Utilization of Computer Programs and Molecular Biology in Documentation of Egyptian clover

Abd alaziz T. Bondok 1, Shereen M. ElNahrawy1, and Ahamed Fouid2

ABSTRACT

Egyptian clover varieties are part of a very important agricultural family, second only to cereals. In this paper, the materials under the present investigation consisted of three biological datasets, containing different protein characteristics annotated from interpro, prints and quick Godatabasesthey were selected and have been applied to data from XML flat files into a local database of size 10MB in oracle tool which is used throughout the experiments Shimaa A. Badawy(2013).Moreover, in the present study, a new program was designed and applied which showed high flexibility and efficient of mining the hierarchy (Molecular Biology data of T. alexandrinum L.). In addition describe protein characteristics including family of protein, finger print of protein of Trifoliumalexandrinum L.

Keywords: Molecular biology, computer programs, documentation, Trifoliumalexandrinum L., hierarchy databases, sequences, fragments, protein.

INTRODUCTION

The genus Trifolium comprises approximately 290 species. While, T. repens, T. Pratense, T. nigrescens, etc., are important constituents of temperate pastures, T. alexandrinum and T. resupinatum are cultivated as winter annual fodder in the tropical and sub-tropical belt. Trifoliumalexandrinum, commonly known as berseem or Egyptian clover is an important winter annual fodder legume cultivated in Egypt Khaled Y. Abdel-Halim (2014).

In the present study an attempt was carried out have tried to develop methods of the Egyptian clover documentation through working on two levels:

- Molecular biology data.
- Databases of computer programs.

Databases are the heart of computer programs essentially they are electronic filing cabinets that offer a convenient and efficient method of storing vast amounts of information. There are many different databases types, depending on the nature of the information being stored (e.g., sequences, structures, etc.). The number of different databases is growing very rapidly. During the year 2000, 55 new databases were created, bringing the total at the end of the year to 2811.

Documentation contains data related to molecular biology characteristics and data bases of computer programs characteristics related to Egyptian Clover.

The Reasons for doing documentation on the Egyptian clover are to:
- Avoiding genetic erosion which has also been caused by the replacement of domestic cultivars of Egyptian clover by improved cultivars with a narrow genetic base (Hawkes, 1983).

Moreover, Gene Banks will have the fact that genetic variability allows populations to adapt to environmental changes is valued for all organisms including Egyptian clover, …..etc and microflora where the evaluation course is quite variable. The evolutionary process that maintained this diversity in the past is unable to surive in the present technological era. The present abundance of genetic diversity which still survives is being threatened by a combination of population pressures, adverse economic conditions and the interaction among these factors and subsequently further deterioration of genetic resources (Egyptian Clovercultivars).

The science of genetics in general and precisely the conservation genetics should play a sizable part to minimize the effect of these risks (Abdel Salam et al., 1994).

In Egypt, The population is rapidly increasing and at the same time these is a big gap between the food and feed needs and the available production of the crops. This due to the limited natural recourses especially the water resources. The visible way to narrow this gap is through vertical development of our agricultural system. High yielding varieties of various crops are playing very efficient roles in increasing the agricultural production vertically. Our objective in current research paper was developing high yielding cultivars of Egyptian clover.

To develop varieties so, this is a need to use of huge information in molecular biology and computer science area to make databases on DNA hierarchy and amino acids sequences levels in Egyptian clover varieties to create a new varieties in a short time.

1Forage crops Res. Department, field Crops Res. institute, ARC.Giza, Egypt.
2Faculty of computer and information, Ain Shams univ., Cairo, Egypt.

Received February 8, 2016, Accepted June 23, 2016
The aims of the present study were:
- Designed and applied new database for mining the hierarchy related to Egyptian Clover. This database will enable to describe protein attributes including family of protein and its fingerprint related to Egyptian clover.
- Determination of best programs which might be be used efficiently in documentation of Egyptian clover germplasm.
- In future goal, update of the breeding methods of Egyptian clover species employing computer programs.

MATERIALS AND METHODS

Plant material:
Plant material of Egyptian clover was obtained from the Forage Crops Res. Depart. Field Crops Res. Institute, ARC, Giza Egypt.

SDS – PAGE electrophoresis:
Total protein content was determined in grounded fine powder seeds of each sample by the method described by Bradford (1976) using bovine serum albumin (96%, Sigma Chemical Co, St. Louis, MO, USA), as standard, then total soluble proteins were extracted with extraction buffer. Fifty µL of the extract were mixed with 50 of SDS, 5% v/v β-mercaptoethanol, 7% v/v glycerol and 0.03% bromphenol blue and boiled for 7 min in a boiling water bath. 14 µL of the sample was loaded on to each well.

Electrophoresis SDS-PAGE was carried out according to the procedures of Laemmli (Laemmli, 1970) in 1.5 mm thick gels with 14% (w/v) separating gel and 4% (w/v) stacking gel in a vertical electrophoresis unit (Cleaver Scientific, England). SDS – PAGE was carried out at 75 volt for 3 hours. After electrophoresis, the gels were over night stained using 0.1% (w/v) Coomassie Brilliant Blue R-250. Then, destained using a 10% (v/v) acetic acid solution until a clear background was achieved. A page ruler pertained protein ladder (thermo – Fisher Scientific) was used as protein molecular weight marker.

Gel documentation system (Geldoc – It² imaging system, uvp, England), was applied for data scoring and documentation. Total lab analysis software (Total Lab TL120,V2008) was employed for constructing binary matrix for SDS PAGE data according to presence or absence of a band of each sample which remarked as I or O.

DATA BASE LOADER
The main issue now is how to get available biological data on to a local database in order to perform different kinds of computation.

In fact, the three discussed databases, above, are available in XML flat files to download via the web, interpro, prints and quick go. The whole process is illustrated in Fig (1). First, XML flat files are parsed by a Java or C++.

Parser program to extract table definitions tages and data. Second extracted data is loaded into oracle database, using a DBLoaderhttp://www.ailab25. engr.uconn.edu: in order to produce a local biological database, called BioDB. Note that one has to determine if he wants to process all data or part of it, database loader application in the present study is performed to the molecular biology data.

![Fig. 1. Preprocessing to create local database (BioDB)](image-url)
Inter pro database:
Inter pro is one of the database with signature diagnostic for protein families, domains, repeats, or functional sites, which amalgamates the efforts of prosite, prints, Pfam (Bat man et al, 2000), and pro dom (Corpet et al, 2000) database projects.

Prints database:
Prints database has been the international project to cooperate with prosite, pfam, and pro Dom databases. This database houses a collection, of fingerprint information for protein families, fingerprints are groups of motifs (i.e., specific protein sequences fragments) that could be inferred by aligning similar sequences (Attwood et al., 2002).

Quick Godatabase:
Quick go contains information about the gene ontology produced by the gene ontology(Go) consortium (Ashburner, et al, 2000). The gene ontology components of IPR000276. Gene ontology part of biological process of IPR 000276 is shown. Each gene ontology components has its own go entryid, in the form of go xxxx with x as digits.

RESULTS AND DISCUSSIONS
Molecular Biology Data
Biological data:
Biological data of Trifoliumalexandrinum an informative -rich domain, where data needs to be analyzed. Biological data of Trifoliumalexandrinum has been chosen for two reasons: the first one is its built-in hierarchy in most of the databases available as will be illustrated in details in the next section: The other reason is the use of biological data is timely as the information community is in great need for using data mining techniques to predict, for example gene, functions by analyzing data cumulated from diversified sources and protein functions as well in the present study used three available databases.

- Interpro database:
Interproof Trifoliumalexandrinum is one of the databases with signatures diagnostic for protein families, domains, repeats, or functional sites (Apweiler et al 2000) which amalgamates the efforts of PROSITE (Hoffman et al, 1999), PRINTS’ (Attwood et al, 2002), Pfam(Batman et al, 2000), and ProDom(Corpet et al, 2000) database projects. It is a vital tool for the computation of functional classification of newly determined sequences of Trifoliumalexandrinum that lack biochemical characterization. Interproof Trifoliumalexandrinum has been developed to rationalize protein family characterization and inherit functional insights in order to discover new functionalities. Interproof Trifoliumalexandrinum provides interface for both text-and sequence-based searches. Overlapping domains, signatures or profiles describing common domains or protein families were merged into a single Interpro entry with a unique accession number (which takes the form IPRxxxx, where x is a digit). An example of an Interpro of Trifoliumalexandrinum search results is shown in Fig. (3) Other links of protein family signatures are also provided, such as Pfam database (PF00001), PRINTS database (PR00237), and PROSITE (PS00237). The type of ID is specified whether it is a protein family domain repeat, or functional site, where it is a family type in the current case.

Figure (3) also illustrates family tree if there exists one, whether it is a-family or a domain type. Hierarchical information in the interproof Trifoliumalexandrinum database includes both is-a- (parent -child) and contains found -in information for example, in 'is-a' relationship if there exists an edge between two nodes we can say that one of them is a parent of the other.

- PRINTS database:
PRINTS of Trifoliumalexandrinum database contains information about the Pfam, and ProDom databases. This database houses a collection of fingerprint information for protein families (Attwood et al, 2002). Fingerprints of Trifoliumalexandrinum are groups of motifs (i.e., specific protein sequences fragments) that could be inferred by aligning similar sequences. Those motifs characterize aligned family and provide specific diagnostic signature. Fingerprints are more powerful than single motif approaches. The technique used to collect fingerprints is to discriminate patterns in a hierarchical form, i.e., protein sequences. Such a hierarchical approach has been used to, resolve G-protein-Coupled -Receptors (GPCRs) super families into their constituent families and receptors subtypes and to classify a variety of channel proteins.
transporters, and enzymes. Fig. (4) shows the results for family ID IPR000276 fingerprint at PRINTS. Each fingerprint of *Trifolium alexandrinum* entry has its own PRINTS entry ID< called accession number, which is PR00237 in this cases.

- **QuickGO Database:**
  QuickGO of *Trifolium alexandrinum* contains information about the Gene ontology, produced by the Gene Ontology (GO) Consortium (Ashburner et al 2000). The Gene ontology components of IPR000276 are shown in Fig. (6) where they are linked to interpro database. In Fig. (5) (a), Gene Ontology part of Biological process IPR000276 is shown. Each Gene Ontology of *Trifolium alexandrinum* component has its own Go-entry id, in the form of GO-XXXXX with x as digits. Representation of hierarchical knowledge has widely seen as an important aspect in the design of a formal ontology. Gene Ontology structure is represented as a directed-Acyclic Graph (DAG) that represents a network rather than a tree. Each node can be a child or a parent; a child may have more than one parent. There are two types of parents in a hierarchical structure that a node can have; is-a/relationship or part-of relationship. A node can have more than one parent of one type ‘is-a’/part-of relationship or a mixed of both and part-of relationships.

<table>
<thead>
<tr>
<th>Access. Number</th>
<th>Matches: 6914 proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature of protein families</td>
<td>Rhodopsin-like GPCR superfamily</td>
</tr>
<tr>
<td>Signature of protein families</td>
<td>PF0001:7Tm-1</td>
</tr>
<tr>
<td>Signature of protein families</td>
<td>PR00237:GPCRRHODOPSIN-1</td>
</tr>
<tr>
<td>Signature of protein families</td>
<td>PS00237:GproteinR Receptor Fl-1</td>
</tr>
<tr>
<td>Signature of protein families</td>
<td>PS50262:Gprotein recap Fl-2</td>
</tr>
</tbody>
</table>

**Fig. 3. Interpro of T. alexandrinum Examples of IPR000276 Search results**

<table>
<thead>
<tr>
<th>No. of motifs</th>
<th>Accession</th>
<th>PR00237</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of moieties</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Creation Date</td>
<td>12-JUL-2010</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Rhodopsin-like superfamily signature (GPCR)</td>
<td></td>
</tr>
<tr>
<td>Database References</td>
<td>Interpro:IPR000276,BLOCKS:BL00237:PFam:PF00001</td>
<td></td>
</tr>
</tbody>
</table>

G protein-coupled receptors (GPCRs) constitute a vast protein family that encompasses a wide range of functions.

**Fig. 4. Shows the result for family ID IPR 000276 fingerprint at PRINTS of T. alexandrinum**
The is-a' relationship or 'part-of relationship, node can have more than one parent of one type is-a'/'part-of relationship refers to 'when a child is an instance of the parent.and the 'pat of relationship. These types of relationships of Trifolium alexandrinum are available for the three, extensions of gene product; the molecular function, the biological processes and the cellular component. For example, in Fig. (5) a tree term, which has a (P) in front of it, means this term' is a part of the above term. However, a tree term, which has an (I) in front of it, means this term is a child of the term above it. Fig. (6)illustrates a part of the biological process of IPRG000276; a biological process is a part of the Gene ontology, but a cellular process is a biological process and a cell communication is a cellular process, etc. In Fig. (5), which shows the description of molecular function of IPR000276, molecular function is a part of the Gene ontology, but signal transducer activity is a molecular function. Finally, each item described in the tree of biological process, molecular function, or cellular component, may have other children, which applies to the same rules as their parents.

**Tables Created at BioDB:**

BioDBof Trifolium alexandrinum contains a database schema from annotating information of the three databases as illustrated in Fig. (7), which consist of 6 Tables as follows:

<table>
<thead>
<tr>
<th>GO:0007186</th>
<th>Name</th>
<th>G-protein coupled receptor protein signaling pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent tree</td>
<td>Tree</td>
<td>Gene ontology (GO:0003673) (P) Biological process (GO:0008150)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1): Cellular process (1) Cell communication (1) signal transduction</td>
</tr>
<tr>
<td>Child terms</td>
<td></td>
<td>- Signal transduction during conjugation with cellular fusion.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Signal transduction during conjugation without cellular fusion</td>
</tr>
<tr>
<td>Part of child terms</td>
<td>Interpro Mappings</td>
<td>- Psin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Chemokine receptor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- G-Protein, Gamma Subunit</td>
</tr>
</tbody>
</table>

**Fig. 5. Part of Quick GO of T.alexandrinum reference for biological process of IPR000276 (GO:0007186) (HTTP://WWW.geneontology.org)**

<table>
<thead>
<tr>
<th>GO:0001584</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhodopsin-like receptor activity</td>
<td>Name</td>
</tr>
<tr>
<td>Gene ontology (GO:0003673)</td>
<td>Tree</td>
</tr>
<tr>
<td>(P) Molecular function (GO: 0008150)</td>
<td>(1): signal transducer activity</td>
</tr>
<tr>
<td>(1) Receptor activity</td>
<td>(1) transmembrane receptor activity</td>
</tr>
<tr>
<td>(1) G-protein coupled receptor Activity</td>
<td>(I) Rhodopsin-like receptor Activity</td>
</tr>
<tr>
<td>- Nucleotide receptor activity</td>
<td>Child terms</td>
</tr>
<tr>
<td>- Viral receptor activity</td>
<td></td>
</tr>
<tr>
<td>- Amine receptor activity</td>
<td>Interpro Mappings</td>
</tr>
<tr>
<td>- Opsin</td>
<td></td>
</tr>
<tr>
<td>- Histamine H4 receptor</td>
<td></td>
</tr>
<tr>
<td>- Perojpsin</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 6. Part of QuickGO reference for molecular function of IPR000276 (GO: 0001584)**
Fig. 7. BioDatabase of *Trifolium alexandrinum* Schema used in experiments
A- Interpro Table of *Trifolium alexandrinum*:

Which includes Interpro entry ID, (primary key), type of the entry (family, domain, etc), parent list/child list (the relationship used to indicate true protein family/subfamily relationships), contains list the relationship used to indicate domain composition. Some domains can be found in more than one type of protein or family of proteins, but is not a SUBTYPE in the family.

B- INTERPRO-GO Table of *Trifolium alexandrinum*:

That includes Interpro entry ID. Interpro entry may have multiple Gene ontology (GO) annotations, hence the possible duplications of Interpro-ID in this table. Not all Interpro entries are annotated by GO. Another attribute in Interpro-GO also includes GO-ID, which may be associated with multiple Interpro entries as different proteins may share similar functions. For example, in Table (1) Interpro-id IPR000003 has three molecular functions; DNA binding (GO:0003677), Ligand-dependent nuclear receptor and GO-category describes functional classification of the entry, where Gene Ontology (GO) terms are described in' three categories: Molecular function, Cellular Component, Biological Process. Finally, "GO-description" gives brief annotation about the function of the interpro entry. An example of the outcome of INTER-PRO-GO table is given in Table (1).

C- FINGERPRINT Table of *Trifolium alexandrinum*:

In Table (2) lists protein fingerprint fields from PROSITE, PFAM, PRODOM, and PRINTS to each Interpro entry. It includes Interpro-id, Fprint-code, Fprint-name2, F print-name3, Fprint-type and print-links. Fprint' code resembles the entries in each fingerprint database. F print-name2 defines the accession number at the PRINTS database. Fprint-name3 is the given name for this particular fingerprint at PRINTS; Fprint-type is the type of the composition of the fingerprint (the number of motifs in the fingerprint). Fprintlinks contains 8 different databases of fingerprint. Table (2), illustrates an example of *Trifolium alexandrinum*Attributes of FINGERPRINT table, such as Interpro-id is IPR000482 its Fprint-code is BRECEPTR Fprint-name is PR00651, Fprint-type is compound (8) and Fprint-like is PR00251.

### Table 1. Example of Attributes of INTERPRO-GO table

<table>
<thead>
<tr>
<th>ROW</th>
<th>INTERPROID</th>
<th>GO ID</th>
<th>GO CATEGORY</th>
<th>GO DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IPR000001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>IPR000002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>IPR000003</td>
<td>0003677</td>
<td>Molecular function</td>
<td>DNA binding</td>
</tr>
<tr>
<td>4</td>
<td>IPR000005</td>
<td>0003700</td>
<td>Molecular function</td>
<td>Transcription factor activity</td>
</tr>
<tr>
<td>5</td>
<td>IPR000006</td>
<td>0046872</td>
<td>Molecular function</td>
<td>Metal ion binding</td>
</tr>
<tr>
<td>6</td>
<td>IPR000007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>IPR000008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>IPR000009</td>
<td>000159</td>
<td>Cellular component</td>
<td>Protein phosphatase type 2A complex</td>
</tr>
<tr>
<td>9</td>
<td>IPR000010</td>
<td>004869</td>
<td>Molecular function</td>
<td>Cysteine protease inhibitor activity</td>
</tr>
<tr>
<td>10</td>
<td>IPR000011</td>
<td>006464</td>
<td>Biological process</td>
<td>Protein modification</td>
</tr>
<tr>
<td>11</td>
<td>IPR000012</td>
<td>005554</td>
<td>Molecular function</td>
<td>Molecular function unknown</td>
</tr>
<tr>
<td>12</td>
<td>IPR000013</td>
<td>005576</td>
<td>Cellular component</td>
<td>Extracellular region</td>
</tr>
<tr>
<td>13</td>
<td>IPR000014</td>
<td>004871</td>
<td>Molecular function</td>
<td>Single transducer activity</td>
</tr>
<tr>
<td>14</td>
<td>IPR000015</td>
<td>005215</td>
<td>Molecular function</td>
<td>Transporter activity</td>
</tr>
<tr>
<td>15</td>
<td>IPR000018</td>
<td>007186</td>
<td>Biological process</td>
<td>G-protein coupled receptor protein signalling pathway</td>
</tr>
<tr>
<td>16</td>
<td>IPR000020</td>
<td>005576</td>
<td>Cellular component</td>
<td>Extracellular region</td>
</tr>
<tr>
<td>17</td>
<td>IPR000021</td>
<td>0016021</td>
<td>Cellular component</td>
<td>Membrane</td>
</tr>
<tr>
<td>18</td>
<td>IPR000022</td>
<td>0016874</td>
<td>Cellular component</td>
<td>Ligase activity</td>
</tr>
<tr>
<td>19</td>
<td>IPR000023</td>
<td>0003872</td>
<td>Molecular function</td>
<td>6- phosphofructokinase activity</td>
</tr>
</tbody>
</table>
Table 2. Example of Attributes of Fingerprint table

<table>
<thead>
<tr>
<th>ROW</th>
<th>Finger print code</th>
<th>Finger print</th>
<th>Finger print name</th>
<th>Finger print type</th>
<th>Interpro</th>
<th>Print links</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11SGLOBULIN</td>
<td>PR00439</td>
<td>11-S seed storage protein family signature</td>
<td>COMPOUND(7)</td>
<td>IPR000459</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1433ZETA</td>
<td>PROO305</td>
<td>14-3-3 protein zeta signature</td>
<td>COMPOUND(6)</td>
<td>IPR000308</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2CENDOP TASE</td>
<td>PROO916</td>
<td>2C endopeptidase (C24) cysteine protease family signature</td>
<td>COMPOUND(4)</td>
<td>IPR000317</td>
<td>PRO017 SRVСYSP TASE</td>
</tr>
<tr>
<td>4</td>
<td>2FENDOP TASE</td>
<td>PROO159</td>
<td>2Fe-4S ferredoxin signature</td>
<td>COMPOUND(2)</td>
<td>IPR000564</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2SGLOBULIN</td>
<td>PROO551</td>
<td>2-Sglobulin family signature</td>
<td>COMPOUND(9)</td>
<td>IPR000677</td>
<td>PRO0439 11SGLOBULIN</td>
</tr>
<tr>
<td>6</td>
<td>3FE4SFRD OXINN</td>
<td>PROO352</td>
<td>3Fe-4S ferredoxin signature</td>
<td>COMPOUND(3)</td>
<td>IPR001080</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4DISULPH CORE</td>
<td>PROO0003</td>
<td>4-disulphide core signature</td>
<td>COMPOUND(4)</td>
<td>IPR002221</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>4FE4SFRD OXIN</td>
<td>PROO0353</td>
<td>4Fe-4S ferredoxin signature</td>
<td>COMPOUND(2)</td>
<td>IPR001450</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>5HT1BREC EPTR</td>
<td>PROO0512</td>
<td>5-hydroxytryptamine 1A receptor signature</td>
<td>COMPOUND(7)</td>
<td>IPR000610</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>10</td>
<td>5HT1DREC EPTR</td>
<td>PROO0513</td>
<td>5-hydroxytryptamine 1D receptor signature</td>
<td>COMPOUND(5)</td>
<td>IPR002147</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>11</td>
<td>5HT1FREC EPTR</td>
<td>PROO0514</td>
<td>5-hydroxytryptamine 1D receptor signature</td>
<td>COMPOUND(5)</td>
<td>IPR000505</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>12</td>
<td>5HT2ARE CEPTR</td>
<td>PROO0515</td>
<td>5-hydroxytryptamine 1F receptor signature</td>
<td>COMPOUND(7)</td>
<td>IPR000450</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>13</td>
<td>5HT2ARE CEPTR</td>
<td>PROO0516</td>
<td>5-hydroxytryptamine 2A receptor signature</td>
<td>COMPOUND(7)</td>
<td>IPR000455</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>14</td>
<td>5HT2BREC EPIR</td>
<td>PRO00651</td>
<td>5-hydroxytryptamine 2B receptor signature</td>
<td>COMPOUND(8)</td>
<td>IPR000482</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>15</td>
<td>5HT2REC EPTR</td>
<td>PRO00517</td>
<td>5-hydroxytryptamine 2C receptor signature</td>
<td>COMPOUND(8)</td>
<td>IPR000377</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>16</td>
<td>5HT4REC EPTR</td>
<td>PR01059</td>
<td>COMPOUND(3)</td>
<td>COMPOUND(11)</td>
<td>IPR001520</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>17</td>
<td>5HT5ARE CEPTR</td>
<td>PRO00518</td>
<td>5-hydroxytryptamine 5A receptor signature</td>
<td>COMPOUND(5)</td>
<td>IPR001397</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
<tr>
<td>18</td>
<td>5HT5BRE CEPTR</td>
<td>PR00519</td>
<td>5-hydroxytryptamine 5B receptor signature</td>
<td>COMPOUND(6)</td>
<td>IPR000431</td>
<td>PRO00251 BACTRLOP SIN</td>
</tr>
</tbody>
</table>
D-Interpro-examples table of *Trifolium alexandrinum*:

This table collects a representative list of the kind of proteins matching the entry. The list shows the diversity of the matches in terms of function and/or taxonomic range for each InterPro entry. The proteins are from either SWISSPROT or TREMBL. INTERPRO-EXAMPLES, table contains four attributes of *Trifolium alexandrinum*: Interpro-ID, Match-pid, match-pdb, and Match-pname; Interpro-ID, is the Interpro entry ID, where Match-pid contains IDS for matched proteins. Match-pids can be used to protein sequences from protein table. Match-pdb describes two protein databases that contain the matching protein entries. Finally, match-pname is the names of the matching protein entries.

E-Quick Go table:

In table (1) consists of four attributes: Go-id, Go-Isarelation, and go-Partofrelation. Go-id is the Go entry id, where Go-Isarelation shows if there exist a parent–child relationship between this go-id and other Go-ID finally. Go-Partofrelation shows whether this go-id is Go; a part of other Go-IDS or not for example if Go-ID is 0004872 the Go-isarelation is Go :0004872 and Go-Partofrelation would have a null value.

F-Motif Table:

This table contains three attributes print-code, match-proteins, and Motifseq3, where it includes the matched proteins for a specific fingerprint code. For example, in Table (3), Finger-Print code 115Globulin has a number of matching proteins such as Gluz-Trifolum and Gu12-Trifolum Motifs sequences are shown in the same table too as well.

REFERENCES


1- Interpro Databases.
2- Prints Databases.
3- Quick Go Databases.

Genetic Erosion

Server Software 10MB)