计算分子生物学(网络版), 2014 年, 第 3 卷, No.11, 1-8
Jisuan Fenzi Shengwuxue (Online), 2014, Vol.3, No.11, 1-8
http://cmb.biopublisher.cn

研究论文
Research Article

古吉拉特邦豆科家庭数据库(GLDB): 存在于印度古吉拉特邦的豆科植物生物信息学数据库

Sagar S. Patel [1]✉, Dipti B. Shah [1]✉, Hetalkumar J. Panchal [2]✉
1 G. H. Patel Post 研究生计算机科学与技术系, Sardar Patel 大学, Vallabh Vidyanagar, Gujarat-388120, 印度;
2 古吉拉特农业生物技术研究所, Navsari 农业大学, 苏拉特, 古吉拉特-395007, 印度
✉ 通讯作者: sgr308@gmail.com; ✉ 作者

计算分子生物学, 2014 年, 第 3 卷, 第 11 篇 doi: 10.5376/cmb.cn.2014.03.00011

推荐引用：

Patel et al., 2014, Gujarat State Leguminosae Family Database (GLDB): Bioinformatics Database of Leguminosae Family present in Gujarat state of INDIA, Computational Molecular Biology, Vol.4, No.12, 1-13 (doi: 10.5376/cmb.2014.04.0012)

摘　要　生物数据库在生物信息学中发挥着重要的作用，它们为科学家提供了集中访问各种生物数据的机会。目前，分子数据可用于对许多不同植物物种进行分类、进化和情缘关系分析。我们尝试为豆科家族物种生成初步的生物信息学数据，尽可能在一种平台上提供特定类型的信息。总之，我们做了一个包括在印度的古杰雷特州发现的豆科家族所有信息的数据库，包含每个物种的植物学信息和生物信息学信息，并在一个平台上进行分析。创建这种类型的数据库反映出生物信息学可以在植物数据库中发挥重要作用的当代的跨学科方法。开发生物信息学数据库并将所有这些相关信息放在社会特别是科学界一个广泛的平台上，使之得到进一步延伸。

关键词　生物信息学, 数据库, 豆科家族

Gujarat State Leguminosae Family Database (GLDB): Bioinformatics Database of Leguminosae Family present in Gujarat state of INDIA

Sagar S. Patel [1]✉, Dipti B. Shah [1]✉, Hetalkumar J. Panchal [2]✉
1 G. H. Patel Post Graduate Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidyanagar, Gujarat-388120, India;
2 Gujarat Agricultural Biotechnology Institute, Navsari Agricultural University, Surat, Gujarat- 395007, India
✉ Corresponding author, sgr308@gmail.com; ✉ Authors

Abstract Biological databases play a central role in bioinformatics. They offer scientists the opportunity to access a wide variety of biologically data centrally. In present days molecular data are available for many plant species which can be analyzed in relation to taxonomic or evolutionary or affinity context for different plant species. Authors have tried to generate preliminary Bioinformatics data for Leguminosae family species. As much as possible of a particular type of information should be available in one plat form. In short we have made a database which includes all information of Leguminosae family which are found in Gujarat state of India, which contains Botanical information of each species and Bioinformatics information with analysis at one platform. The creation of such kind of database is reflection of interdisciplinary approach for current era where bioinformatics can play major role for botanical database. Which is further extended by developing bioinformatics database and placed all such relevant information on a wide platform for community in general as well as for scientific society in particular.

Keywords Bioinformatics, Database, Leguminosae family

收稿日期：2014 年 12 月 10 日
接受日期：2014 年 12 月 30 日
发表日期：2015 年 01 月 07 日
基金项目：无

生物信息学已经发展成为一个应用于生物技术和生物科学的信息和计算机技术综合发展的完整的多学科课题。生物信息学使用计算机软件工具进行数据库创建、数据管理、数据储存、数据挖掘

和全球网络通信，包括核酸序列(基因和 RNA)蛋白质序列和结构信息的注释、存储、分析和检索。包括序列和结构信息的数据库以及访问、搜索、可视化的方法和检索信息。生物信息学致力于创建和维护生物信息数据库，研究人员可以访问现有信息也可以输入新的内容。

功能基因组学，分子结构，蛋白质组分析，细胞代谢，生物多样性，化学工程下游加工，药物和疫苗设计都是生物信息学领域的组成部分。豆科家族是最大的家族之一，包含数千种植物，草本、灌木、树木，其中包含超过 250 种在古吉拉特邦发现的物种。豆科分为 3 个亚科：蝶形花科(Papilionaceae)、含羞草科和苏木科。

1 方法

该数据库包括豆科家族每个物种的信息，如其描述、经典分类，在古吉拉特州分布的地方、本地名称、图片、用途等。该数据库还包括古吉拉特州部分的豆科家庭成员的分布，包括其在中古吉拉特邦不同地区的谷歌地图分布以及在该特定地区发现的物种的列表。该数据库中包含的信息来自 NCBI 数据库和五个 RNA-Seq 序列的生物信息学分析。以这样的方式在这个数据库中提供的数据信息将会对进一步研究或分析非常有用。对任何植物物种进行详细研究都需要它可用的核心数据，创建可以集中提供并根据未来需求定期更新的数据库是集中数据的方法之一。

1.1 数据收集

为了创建数据库，我们需要被称为 Data.Information 的所有豆科家族的物种信息，比如它的描述、经典分类、分布的地方、本地名称。用途和图片是从博士论文，书籍等各种资源收集的。编辑和收集每个豆科家庭物种后，我们需要创建一个数据库来访问和检索每个物种的数据。我们在 XAMPP、Dreamweaver 和脚本 PHP 语言的帮助下创建了数据库。

2 结论

2.1 物种信息检索工具

作为研究的一部分，设计一种名为“物种信息检索工具”的工具，用来获得在古吉拉特州发现的豆科物种的信息。用户必须单击 GDB 数据库左侧的植物信息选项。图 1 是 GDB 数据库的主页，图 2 是物种信息检索工具的主页，选择物种点击提交按钮之后，它将给出所选择的特定物种的完整结果(图 3)。

About Leguminosae Family

图1 GDB数据库主页

Figure 1 Home page of GLDB database

Species Information Retrieval Tool

Sub-Family : Fabaceae (Papilionaceae)

Genus : Butea

Species : Butea monosperma (Lam.) Taub.

图2物种信息检索工具主页

Figure 2 Home page of Species Information Retrieval Tool

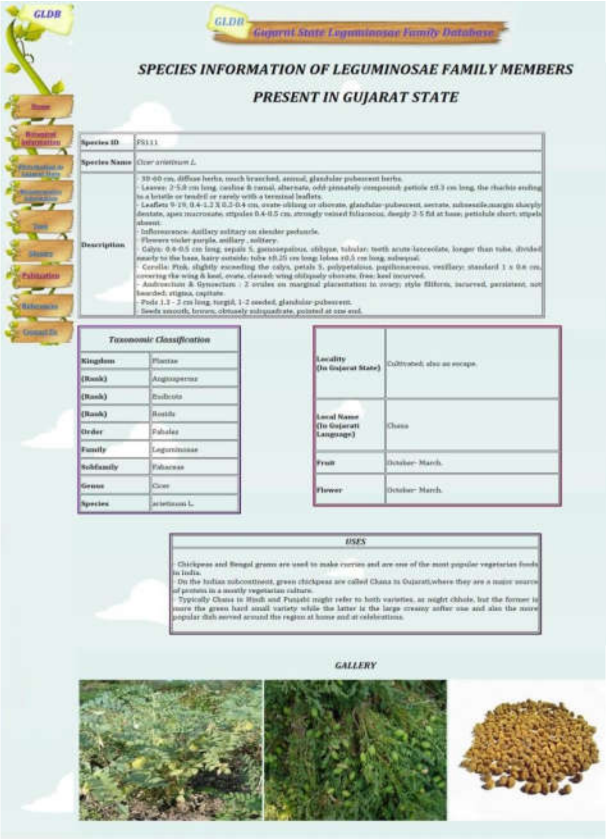
2.2古吉拉特邦的豆科家庭成员的分布

古吉拉特邦分为几个子区：喀奇、骚拉什特拉、北古吉拉特邦、古吉拉特邦中部、南古吉拉特邦和其他地区(包括在森林、花园中发现的物种)等。此外，创建的区域分配合理的网页需要提供在该特定地区发现的那些物种的信息。用户必须点击GDB数据库左侧的古吉拉特州分布选项，这些数据都是通过使用生物信息学方法开发的，所以用户还可以在Google地图中获取每个区域的位置，根据选择和点击各自的物种后，它将显示物种的完整描述。图4是显示了古吉拉特州和豆科家族的区域分布的信息的首页；图5是Kutch地区的信息，如果用户点击一种物种，它将显示如图6的结果页面。

2.3豆科家族的数据和分析

本节通过工具和数据分析介绍了豆科家族物

种的生物信息学数据分析(图7)。一种名为“豆科工具”的工具被开发用来获得在古吉拉特邦发现的每个豆科家庭如DNA，蛋白质，基因组等物种信息。名为“ConSeq工具”的一种工具被开发用来找出提交的序列中的保守序列。本文还介绍了五种豆科家族物种的RNA-Seq数据分析，一种被描述为“De novo RNA-Seq”的数据分析(表1; 2)。


GLDB

Gujarat State Leguminosae Family Database

SPECIES INFORMATION OF LEGUMINOSAE FAMILY MEMBERS PRESENT IN GUJARAT STATE

Species ID	FS111
Species Name	*Cicer arietinum L.*
Description	- 30-60 cm, diffuse herbs, much branched, annual, glandular pubescent herbs. - Leaves: 2-5.8 cm long, cauline & ramal, alternate, odd-pinnately compound; petiole ±0.3 cm long, the rhachis ending to a bristle or tendril or rarely with a terminal leaflets. - Leaflets 9-19, 0.4-1.2 X 0.2-0.4 cm, ovate-oblong or obovate, glandular-pubescent, serrate, subsessile,margin sharply dentate, apex mucronate; stipules 0.4-0.5 cm, strongly veined foliaceous, deeply 2-5 fid at base; petiolule short; stipels absent. - Inflorescence: Axillary solitary on slender peduncle. - Flowers violet-purple, axillary , solitary. - Calyx: 0.4-0.5 cm long; sepals 5, gamosepalous, oblique, tubular; teeth acute-lanceolate, longer than tube, divided nearly to the base, hairy outside; tube ±0.25 cm long; lobes ±0.5 cm long, subequal. - Corolla: Pink, slightly exceeding the calyx, petals 5, polypetalous, papilionaceous, vexillary: standard 1 x 0.6 cm, covering the wing & keel, ovate, clawed; wing obliquely obovate, free; keel incurved. - Androecium & Gynoecium : 2 ovules on marginal placentation in ovary; style filiform, incurved, persistent, not bearded; stigma capitate. - Pods 1.3 - 2 cm long, turgid, 1-2 seeded, glandular-pubescent. - Seeds smooth, brown, obtusely subquadrate, pointed at one end.

Taxonomic Classification	
Kingdom	Plantae
(Rank)	Angiosperms
(Rank)	Eudicots
(Rank)	Rosids
Order	Fabales
Family	Leguminosae
Subfamily	Fabaceae
Genus	Cicer
Species	arietinum L.

Locality (In Gujarat State)	Cultivated; also an escape.
Local Name (In Gujarati Language)	Chana
Fruit	October- March.
Flower	October- March.

USES
- Chickpeas and Bengal grams are used to make curries and are one of the most popular vegetarian foods in India. - On the Indian subcontinent, green chickpeas are called Chana in Gujarati,where they are a major source of protein in a mostly vegetarian culture. - Typically Chana in Hindi and Punjabi might refer to both varieties, as might chhole, but the former is more the green hard small variety while the latter is the large creamy softer one and also the more popular dish served around the region at home and at celebrations.

GALLERY

图3物种信息检索工具的结果页

Figure 3 Result page of Species Information Retrieval Tool

2.3.1豆科植物工具

在印度古吉拉特邦发现这个家庭有超过250个种类，来自NCBI数据库的约149种豆科家庭物种信息已经被收集和编制。豆科分为3个亚科：蝶形花科(Papilionaceae)、含羞草科和苏木科。

在这个“Leguminoase工具”中，用户必须选择相应的选项，单击提交按钮后，它将直接从NCBI数据库获取各种信息，比如它的物种名称、PubMed、Pubmed Central、核苷酸等。

用户必须在GDB数据库的左侧选择Bioinformatics Information选项，然后单击Leguminoase工具。图8显示了点击提交按钮之后选择的一个种类，它给出了各种选项(图9)，无论点击什么选项，它都将从NCBI数据库获取数据到这个VLDB数据库。图10显示了该特定物种的一个Pubmed选项。


GLDB

Gujarat State Leguminosae Family Database

Distribution of Leguminosae Family Members in Gujarat State

- *About Gujarat State* : On May 1, 1960, Gujarat was created out of 17 northern districts of the former state of Bombay Located on the western coast of India, has the longest coastline of 1,600 Km.

  - Population of approximately 6.03 Crore (4.99% of Indian Population)
  - Gandhinagar, the capital city of Gujarat is located close to Amdavad (Ahmedabad), the commercial capital
  - Ahmedabad is the most populated District in the State, with 7.20 million people
  - Diverse climatic conditions with mild and pleasant winters and hot and dry summers and heavy monsoon.
  - The state currently has 26 districts (226 talukas, 18,618 villages, 242 towns)

- *Region wise Distribution* : It includes Information of Leguminosae Family members present in Gujarat state. In addition, Region wise web pages created which gives information of those species which are found in that particular region.

For Further Information, Click on Respective Region :

1. *Kutch:* It is a largest district of Gujarat state, covering an area of 45,652 km².
2. *Saurashtra* : The Saurashtra or Kathiwar region comprises the south western part of Gujarat state and the districts included in this region are Rajkot, Junagadh, Bhavnagar, Porbandar, Jamnagar, Amreli, Surendranagar.
3. *North Gujarat* : The North part of Gujarat is called North Gujarat. It includes Gandhinagar, Banaskantha, Sabarkantha, Mehsana and Patan districts.
4. *Central Gujarat* : This region is located in central region in Gujarat state which includes Nadiad, Anand, Vallabh Vidyanagar, Vasad, Vadodara, Rajpipla districts.
5. *South Gujarat* : The Districts in this region are Surat district, Navsari district, Dang district, Valsad district, and the newly formed Tapi district. The main cities other than Surat are Navsari, Vyara, Valsad, Bardoli, Vapi, Bilimora, Chikhli, Songadh.
6. *Other Region* : This region includes species are found in throughout, wild in deciduous forest; also planted in gardens, scrub forest etc.

图4古吉拉特邦和豆科家族区域分布主页

Figure 4 Home Page of Gujarat State and Leguminosae Family with region wise distribution


GLDB

Gujarat State Leguminosae Family Database

Distribution of Leguminosae Family Members in Gujarat State

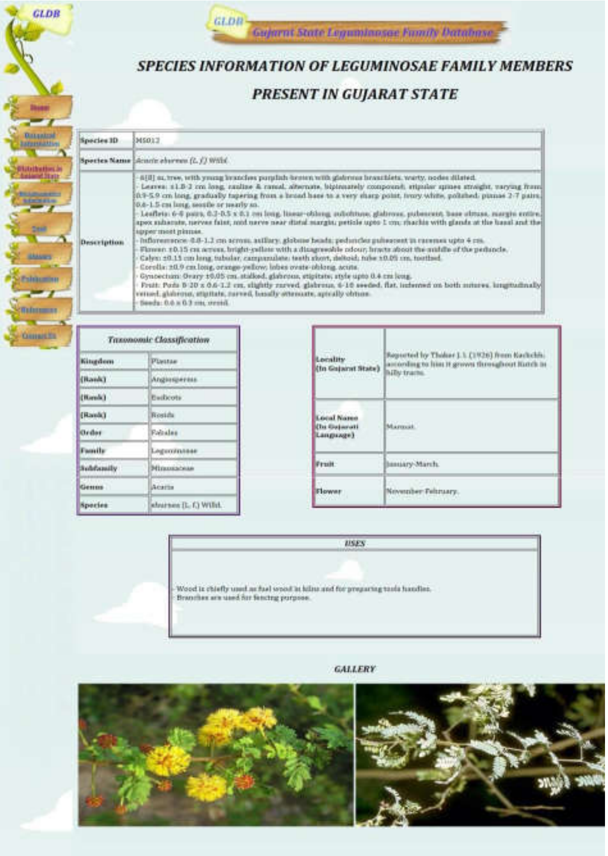
Region : Kutch

List of Species Found at Kutch

Sub-Family Name	Species Name	Location
Caesalpiniareace	*Cassia sophera L. var. sophera.*	Kutch
Fabaceae	*Alysicarpus hamosus W. & A. Prodr.*	Kutch
Fabaceae	*Astragalus prolixus Sieb.*	Kutch- Mundra
Fabaceae	*Canavalia maritime (Aubl.) Thou.*	Kutch
Fabaceae	*Crotalaria antonii Wt. & Arn.*	Kutch,in cultivated fields.
Fabaceae	*Crotalaria evolvuloides Wt. ex Wt. & Arn.*	Kutch
Fabaceae	*Crotalaria filipes Benth. var. filipes*	Kutch
Fabaceae	*Crotalaria hirsuta Willd.*	Kutch
Fabaceae	*Indigofera caerulea Roxb. var. monosperma Sant.*	Kutch
Fabaceae	*Indigofera hochstetteri*	Kutch
Fabaceae	*Lotus garcini DC.*	Kutch

图5 Kutch地区信息

Figure 5 Kutch region Information



SPECIES INFORMATION OF LEGUMINOSAE FAMILY MEMBERS PRESENT IN GUJARAT STATE

Species ID	MS012
Species Name	*Acacia eburnea (L. f.) Willd.*

Taxonomic Classification	
Kingdom	Plantae
(Rank)	Angiosperms
(Rank)	Eudicots
(Rank)	Rosids
Order	Fabales
Family	Leguminosae
Subfamily	Mimosaceae
Genus	Acacia
Species	eburnea (L. f.) Willd.

Locality (In Gujarat State)	Reported by Thaker J. I. (1926) from Kachchh; according to him it grows throughout Kutch in hilly tracts.
Local Name (In Gujarati Language)	Marmat.
Fruit	January-March.
Flower	November-February.

USES

- Wood is chiefly used as fuel wood in kilns and for preparing tools handles.
- Branches are used for fencing purpose.

图6在Kutch地区发现的物种信息

Figure 6 Information of species found in Kutch region

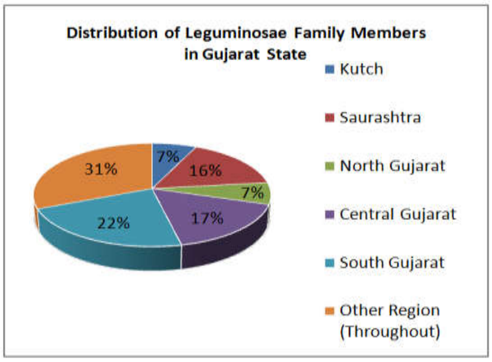


图7豆科家族的分布信息饼图

Figure 7 Information of Distribution of Leguminosae family as pie chart

2.3.2 ConSeq工具

使用这个工具时，用户必须提供rbcL或matK蛋白质序列作为输入项，然后输出结果，这与用户的序列长度有关，在该序列中发现保守序列决定该物种是否包括在豆科家族的三个亚家族中以及rbcL或matK蛋白序列。用户必须单击GDB数据库左侧的工具选项重定向到ConSeq工具的主页(图11)。如果在该序列中发现保守区，它会显示如图12的结果。但如果在序列中找不到保守区域，则ConSeq工具输出的结果如图13

表1豆科家族的分布信息

Table 1 Information of Distribution of Leguminosae family

Location	Region specific Species
Kutch	25
Saurashtra	60
North Gujarat	25
Central Gujarat	61
South Gujarat	82
Other Region (Throughout)	115

表2豆科工具信息

Table 2 Information of Leguminobase tool

Sub family	DNA	Protein	SRA
Fabaceae (Papilionaceae)	104	67	15
Caesalpiniaceae	26	26	-
Mimosaeae	19	14	-
TOTAL	149	107	15

图8豆科植物物种选择工具

Figure 8 Leguminobase tool species selection

2.3.3 RNA-Seq数据分析

从头意味着组装短的片段以获得全长(有时是新的)的序列，从头测序包括第一次对新的基因组测序并需要测序读数的专业装配。读取长度，读取深度和灵活的成对末端插入物大小的独特组合使得Illumina成为理想的从头测序仪。无与伦比的原始读取精度使得高质量，长重叠群装配实现高效的生产。使用Illumina和Roche 454分析五种不同物种的数据，对数据集的详细分析提供了五种物种的几个重要特征信息，例如GC含量。豆类和其他植物物种保守基因通过GO术语分配功能类别以及通过

MISA工具识别SSR。

BIOINFORMATICS INFORMATION FROM NCBI DATABASE

Species Name:	*Cicer arietinum L.*

Literature :

PubMed	Scientific & medical abstracts/citations
Pubmed Central	full-text journal articles

Nucleotide Sequences :

Nucleotide	DNA and RNA sequences
SRA	High-throughput DNA and RNA sequence read archive
PopSet	Sequence sets from phylogenetic and population studies

Genomes :

Genome	Genome sequencing projects by organism
BioProject	Biological projects providing data to NCBI

Proteins :

Protein	Protein sequences
Structure	Experimentally-determined biomolecular structures

图9豆科工具结果页

Figure 9 Result page of Leguminobase tool

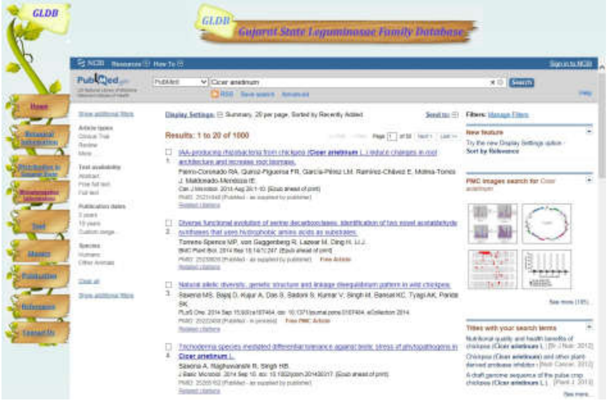
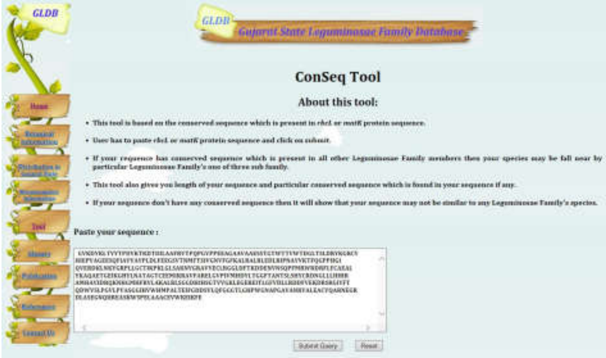


图10 GLDB数据库中NCBI数据库Pubmed信息的屏幕截图

Figure 10 Screenshot of Pubmed information from NCBI database in GLDB database



ConSeq Tool

About this tool:

- This tool is based on the conserved sequence which is present in rbcL or matK protein sequence.
- User has to paste rbcL or matK protein sequence and click on submit.
- If your sequence has conserved sequence which is present in all other Leguminosae Family members then your species may be fall near by particular Leguminosae Family's one of three sub family.
- This tool also gives you length of your sequence and particular conserved sequence which is found in your sequence if any.
- If your sequence don't have any conserved sequence then it will show that your sequence may not be similar to any Leguminosae Family's species.

Paste your sequence :

图11 ConSeq工具的屏幕截图

Figure 11 Screen shot of ConSeq Too

值得注意的是，对五种不同豆科物种Arachis hypogaea L.、 Cicer arietinum L.、Phaseolus vulgaris L.、Trigonella foenum-graecum L.、和Vicia sativa L.的研究将有助于进一步进行功能基因组学研究，因为它包括具有完整注释的每个物种的有用信息。图14表示De novo RNA-Seq数据页的主页。首先用户需要选择一个平台Roche 454或Illumina，然后在第二个选项中选择物种，最后选择重叠群。单击提交按钮后，它会给出如图15所示的完整的重叠群信息。图15为Arachis hypogea L.的重叠群数10017的结果。结果显示从数据库中获得的一个重叠群的各种不同的信息，有用的信息：物种名称、重叠群名称、Fasta序列、序列长度、比对的最小E值、比对的相似性意义、比对的GO号、比对上可能性较高的物种、比对结果的描述(HSP)、比对上基因的名称、比对目标的登录号、比对结果的E值、比对的长度、比对匹配上的序列长度、比对上的活性部分，比对上的相似度、比对的描述/重合、比对的描述/查询、比对上的查询帧、比对上的数据库、比对上的得分、

ConSeq Tool - RESULT

- Your Sequence is:
- Length of Your Sequence is: 504
- Your Species Sequence may not contain Conserved sequence which is found in Leguminosae Family Species.

图12 ConSeq工具的结果页

Figure 12 Result page of ConSeq Tool

ConSeq Tool - RESULT

- Your Sequence is:
- Length of Your Sequence is: 504
- Your Species Sequence may not contain Conserved sequence which is found in Leguminosae Family Species.

图13 未找到保守序列时ConSeq工具的结果页

Figure 13 Result page of ConSeq Tool when no conserved sequence found

比对上的GOs、GO的加入、GO名称、酶代码、Ids进程、GO加入的进程、GO名称的进程、主题细节的进程、主题匹配的进程、比对上的数量、酶，如果有重叠群参与通路，就会产生KeggMaps，其次是KEGG通路图像。总共有82,505条重叠群插入的记录，包括10824个Arachis hypogaea L.的重叠群，34678个Cicer arietinum L.的重叠群，6999个Phaseolus vulgaris L.的重叠群，7256个Trigonella foenum-graecum L.的重叠群以及22748个Vicia sativa L.的重叠群。

图14 De novo RNA-Seq数据检索的主页

Figure 14 Home page for De novo RNA-Seq data retrieval

3词汇表

讨论了各种于鉴定植物物种有用的植物术语，用户需要选择GDB数据库中左侧的词汇表选项。

4出版物

本文包含作为此数据库结果的出版物列表。

5参考文献

本文包含创建此数据库的参考论文、在线网站和图书等的列表。

6联系我们

本文包含联系我们表格，用户可以向我们发送任何问题。

7结论

设计古吉拉特邦的豆科家庭成员的数据库时考虑了以下几点：

1桥接植物信息与生物信息学信息以及分析。

2利用生物信息学的各种工具分析豆科家族物种的目的。

3通过各种工具和软件的帮助，从上述工作中获取次级信息。

4以数据库的形式向公众提供生物信息学信息。

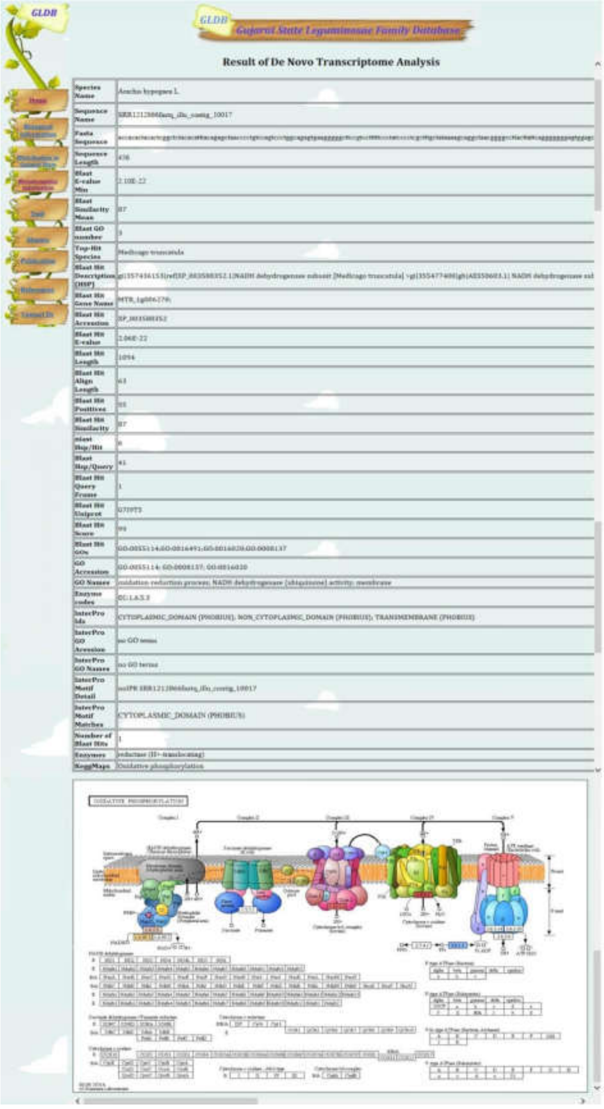


图15 de novo RNA-Seq的结果页

Figure 15 Result page of de novo RNA-Seq

命名为“古吉拉特邦豆科家庭数据库(GLDB)”的豆科家族的综合数据库创建了每个豆科家族的有用信息，此数据库中开发了许多内置工具以获得特定物种的信息，比如古吉拉特州每个种类的豆科家族在谷歌地图上的分布的完整的信息。生物信息学包括许多工具，比如一个工具被设计用来获得特定物种来自NCBI数据库的DNA、PROTEIN、GENOME等信息。ConSeq工具被用来找出保守的序列，5个豆科家族的RNA-Seq数据分析做了序列拼装及注释。豆科家族的这个特殊数据库，为当前植物学科学界的需求提供服务。因为也需要服务于

他们的需要，到目前为止，这样的信息在一个平台是不可用的。豆科家族各种分散数据以一种任何人希望找到这些特定物种的信息只要单击鼠标就可以获得的方式储存。

参考文献

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

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

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. & King, C. 1997. Common Families of Flowering Plants. Cambridge: Cambridge University Press

J. L. Collins, J. P. Biggs, C. Voelckel and S. Joly, 2008, An approach to transcriptome analysis of non-model organisms using short-read sequences, Genome Informatics 21:3-14

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

Jianan Zhang, Shan Liang, Jialei Duan, Jin Wang, Silong Chen, Zengshu Cheng, Qiang Zhang, Xuanqiang Liang and Yurong Li, 2012, De novo assembly and Characterisation of the Transcriptome during seed development, and generation of genic-SSR markers in Peanut (Arachis hypogaea L.), BMC Genomics 2012 13:90

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

Libault, M., Joshi, T., Benedito, V.A., Xu, D., Udvardi, M.K., and Stacey, G., 2009, Legume Transcription Factor Genes: What makes legumes so special?. Plant Physiology 151: 991-1001

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 5(7): 621-8

Ness, R.W., Siol, M., and Barrett S.C.H., 2011, De novo sequence assembly and characterization of the floral transcriptome in cross and self-fertilizing plants, BMC Genomics 12: 298

Patel RK, Jain M, 2012, NGS QC Toolkit: A Toolkit for Quality Control of Next Generation Sequencing Data, PLoS ONE 7(2): e30619. doi:10.1371/journal.pone.0030619

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

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

Rohini Garg, Ravi K. Patel, Akhilesh K. Tyagi, and Mukesh Jain., 2011, De Novo Assembly of Chickpea Transcriptome Using Short Reads for Gene Discovery and Marker Identification, DNA RESEARCH 18, 53–63; doi:10.1093/dnares/dsq028

Sagar Patel, 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: Vol 4, Issue3, 2013, Pages. 54-59. ISSN 0976-2604. Online ISSN 2278-6007

Sagar Patel, 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, Pages- 2184-2189 Vol 2 Issue 4 April 2013, ISSN 2278-733X

Sagar Patel, 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: Vol.1 n.2 Pages.187-194 June, 2013 Print ISSN 2319-5169

Sagar Patel, Shah D., Panchal H., Comparative study of five Legume species based on De Novo Sequence Assembly and Annotation, Computational Molecular Biology, Vol.4, No.9, 1-6 (doi: 10.5376/cmb.2014.04.0009)

Sagar Patel, Shah D., Panchal H., Conseq Tool: A Tool to Find Conserved Region in Protein Sequences of Leguminosae Family. Journal of Advanced Bioinformatics Applications and Research, Vol 5, Issue3, 2014, pp134-139, ISSN

0976-2604.Online ISSN 2278–6007

Sagar Patel, Shah D., Panchal H., De Novo RNA Seq Assembly and Annotation of Cicer arietinum L. (SRR627764). Legume Genomics and Genetics, 2014, Vol. 5, No. 6. (doi: 10.5376/lgg.2014.04.0006)

Sagar Patel, Shah D., Panchal H., De Novo RNA Seq Assembly and Annotation of Phaseolus vulgaris L. (SRR1283084), Genomics and Applied Biology, Vol.5, No.5, 1-6 (doi: 10.5376/gab.2014.05.0005)

Shi, C.Y., Yang, H., and Wei, C.L., 2011, Deep sequencing of the Camellia sinensis transcriptome revealed candidate genes for major metabolic pathways of tea-specific compounds, BMC Genomics 12: 131

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

Vaidya K., Ghosh A., Kumar V, Chaudhary S, Srivastava N, Katudia K, Tiwari T and Chikara K., 2012, De novo transcriptome sequencing in Trigonella foenum-graecum to identify genes involved in the biosynthesis of diosgenin. The Plant Genome:doi: 10.3835/ lantgenome2012.08.0021

Wang, X.W., Luan, J.B., Li, J.M., Bao, Y.Y., Zhang, C.X., and Liu, S.S., 2010, De novo characterization of a whitefly transcriptome and analysis of its gene expression during development, BMC Genomics 11: 400

Wang, Z., Gerstein, M., and Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics, Nat Rev Genet. 10(1): 57-63

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

http://www.blast.ncbi.nlm.nih.gov/Blast.cgi

http://www.blast2go.com/b2ghome

http://www.clcbio.com/products/clc-genomics-workbench/

http://www.genome.jp/kegg/pathway.html

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

http://www.ildis.org/

http://www.kew.org/

http://www.missouribotanicalgarden.org/

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

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

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

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

http://www.nipgr.res.in/ngsqctoolkit.html

http://www.pgrc.ipk-gatersleben.de/misa/misa.html

http://www.plantgrn.noble.org/PlantTFcat/

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

http://www.trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software

http://www.wego.genomics.org.cn