NP

Biotechnology

Bioinformatics Information of Leguminosae Family in Gujarat State

Sagar Patel* and Hetal Kumar Panchal

G. H. Patel Post Graduate Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidyanagar, Gujarat-388120, India.

*Email: sgr308@gmail.com

Abstract

Bioinformatics is the computational analysis of biological data, consisting of the information stored in the form of DNA, Protein and Genome sequences in various biological databases. Leguminosae family is one of the largest families that contain thousands of species of Plants, Herbs, Shrubs and Trees worldwide. There are more than 250 species of this family which are found in Gujarat state. There are three subfamilies of Leguminosae family which are Fabaceae (Papilionoideae), Caesalpiniaceace and Mimosaceae. Gujarat is a state in north-western India. It has an area of 78,687 sq mi (203,800 km2) with a coastline of 1,600 km; its capital city is Gandhinagar, while its largest city is Ahmedabad. In this paper authors have collected Bioinformatics information available of 148 out of 266 Leguminosae family species like DNA, PROTEIN and Genome information of Leguminosae family from NCBI database.

Highlights

- There are more than 250 species of Leguminosae family which are found in Gujarat state.

- Various information like DNA sequence, Protein sequence and Genome information of 148 out of 266 species of Leguminosae family found in NCBI database.

- Data are provided in such way that it will very useful to many people for further study or analysis purpose.

Keywords: Leguminosae family, bioinformatics, Gujarat state, NCBI

Introduction

This Leguminosae family contains species of Plants, Herbs, Shrubs, and Trees. Legumes are used as crops, forages and green manures; they also synthesize a wide range of natural products such as flavours, drugs, poisons and dyes. Legumes are able to convert atmospheric nitrogen into nitrogenous compounds useful to plants. This is achieved by the presence of root nodules containing bacteria of the genus Rhizobium. These bacteria have a symbiotic relationship with Legumes, fixing free nitrogen for the plants; in return legumes supply the bacteria with a source of fixed carbon produced by photosynthesis. This enables many legumes to survive and compete effectively in nitrogen poor conditions. Leguminosae family is further classified into three subfamilies; 1. Fabaceae (Papilionoideae), 2. Caesalpiniaceace and 3. Mimosaceae.

NCBI (The National Center for Biotechnology Information)

A database is collection of information stored in a computer

in a systematic way, such that a computer can consult it to answer questions. A biological database is a large, organised body of persistent data, usually associated with computerized software designed to update, query and retrieve components of the data stored within the system. A simple database might be a single file containing many records, each of which includes the same set of information.

The National Center for Biotechnology Information (NCBI) is part of the United States National Library of Medicine (NLM), a branch of the National Institutes of Health. The NCBI houses a series of databases relevant to biotechnology and biomedicine. Major databases include GenBank for DNA sequences, Protein, Genome, EST etc. All these databases are available online through the Entrez search engine. http://www.ncbi.nlm.nih.gov/

Bioinformatics Information

The National Center for Biotechnology Information (NCBI 2001) defines bioinformatics as:"Bioinformatics is the field of science in which biology, computer science, and information technology merge into a single discipline. There are three important sub-disciplines within bioinformatics: the development of new algorithms and statistics which assess relationships among members of large data sets, the analysis and interpretation of various types of data including nucleotide and amino acid sequences, protein domains and the development and implementation of tools that enable efficient access and management of different types of information." In this paper we have collected various bioinformatics information like DNA, Protein and Genome sequence information of Leguminosae family members which are found in Gujarat State of India.

DNA (Deoxyribonucleic acid)

The Deoxyribonucleic acid (DNA) is a molecule that encodes the genetic instructions used in the development and functioning of all known living organisms and many viruses. Along with RNA and proteins, DNA is one of the three major macromolecules essential for all known forms of life. Genetic information is encoded as a sequence of nucleotides (guanine, adenine, thymine, and cytosine) recorded using the letters G, A, T, and C. Most DNA molecules are double-stranded helices, consisting of two long polymers of simple units called nucleotides, molecules with backbones made of alternating sugars (deoxyribose) and phosphate groups (related to phosphoric acid), with the nucleobases (G, A, T, C) attached to the sugars. DNA is well-suited for biological information storage, since the DNA backbone is resistant to cleavage and the double-stranded structure provides the molecule with a built-in duplicate of the encoded information.

Protein

Proteins are large biological molecules consisting of one or more chains of amino acids. Proteins perform a vast array of functions within living organisms, including catalyzing metabolic reactions, replicating DNA, responding to stimuli, and transporting molecules from one location to another. Proteins differ from one another primarily in their sequence of amino acids, which is dictated by the nucleotide sequence of their genes, and which usually results in folding of the protein into a specific three-dimensional structure that determines its activity.

Genome

The Genome is the entirety of an organism's hereditary information. It is encoded either in DNA or, for many types of viruses, in RNA. The genome includes both the genes and the non-coding sequences of the DNA/RNA.

Materials and Methods

Collection of Data

Here we have searched each species in NCBI database so all information are collected from NCBI database only. Details are given below in Result section.

Results and Discussion

In this paper we have differentiated Bioinformatics information such as DNA, Propein and Genome into three tables as per Leguminosae family's subfamily name wise. Species data are available in online NCBI database.

Table 1: Fabaceae (Papilionoideae)

Bioinformatics Information	Leguminosae Family Species
DNA	*Abrus precatorius L., Aeschynomene aspera L., Aeschynomene indica L., Alhagi pseudalhagi, Alysicarpus vaginalis L., Arachis hypogaea L.,Butea monosperma, Butea superba, Cajanus cajan, Canavalia ensiformis, Canavalia gladiata, Cicer arietinum L., Clitoria ternatea L.,Crotalaria albida, Crotalaria juncea L., Crotalaria medicaginea,*
Bioinformatics Information	**Leguminosae Family Species**
DNA	*Crotalaria mysorensis, Crotalaria nana, Crotalaria orixensis, Crotalaria prostrata, Crotalaria bifaria, Crotalaria retusa L., Crotalaria spectabilis, Crotalaria triquetra, Crotalaria verrucosa L., Crotalaria burhia, Crotalaria calycina, Crotalaria evolvuloides, Crotalaria filipes Benth. var. filipes, Crotalaria filipes Benth. var. trichophora, Dalbergia lanceolaria L., Dalbergia latifolia, Dalbergia sissoo, Derris scandens, Desmodium rotundifolium, Desmodium gangeticum, Desmodium heterocarpon, Desmodium renifolium, Dolichos trilobus L., Erythrina suberosa, Erythrina variegata L., Gliricidia sepium,Glycine max, Indigofera angulosa, Indigofera glandulosa, Indigofera hochstetteri, Indigofera linifolia, Indigofera linnaei, Indigofera oblongifolia, Indigofera prostrata, Indigofera sessiliflora, Indigofera spicata, Indigofera tinctoria, Indigofera coerulea, Indigofera cassioides, Indigofera colutea, Indigofera cordifolia, Indigofera glabra L., Lablab purpureus, Lathyrus aphaca L., Lathyrus odoratus L., Lathyrus sativus L., Lupinus albus L, Medicago laciniata, Medicago lupulina L., Medicago polymorpha L., Medicago sativa L., Melilotus alba, Melilotus indica, Phaseolus vulgaris L., Pisum sativum L., Pseudarthria viscida, Pterocarpus marsupium, Pueraria lobata, Pueraria tuberosa, Rhynchosia minima , Rothia indica, Sesbania bispinosa, Sesbania cannabina, Sesbania grandiflora, Sesbania sesban, Smithia sensitiva, Taverniera cuneifolia, Tephrosia strigosa, Tephrosia villosa, Tephrosia candida, Tephrosia collina, Tephrosia jamnagarensis, Tephrosia purpurea, Teramnus labialis, Trifolium repens, Trigonella corniculata, Trigonella foenuam-graecum, Uraria picta, Vicia hirsuta, Vicia sativa L., Vigna aconitifolia, Vigna unguiculata, Vigna vexillata, Vigna angularis, Vigna radiata, Vigna trilobata, Vigna umbellate L., Zornia gibbosa.*
Protein	*Abrus precatorius L., Aeschynomene indica L., Alhagi pseudalhagi, Arachis hypogaea L., Butea monosperma, Butea superba, Cajanus cajan, Canavalia ensiformis, Canavalia gladiata, Canavalia virosa, Cicer arietinum L., Clitoria ternatea L., Crotalaria juncea L., Crotalaria medicaginea, Crotalaria retusa L., Crotalaria spectabilis, Crotalaria verrucosa L., Dalbergia lanceolaria L., Dalbergia latifolia, Dalbergia sissoo, Derris scandens, Desmodium rotundifolium, Desmodium heterocarpon, Dolichos trilobus L., Erythrina suberosa, Erythrina variegata L., Gliricidia sepium, Glycine max,*
Bioinformatics Information	**Leguminosae Family Species**
Protein	*Indigofera linifolia, Indigofera spicata, Indigofera tinctoria, Lablab purpureus, Lathyrus aphaca L., Lathyrus odoratus L., Lathyrus sativus L., Lupinus albus L, Medicago laciniata, Medicago lupulina L, Medicago polymorpha L., Medicago sativa L., Melilotus alba, Melilotus indica, Phaseolus vulgaris L., Pisum sativum L., Pterocarpus marsupium, Pueraria lobata, Sesbania bispinosa, Sesbania cannabina, Sesbania grandiflora, Sesbania sesban, Smithia sensitiva, Taverniera cuneifolia, Tephrosia villosa, Tephrosia purpurea, Teramnus labialis, Trifolium repens, Trigonella corniculata, Trigonella foenuam-graecum, Vicia hirsuta, Vicia sativa L., Vigna aconitifolia, Vigna unguiculata, Vigna vexillata, Vigna angularis, Vigna radiata, Vigna trilobata, Vigna umbellata L.*
Genome	*Arachis hypogaea L., Cajanus cajan L., Cicer arietinum L., Glycine max, Lablab purpureus L., Lupinus albus L., Medicago sativa L., Melilotus alba, Phaseolus vulgaris L., Trifolium repens L., Trigonella foenuam-graecum L., Vicia sativa L., Vigna unguiculata L., Vigna radiata L., Vigna umbellata L.*

Table 2: Caesalpiniaceace

Bioinformatics Information	Leguminosae Family Species
DNA	*Bauhinia purpurea L., Bauhinia tomentosa L., Caesalpinia coriaria, Caesalpinia decapetala, Caesalpinia pulcherrima L., Cassia absus L.,Cassia mimosoides L., Cassia Montana, Cassia obtusifolia L.,Cassia occidentalis L., Cassia roxburghii, Cassia siamea, Cassia sophera, Cassia surattensis, Cassia tora L. Cassia fistula L., Cassia italica, Cassia javanica L., Delonix elata L., Delonix regia, Hardwickia binata, Parkinsonia aculeata, Peltophorum pterocarpum, Saraca asoca, Tamarindus indica L..*
Bioinformatics Information	**Leguminosae Family Species**
Protein	*Bauhinia purpurea L., Bauhinia tomentosa L., Caesalpinia coriaria, Caesalpinia decapetala, Caesalpinia pulcherrima L., Cassia absus L., Cassia mimosoides L., Cassia Montana, Cassia obtusifolia L.,Cassia occidentalis L., Cassia roxburghii, Cassia siamea, Cassia sophera, Cassia surattensis, Cassia tora L. Cassia fistula L., Cassia italica, Cassia javanica L., Delonix elata L., Delonix regia, Hardwickia binata, Parkinsonia aculeata, Peltophorum pterocarpum, Saraca asoca,Tamarindus indica L.*

Table 3: Mimosaceae

Bioinformatics Information	Leguminosae Family Species
DNA	*Acacia polyacantha, Acacia senegal L., Acacia farnesiana L., Acacia ferruginea, Acacia leucophloea, Acacia nilotica L., Acacia planifrons L., Albizia amara, Albizia lebbeck L., Albizia procera, Leucaena leucocephala, Mimosa hamata, Mimosa pudica L., Mimosa rubicaulis, Neptunia oleracea, Pithecellobium dulce L., Prosopis cineraria L., Prosopis juliflora, Samanea saman.*
Protein	*Acacia polyacantha, Acacia senegal L., Acacia ferruginea, Acacia nilotica L., Acacia planifrons L., Albizia amara, Albizia lebbeck L., Albizia procera, Leucaena leucocephala, Mimosa pudica L., Pithecellobium dulce L., Prosopis cineraria L., Prosopis juliflora, Samanea saman.*
Genome	*No Record Found.*

Table 4: Overall Bioinformatics information

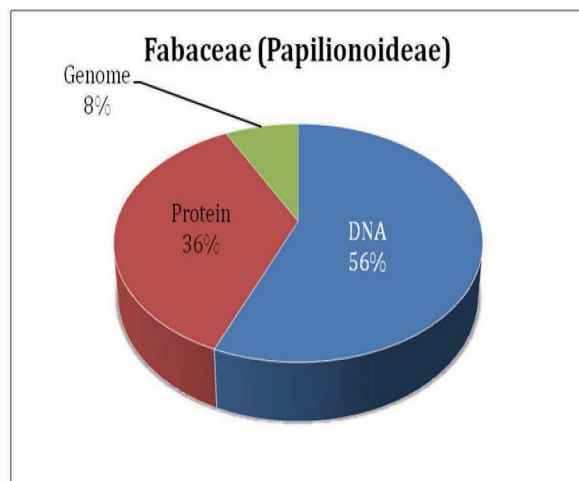| Sub family	DNA	Protein	Genome	Total
Fabaceae (Papilionoideae)	104	67	15	186
Caesalpiniaceae	25	25	-	50
Mimosaeae	19	14	-	33
Total	148	106	15	

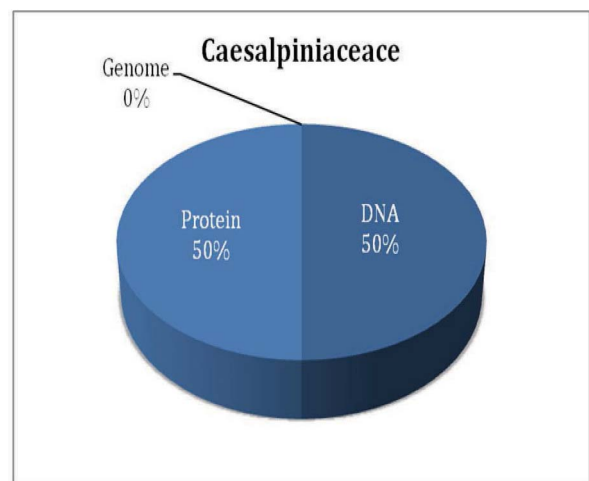**Fig. 1:** Fabaceae (Papilionoideae)
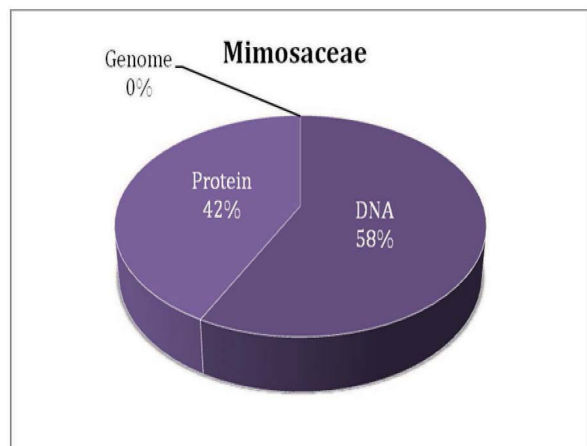

Fig. 2: Caesalpiniaceace

Fig. 3: Mimosaceae

Conclusion

There are about 266 Leguminosae Family species are found in Gujarat state then we searched every species in NCBI database we got information of 148 species out of 266 species.

Acknowledgements

References

Shah,G.L. 1978. Flora of Gujarat State. Publ. by Sardar Patel University, Vallabh Vidyanagar, Anand, India.

Kalpesh Anjaria 2002. Ph. D. Thesis: Floristic studies of Anand District. Submitted to Sardar Patel University, Vallabh Vidyanagar, Anand, India.

Sagar Patel,Sagar. 2011. Some Leguminous trees in Anand District (M.Sc., Project work) Sardar Patel University, Vallabh Vidyanagar, Gujarat, India.

Jean-Mchel Claverie and Cedric Notredame 2003.Bioinformatics – A Beginner's Guide. Publ. by Wiley Publishing, Inc. USA.

Oza,G.M and Kishore S. Rajput 2006. Biodiversity of Gujarat Forest Trees.Publ. By INSONA, Vadodara, India.

Patel, Anjaria, Panchal 2012. Leguminous Trees In Anand District: Collection and Analysis With Bioinformatics Applications. LAP Lambert Academic Publishing, Germany.

Harborne, J.B. 1994. Phytochemistry of the Leguminosae. In Phytochemical Dictionary of the Leguminosae, eds Bisby,F.A. *et al.*, London: Chapman & Hall

Heywood, V.H.(ed) 1993. Flowering Plants of the World. London: B T Batsford

Hickey, M. and King, C. 1997. Common Families of Flowering Plants. Cambridge: Cambridge University Press

Polhill, R.M. and Raven, P.H. (eds) 1981. Advances in Legume Systematics. Royal Botanic Gardens, Kew

Smartt, J. and Simmonds, N.W. (eds) 1995. Evolution of Crop Plants. Harlow: Longman Scientific & Technical

Sagar Patel and Panchal H., 2013. Leguminobase: A Tool to Get Information of Some Leguminosae Family Members From NCBI Database in *Journal of Advanced Bioinformatics Applications and Research* **4** :54-59.

Patel,Sagar., Panchal H., Anjaria K.,2012. DNA Sequence analysis by ORF Finder and Genomatix Tool: Bioinformatics Analysis of some tree species of Leguminosae Family. Publication Year: 2012, Page(s): 922- 926. *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine* (BIBM), Philadelphia, USA.

Patel,Sagar., Panchal H., Anjaria K., Phylogenetic analysis of some leguminous trees using ClustalW2 Bioinformatics Tool. Publication Year: 2012, Page(s): 917- 921. *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine* (BIBM), Philadelphia, USA.

Patel,Sagar., Panchal H., Smart J., Anjaria K., 2013.Distribution of Leguminosae family members in Gujarat State of India: Bioinformatics Approach in *International Journal of Computer Science and Management Research* **2**:2184-2189

Patel,Sagar., Panchal H., Smart J., Anjaria K., 2013. Species Information Retrieval Tool: A Bioinformatics tool for Leguminosae family in *International Journal of Bioinformatics and Biological Science* **1** (2):187-194

http://www.en.wikipedia.org

http://www.theplantlist.org/browse/A/Leguminosae

http://www.ildis.org/

http://www.ncbi.nlm.nih.gov/

http://www.kew.org/

http://www.missouribotanicalgarden.org/

http://plantnet.rbgsyd.nsw.gov.au/iopi/iopihome.htm

http://www.gujaratindia.com/state-profile/demography.htm

http://plantnet.rbgsyd.nsw.gov.au/iopi/iopihome.htm

http://www.ncbi.nlm.nih.gov/genome/

http://www.ncbi.nlm.nih.gov/nuccore/

http://www.ncbi.nlm.nih.gov/protein/