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

Paper No. 173 Received: September 20, 2013 Accepted: January 10, 2014 Published: March 09, 2014

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. Gujarath 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 computerised 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, Protein 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)

<table>
<thead>
<tr>
<th>Bioinformatics Information</th>
<th>Leguminosae Family Species</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DNA</strong></td>
<td><strong>Leguminosae Family Species</strong></td>
</tr>
<tr>
<td>DNA</td>
<td>Abrus precatorius L., Aeschynomone aspera L., Aeschynomone 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,</td>
</tr>
</tbody>
</table>
**Table 2: Caesalpiniaceae**

<table>
<thead>
<tr>
<th>Bioinformatics Information</th>
<th>Leguminosae Family Species</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DNA</strong></td>
<td>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.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Protein</strong></th>
<th><strong>Leguminosae Family Species</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3: Mimosaceae**

<table>
<thead>
<tr>
<th>Bioinformatics Information</th>
<th>Leguminosae Family Species</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DNA</strong></td>
<td>Acacia polyacantha, Acacia senegal L., Acacia farnesiana L., Acacia ferruginea, Acacia leucomphloe, Acacia nilotica L., Acacia planifrons L., Albizia amara, Albizia lebbeck L., Albizia procera, Leucaena leucocephala, Mimosa hamata, Mimosa pudica L., Mimosa rubiculis, Neptunia oleracea, Pithecellobium dulce L., Prosopis cineraria L., Prosopis juliflora, Samanea saman.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Protein</strong></th>
<th><strong>Leguminosae Family Species</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4: Overall Bioinformatics information**

<table>
<thead>
<tr>
<th>Sub family</th>
<th>DNA</th>
<th>Protein</th>
<th>Genome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabaceae (Papilionoideae)</td>
<td>104</td>
<td>67</td>
<td>15</td>
<td>186</td>
</tr>
<tr>
<td>Caesalpiniaceae</td>
<td>25</td>
<td>25</td>
<td>-</td>
<td>50</td>
</tr>
<tr>
<td>Mimosaceae</td>
<td>19</td>
<td>14</td>
<td>-</td>
<td>33</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>148</td>
<td>106</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

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

**Fig. 2: Caesalpiniaceae**
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
We would like to thank Sardar Patel University and PURSE program of DST-SPU for their support.

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

http://www.en.wikipedia.org
http://www.theplantlist.org/browse/A/Leguminosae
http://www.ildis.org/
http://www.kew.org/
http://www.missouribotanicalgarden.org/
http://www.gujaratindia.com/state-profile/demography.htm