生物数据库-以关键词为主的生物信息 数据库检索

日期: 2022-10-26

实验者: 生信 2001 张子栋

MarkdownNotes/软件第1次作业.md at main · Bluuur/MarkdownNotes

(github.com)

生物信息学原理/软件第1次作业.md·blur/MarkdownNotes - 码云 - 开源中国

(gitee.com)

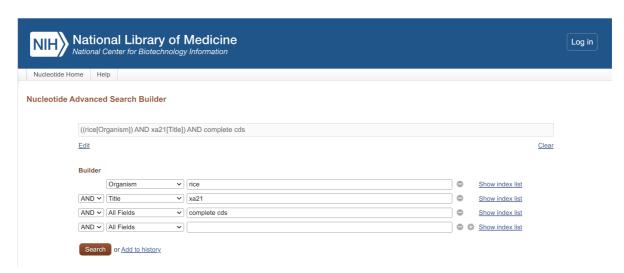
实验目的与内容

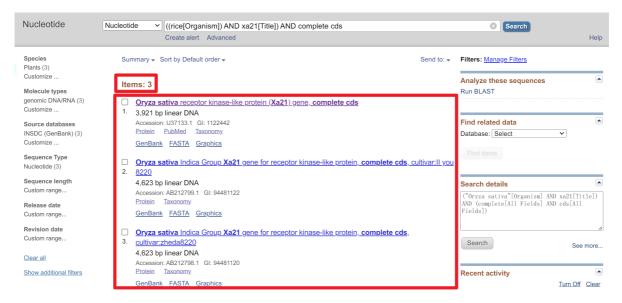
- 熟悉 NCBI 数据库 Entrez 检索系统
- 会使用关键词检索 NCBI、UnitProtKB、PubMed 等数据库,能理解检索结果页面各条目含义
- 了解文献管理工具

实验流程和结果

- 1. 查找与水稻抗病金银 Xa21 有关的资料 (GenBank, UnitProt 等)
 - 1. 有多少条序列具有全长 CDS, 分别由多少碱基构成? 编码多少个氨基酸?
 - 2. 选择修改时间最早的一条序列,指出该基因 exon 和 intron 的位置。该基因编码的蛋白质有多少个氨基酸残基?该蛋白质可能位于细胞中的哪个位置? 是否有三维结构信息?

1.1:





有 3 条序列具有全长 CDS

第1条序列:

- 有 3921 个碱基
- 编码 1025 个氨基酸



receptor kinase-like protein [Oryza sativa Indica Group]

GenBank: AAC49123.1

Identical Proteins FASTA Graphics

编码 1025 个氨基酸 Go to: ✓ 1025 aa LOCUS linear PLN 14-DEC-1995 DEFINITION receptor kinase-like p rotein [Oryza sativa Indica Group]. ACCESSION AAC49123 VERSION AAC49123.1 locus OSU37133 accession <u>U37133.1</u> DBSOURCE KEYWORDS SOURCE Oryza sativa Indica Group (long-grained rice) ORGANISM <u>Oryza sativa Indica Group</u> Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; BOP clade; Oryzoideae; Oryzeae; Oryzinae; Oryza; Oryza sativa. REFERENCE 1 (residues 1 to 1025) AUTHORS Song, W.-Y., Wang, G.-L., Chen, L.-L., Kim, H.-S., Pi, L.-Y., Holsten, T., Wang, G.B., Zhai, W.-X., Zhu, L.-H., Fauquet, C. and Ronald, P. TITLE A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21 JOURNAL Science 270 (5243), 1804-1806 (1995) 8525370 PUBMED REFERENCE 2 (residues 1 to 1025) AUTHORS Ronald, P., Song, W.-Y., Wang, G.-L., Kim, H.-S. and Pi, L.-Y. TITLE Direct Submission

第2条序列:

- 有 4623 个碱基
- 编码 1025 个氨基酸

GenBank → Send to: •

Oryza sativa Indica Group Xa21 gene for receptor kinase-like protein, complete cds, cultivar:II you 8220

GenBank: AB212799.1

FASTA Graphics

```
Go to: ✓
                                   4623 bp
                                                       linear PLN 14-FEB-2008
DEFINITION Oryza sativa Indica
                                                       ceptor kinase-like
           protein, complete cds, cultivar:II you 8220.
ACCESSION
           AB212799
VERSION
            AB212799, 1
KEYWORDS
SOURCE
           Oryza sativa Indica Group (long-grained rice)
  ORGANISM Oryza sativa Indica Group
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; BOP
            clade; Oryzoideae; Oryzeae; Oryzinae; Oryza; Oryza sativa.
REFERENCE
  AUTHORS
           Xue, Q. and Xu, J.
```

receptor kinase-like protein [Oryza sativa Indica Group]

GenBank: BAE93934.1

Identical Proteins FASTA Graphics



clade; Oryzoideae; Oryzeae; Oryzinae; Oryza; Oryza sativa.

第3条序列:

- 有 4623 个碱基
- 编码 1025 个氨基酸

Oryza sativa Indica Group Xa21 gene for receptor kinase-like protein, complete cds, cultivar:zheda8220

GenBank: AB212798.1

FASTA Graphics

```
Go to: ✓
LOCUS
                                               DNA
                                                       linear PLN 14-FEB-2008
DEFINITION Oryza sativa Indica Gr
                                              ne for receptor kinase-like
            protein, complete cds, cultivar:zheda8220.
ACCESSION
            AB212798
VERSION
            AB212798. 1
KEYWORDS
            Oryza sativa Indica Group (long-grained rice)
SOURCE
           <u>Oryza sativa Indica Group</u>
 ORGANISM
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; BOP
            clade; Oryzoideae; Oryzeae; Oryzinae; Oryza; Oryza sativa.
REFERENCE
           Xue, Q. and Xu, J.
  TITLE
            Xa21 gene with bacterial blight resistance from wild rice (Oryza.
            longistaminata) to hybrid rice(Oryza.sativa)
  JOURNAL
           Unpublished
REFERENCE
           2 (bases 1 to 4623)
  AUTHORS
           Qingzhong, X. and Jianhong, X.
  TITLE
            Direct Submission
  JOURNAL
           Submitted (02-MAY-2005) Contact: Xue Qingzhong zhejiang university,
            agronomy; kaixuan road 268, hangzhou, zhejiang 310029, China
```

receptor kinase-like protein [Oryza sativa Indica Group]

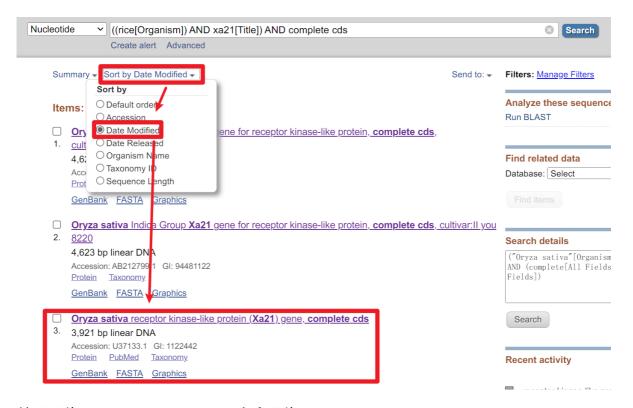
GenBank: BAE93933.1

Identical Proteins FASTA Graphics



1.2:

找到修改时间最早的一条序列



外显子为 1-2677, 3521-3921; 内含子为 2678-3520

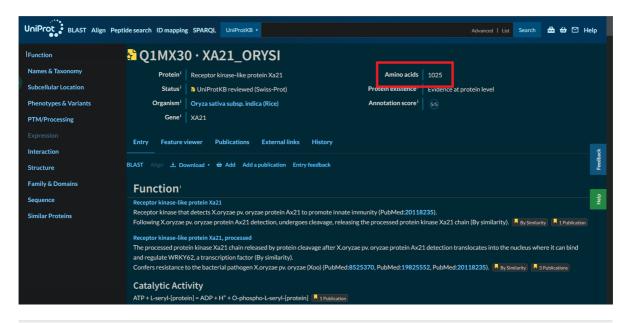
```
/map="11q, RG103"

1..3921
/gene="Xa21"

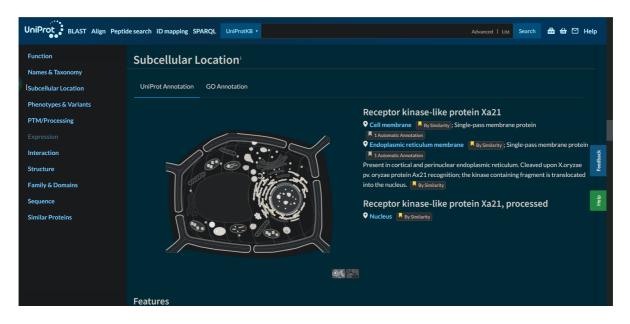
join(1..2677, 3521..3921)

/gene- xaz1
/note="Xa21 disease resistance gene"
/codon_start=1
/product="receptor kinase-like protein"
/protein_id="AAC49123.1"
/translation="MISLPLLLFVLLFSALLLCPSSSDDDGDALYQGGQSLASWNTSGHGQHCTWVGVVCGRRRRRHPHRVVKLLLRSSFLRELDLGDNYLSGEIPPELSRLSRLQLLELSDNSIQGSIPAL
```

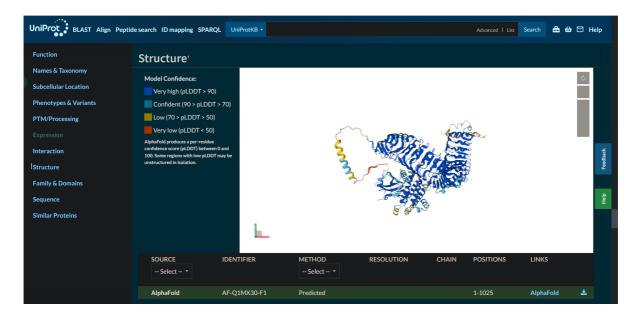
氨基酸残基数 1025



亚细胞定位:细胞膜,内质网膜

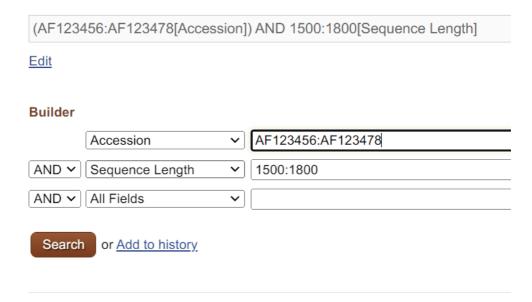


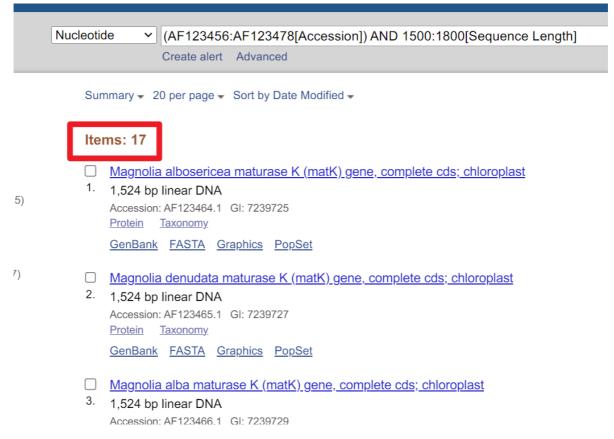
三维结构信息



2. 检索注册号在 AF123456-AF123478 之间并且序列长度在 1500 到 1800 bp 之间的核苷酸数据,共有多少条?如何批量下载它们 FASTA 的序列? 共 17 条

Nucleotide Advanced Search Builder



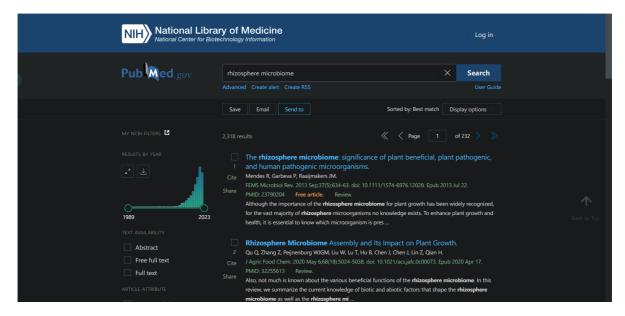


勾选后点击 Send to,格式选择 FASTA

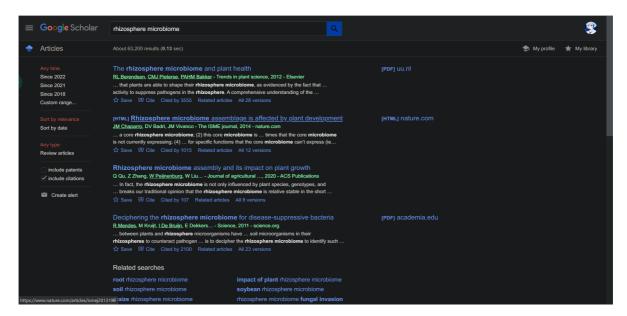


3. 通过关键词,在 PubMed 等数据库搜索跟你「项目育人」项目相关的文献

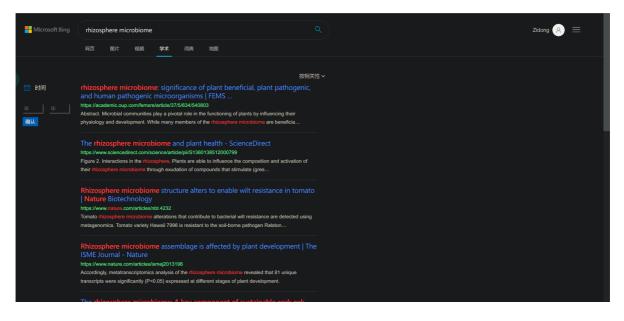
PubMed:



Google Scholar



Bing Scholar





在本次上机实验中,熟悉了几种生信数据库的查询操作以及文献检索。