

Mathivanan, S. et al. An evaluation of human protein–protein interaction data in the public domain. BMC Bioinformatics. 2006. 7, S19

Mishra, G. et al. Human protein reference database—2006 update. Nucleic Acids Research. 2006. 34, 411–414

Mathivanan, S. et al. Human Proteinpedia enables sharing of human protein data. Nature Biotechnology. 2008. 26, 164–167

Amanchy, R. et al. A compendium of curated phosphorylation-based substrate and binding motifs. Nature Biotechnology. 2007. 25, 285–286

Mathivanan S, Periaswamy B, Gandhi TK et al. (2006). "An evaluation of human protein-protein interaction data in the public domain". BMC Bioinformatics. 7 Suppl 5: S19.doi:10.1186/1471-2105-7-S5-S19. PMC 1764475. PMID 17254303.

Brocker, C; Thompson, D; Matsumoto, A; Nebert, DW; Vasiliou, V (Oct 2010).

"Evolutionary divergence and functions of the human interleukin (IL) gene family.". Human Genomics 5 (1): 30–55. doi:10.1186/1479-7364-5-1-30. PMC 3390169. PMID 21106488.
Khadka, A (2014). "Interleukins in Therapeutics". PharmaTutor 2 (4): 67–72.
Priestle JP, Schär HP, Grütter MG (December 1989). "Crystallographic refinement of interleukin 1 beta at 2.0 A resolution". Proc. Natl. Acad. Sci. U.S.A. 86 (24): 9667–71. doi:10.1073/pnas.86.24.9667. PMC 298562. PMID 2602367.
Arai K, Yokota T, Arai N, Lee F, Rennick D, Mosmann T (1985). "Use of a cDNA expression vector for isolation of mouse interleukin 2 cDNA clones: expression of T-cell growth-factor activity after transfection of monkey cells". Proc. Natl. Acad. Sci. U.S.A. 82 (1): 68–72. doi:10.1073/pnas.82.1.68. PMC 396972. PMID 3918306
ALSTON, y., COOMBS, j. (1992), Biosciences, Information Sources and Services, New York: Stockton Press.
CRAFTS-LIGHTLY, A. (1986), Information Sources in Biotechnology, Weinheim: VCH Verlagsgesell-schaft.
GRINDLEY, J.N., BENNETT, D.J. (1993), Public perception and the socio-economic integration of biotechnology, in: Biotechnologia 20, 89-102.
LŰCKE, E.-M., POETZSCH, E. (1993), Biotechnology Directory Eastern Europe, Berlin-New York: de Gruyter.
MARCACCIO, K. Y. (1993), Gale Directory of Databases, Vol. 1: Online Database, Detroit: Gale Research Inc.
MEWES, H.-W. (1990), Workshop Computer Applications in Biosciences, Book of Abstracts, p. 11, Martinsried.
POETZSCH, E. (1986), Faktographische I nformationsfonds zur Biotechnologie, Berlin: WIZ