实用生物信息技术课程总结交流会

SRS应用初探

2008年6月22日

报告人：颜林林(yanlinlin82@gmail.com)

其他组员：颜志辉、郭凤璔、王诵涛

主要内容

1.SRS简介

2.SRS使用实例

3.SRS高级应用技巧

4.总结和展望

SRS简介

- Sequence Retrieval System，意即“序列提取系统”
- 生物信息数据库检索系统
- 集“查询、分析、保存”于一体的信息平台

SRS产生背景

- 生物信息种类繁多且呈指数增长
- 大多数信息以文本文件存储
- 各数据库的查询功能有限
- 难以进行跨数据库的查询

文本文件存储

- 利于不同类型数据的扩展
- 利于迅速增长的数据的存储和维护
- 方便计算机处理
- 能够直接供人阅读

EMBL的数据格式

```
ID   Z66500; SV 1; linear; genomic DNA; STD; INV; 33780 BP.
XX
AC   Z66500;
XX
DT   25-OCT-1995 (Rel. 45, Created)
DT   02-JUN-2008 (Rel. 96, Last updated, Version 86)
XX
DE   Caenorhabditis elegans Cosmid T05C12
XX
KW   HTG.
XX
OS   Caenorhabditis elegans
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
XX
RN   [1]
RP   1-33780
RX   DOI; 10.1126/science.282.5396.2012
RX   PUBMED; 9851916.
RG   WormBase Consortium
RA   ;
RT   "Genome sequence of the nematode C. elegans: a platform for investigating
RT   biology";
RL   Science 282(5396):2012-2018(1998).
XX
RN   [2]
```

GeneBank的数据格式

```
LOCUS       NM_063315               1503 bp    mRNA    linear   INV 07-APR-2008
DEFINITION  Caenorhabditis elegans AcetylCholine Receptor family member
            (acr-14) (acr-14) partial mRNA.
ACCESSION   NM_063315
VERSION     NM_063315.1  GI:17531182
KEYWORDS    .
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 1503)
  CONSRTM   The C.elegans Sequencing Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-2003) Nematode Sequencing Project, Wellcome Trust
            Sanger Institute, Hinxton, Cambridge CB10 1SA, UK, and Genome
            Sequencing Center, Washington University, St. Louis, MO 63110, USA.
            E-mail: sequence@wormbase.org
COMMENT     REVIEWED REFSEQ: This record has been curated by WormBase. This
            record is derived from an annotated genomic sequence (NC_003280).
            The reference sequence was derived from T05C12.2.
            Expression: acr-14 encodes a protein that contains
            neurotransmitter-gated ion-channel ligand-binding and transmembrane
            domains.

            RNAi results: [Rual JF] Not abnormal based on scoring phenotypes
            postembryonic_development_abnormal, lethal. [Sonnichsen B] Not
            abnormal based on scoring phenotypes embryonic_lethal,
            organism_morphology_abnormal, maternal_sterile. [Kamath RS] Not
```

SRS的特点

- 查询针对文本文件数据
- 提供统一的网页操作界面
- 上千个数据库与上百个分析工具
- 能同时关联查询多个数据库
- 任意组合各种查询条件
- 可直接对查询结果进行分析

SRS与BLAST

- 两者都可以查询和提取序列数据
- SRS根据序列的不同字段信息进行关键词查找
- BLAST通过序列相似性比对算法来查找
- SRS通过整合BLAST工具，也可以实现BLAST功能

公用SRS服务器

- EMBL (http://srs.embl.de/srs/)
  - SRS 8.2，102个库，6个工具
- DKFZ (http://www.dkfz-heigelberg.de/srs)
  - SRS 7.13，769个库
- EBI (http://srs.ebi.ac.uk)
  - SRS 7.13，112个库，161个工具

(信息来自：http://downloads.biowisdomsrs.com/publicsrs.html）

2、SRS使用实例

1.SRS快速查询

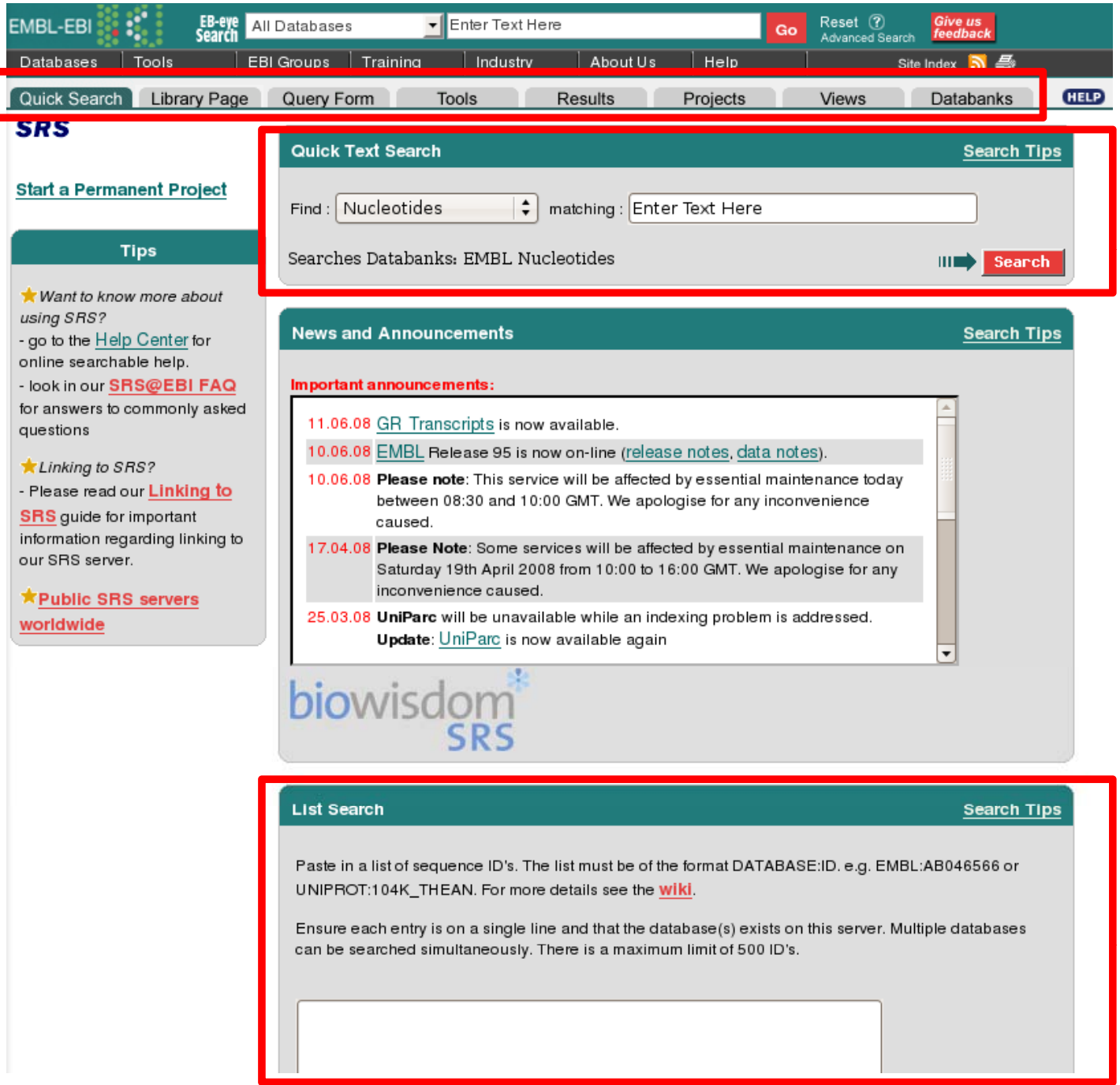
2.SRS标准查询

3.SRS扩展查询

4.保存查询结果

EBI-SRS网站首页

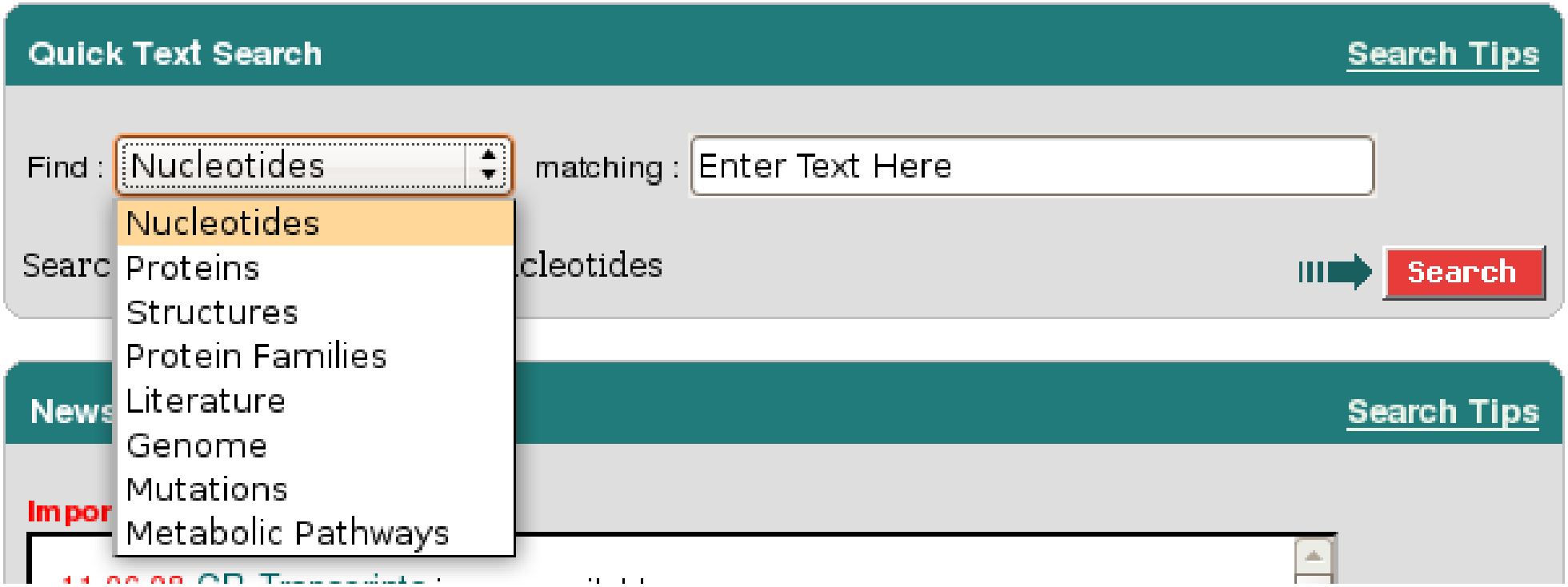
工具栏



快速查询

批量提取

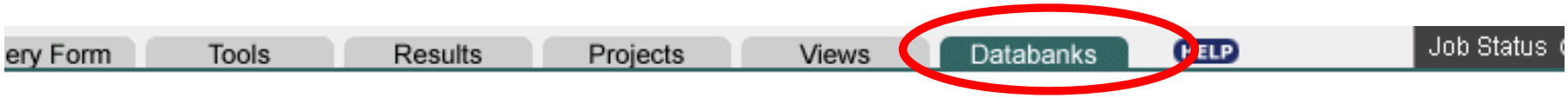
SRS快速查询


Quick Text Search
Search Tips
Find :
Nucleotides
matching :
Enter Text Here
Nucleotides
Proteins
Structures
Protein Families
Literature
Genome
Mutations
Metabolic Pathways
Search
Search Tips

快速查询的类型

选项	含义	数据库
Nucleotide	核酸序列	EMBL EMBLCON EMBLCDS EMBLANN
Proteins	蛋白序列	UNIPROT UNIPARC
Structures	蛋白结构	PDB
Protein Families	蛋白家族	INTERPRO
Literature	文献	MEDLINE
Genome	基因组	ENTREZGENE
Mutations	突变	OMIM
Metabolic Pathways	代谢途径	PATHWAY

EBI-SRS的数据库列表

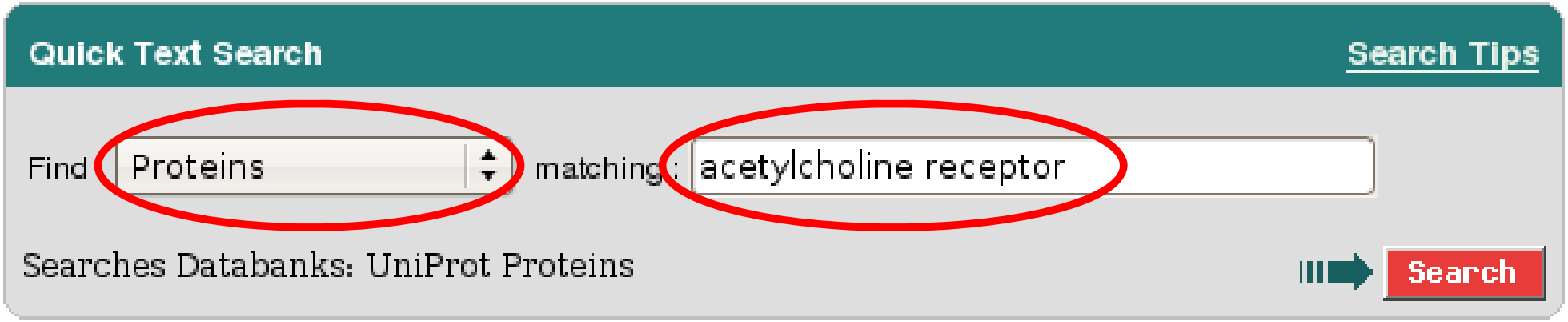


Databank Information

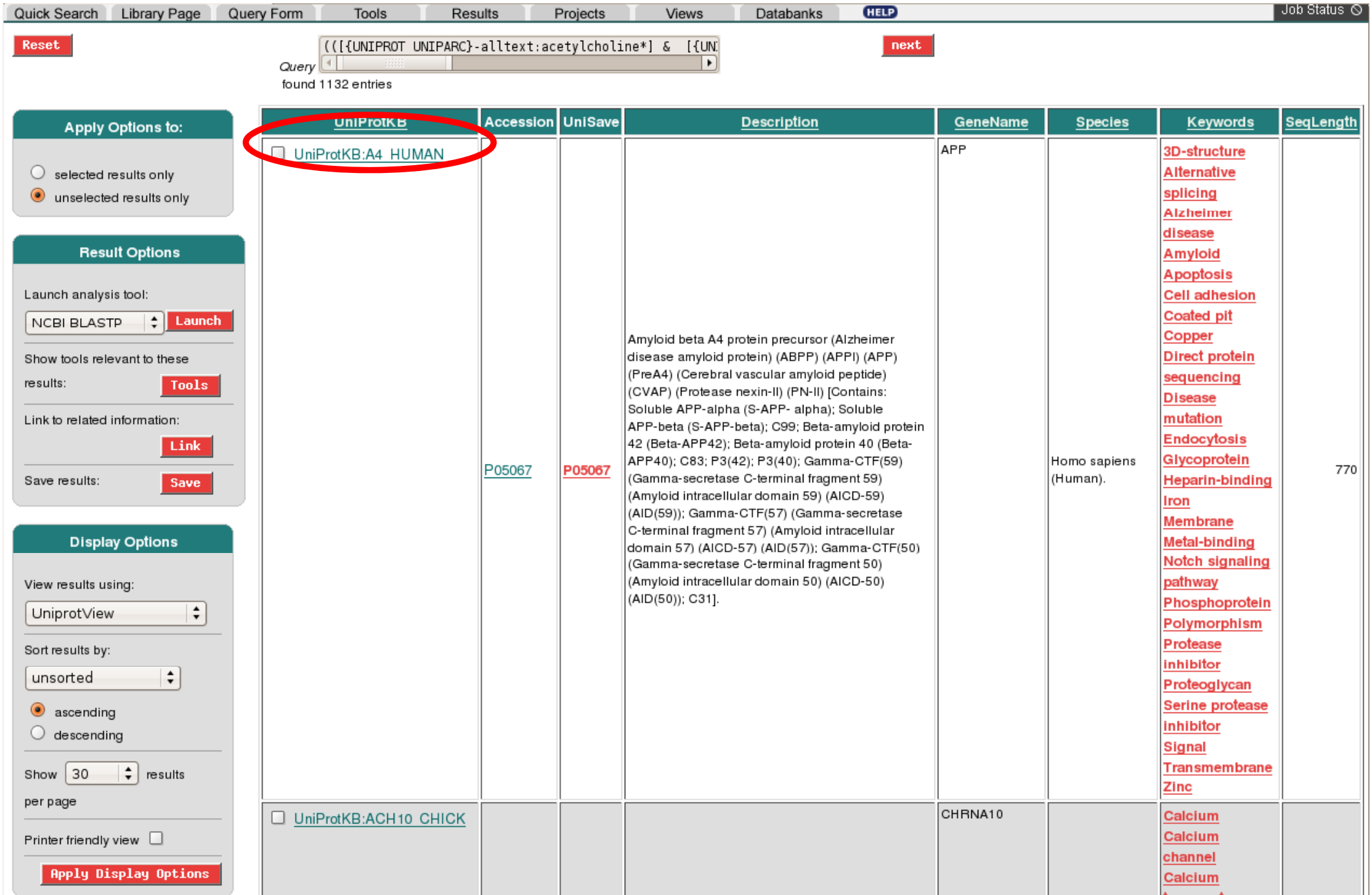
Databank	Release	No. of Entries	Indexing Date	Group	Availability
MEDLINE		virtual databank		References	
TAXONOMY		486851	18-Jun-2008	References	OK
OMIM		19622	17-Jun-2008	References	OK
OMIMMORBIDMAP		4892	12-Jun-2008	References	OK
PATABS		1500000	28-Mar-2008	References	OK
KG		274	10-Apr-2008	References	OK
MEDLINENEW		1698810	18-Jun-2008	ReferencesSub	OK
MEDLINE2008		16880015	19-Mar-2008	ReferencesSub	OK
MED2PUB		14116195	18-Mar-2008	ReferencesSub	OK
UNILIB		23715	18-Jun-2008	GENEDICT	OK
LOCUSLINK		287853	19-Mar-2008	GENEDICT	OK
UNIGENE		1784319	14-Jun-2008	GENEDICT	OK
HGNC		28100	30-Mar-2008	GENEDICT	OK
ENTREZGENE		4431579	14-Jun-2008	GENEDICT	OK
GOA		30154039	28-May-2008	Ontology	OK
GO		24708	18-Mar-2008	Ontology	OK
EMBL		virtual databank		DNASequence	

SRS快速查询实例

以研究“乙酰胆碱受体蛋白”为例



查询结果



Quick Search | Library Page | Query Form | Tools | Results | Projects | Views | Databanks | HELP | Job Status

Reset

Query ([{UNIPROT_UNIPARC}-alltext:acetylcholine*] & [{UN

found 1132 entries

next

Apply Options to:

- selected results only
- unselected results only

Result Options

Launch analysis tool: NCBI BLASTP Launch

Show tools relevant to these results: Tools

Link to related information: Link

Save results: Save

Display Options

View results using: UniprotView

Sort results by: unsorted

- ascending
- descending

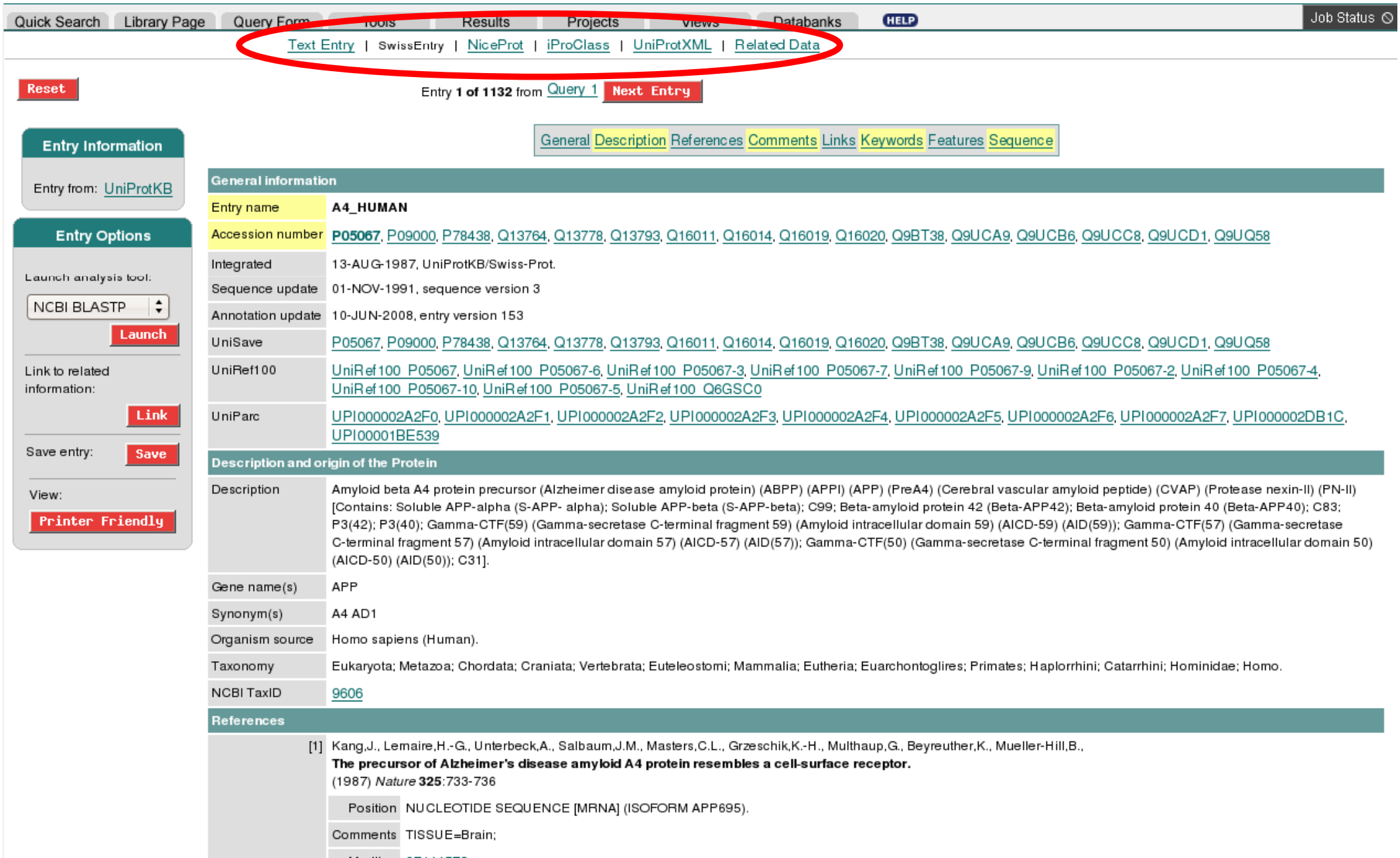
Show 30 results per page

Printer friendly view

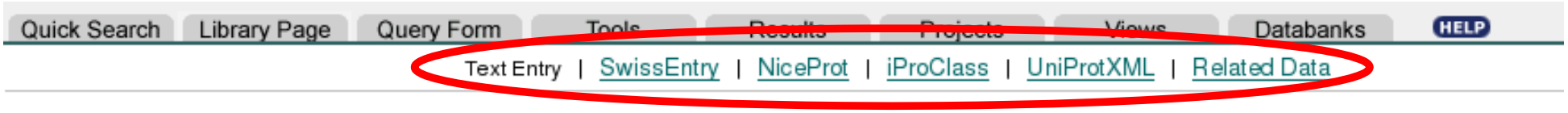
Apply Display Options

UniProtKB	Accession	UniSave	Description	GeneName	Species	Keywords	SeqLength
UniProtKB:A4_HUMAN	P05067	P05067	Amyloid beta A4 protein precursor (Alzheimer disease amyloid protein) (ABPP) (APPI) (APP) (PreA4) (Cerebral vascular amyloid peptide) (CVAP) (Protease nexin-II) (PN-II) [Contains: Soluble APP-alpha (S-APP- alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59) (Amyloid intracellular domain 59) (AICD-59) (AID(59)); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57) (Amyloid intracellular domain 57) (AICD-57) (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain 50) (AICD-50) (AID(50)); C31].	APP	Homo sapiens (Human).	3D-structure; Alternative splicing; Alzheimer disease; Amyloid; Apoptosis; Cell adhesion; Coated pit; Copper; Direct protein sequencing; Disease mutation; Endocytosis; Glycoprotein; Heparin-binding; Iron; Membrane; Metal-binding; Notch signaling pathway; Phosphoprotein; Polymorphism; Protease inhibitor; Proteoglycan; Serine protease inhibitor; Signal; Transmembrane; Zinc	770
UniProtKB:ACH10_CHICK				CHRNA10		Calcium; Calcium channel; Calcium	

查询结果条目(Entry Page)



结果条目的不同显示方式



Reset

Entry **1 of 1132** from Query 1 Next Entry

Entry Information

Entry from: UniProtKB

Entry Options

Launch analysis tool:

NCBI BLASTP

Launch

Link to related information:

Link

Save entry: Save

View:

Printer Friendly

```
ID   A4_HUMAN                Reviewed;         770 AA.
AC   P05067; P09000; P78438; Q13764; Q13778; Q13793; Q16011; Q16014;
AC   Q16019; Q16020; Q9BT38; Q9UCA9; Q9UCB6; Q9UCC8; Q9UCD1; Q9UQ58;
DT   13-AUG-1987, integrated into UniProtKB/Swiss-Prot.
DT   01-NOV-1991, sequence version 3.
DT   10-JUN-2008, entry version 153.
DE   Amyloid beta A4 protein precursor (Alzheimer disease amyloid protein)
DE   (ABPP) (APPI) (APP) (PreA4) (Cerebral vascular amyloid peptide) (CVAP)
DE   (Protease nexin-II) (PN-II) [Contains: Soluble APP-alpha (S-APP-
DE   alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42
DE   (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42);
DE   P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59)
DE   (Amyloid intracellular domain 59) (AICD-59) (AID(59)); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57) (Amyloid intracellular domain
DE   57) (AICD-57) (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal
DE   fragment 50) (Amyloid intracellular domain 50) (AICD-50) (AID(50));
DE   C31].
GN   Name=APP; Synonyms=A4, AD1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
OC   Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=87144572; PubMed=2881207; DOI=10.1038/325733a0;
RA   Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,
```

快速查询存在的假阳性结果

并不是我们要查找的蛋白

Entry name	**A4_HUMAN**
Accession number	**P05067**, P09000, P78438, Q13764, Q13778, Q13793, Q16011, Q16014, Q16019, Q16020, Q9BT38, Q9UCA9, Q9UCB6, Q9UCC8, Q9UCD1, Q9UQ58
Integrated	13-AUG-1987, UniProtKB/Swiss-Prot.
Sequence update	01-NOV-1991, sequence version 3
Annotation update	10-JUN-2008, entry version 153
UniSave	P05067, P09000, P78438, Q13764, Q13778, Q13793, Q16011, Q16014, Q16019, Q16020, Q9BT38, Q9UCA9, Q9UCB6, Q9UCC8, Q9UCD1, Q9UQ58
UniRef100	UniRef100_P05067, UniRef100_P05067-6, UniRef100_P05067-3, UniRef100_P05067-7, UniRef100_P05067-9, UniRef100_P05067-2, UniRef100_P05067-4, UniRef100_P05067-10, UniRef100_P05067-5, UniRef100_Q6GSC0
UniParc	UPI000002A2F0, UPI000002A2F1, UPI000002A2F2, UPI000002A2F3, UPI000002A2F4, UPI000002A2F5, UPI000002A2F6, UPI000002A2F7, UPI000002DB1C, UPI00001BE539
Description and origin of the Protein	
Description	Amyloid beta A4 protein precursor (Alzheimer disease amyloid protein) (ABPP) (APPI) (APP) (PreA4) (Cerebral vascular amyloid peptide) (CVAP) (Protease nexin-II) (PN-II) [Contains: Soluble APP-alpha (S-APP- alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59) (Amyloid intracellular domain 59) (AICD-59) (AID(59)); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57) (Amyloid intracellular domain 57) (AICD-57) (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain 50) (AICD-50) (AID(50)); C31].

CiteXplore

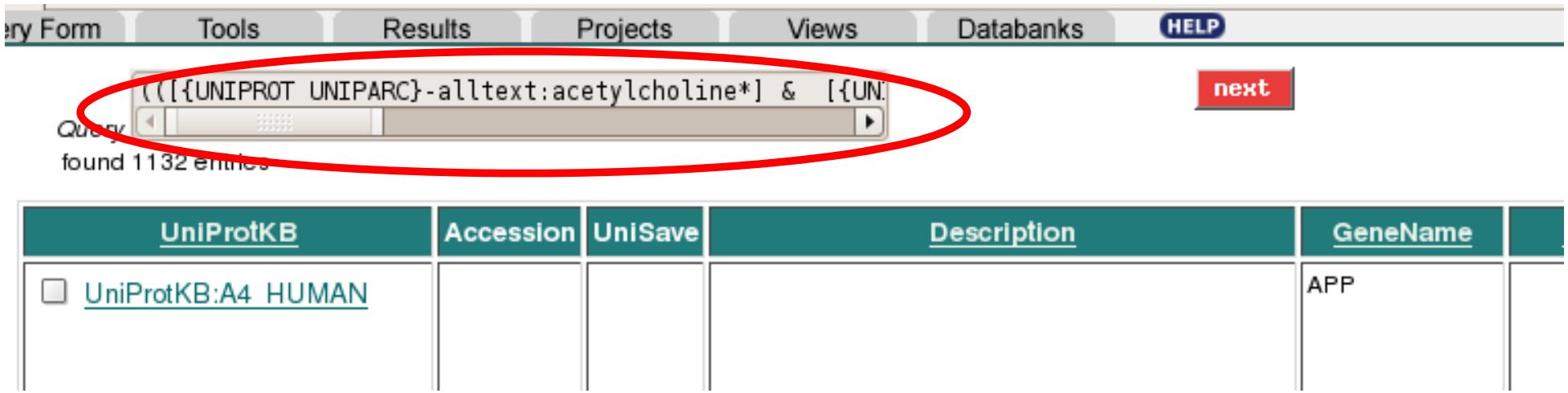
[52] Wang,H.-Y., Lee,D.H.S., D'Andrea,M.R., Peterson,P.A., Shank,R.P., Reitz,A.B.,
Beta-amyloid(1-42) binds to alpha7 nicotinic acetylcholine receptor with high affinity. Implications for Alzheimer's disease pathology.
(2000) *J. Biol. Chem.* **275**:5626-5632

Position	INTERACTION OF BETA-APP42 WITH CHRNA7.
Medline	20148766
DOI	10.1074/jbc.275.8.5626;
PubMed	10681545 HTML PDF ABS DOI CiteXplore

引用文献中出现了我们查询的关键词

[53] Passer,B., Pellegrini,L., Russo,C., Siegel,R.M., Lenardo,M.J., Schettini,G., Bachmann,M., Tabaton,M., D'Adamio,L.,
Generation of an apoptotic intracellular peptide by gamma-secretase cleavage of Alzheimer's amyloid beta protein precursor.

为什么会出现假阳性结果

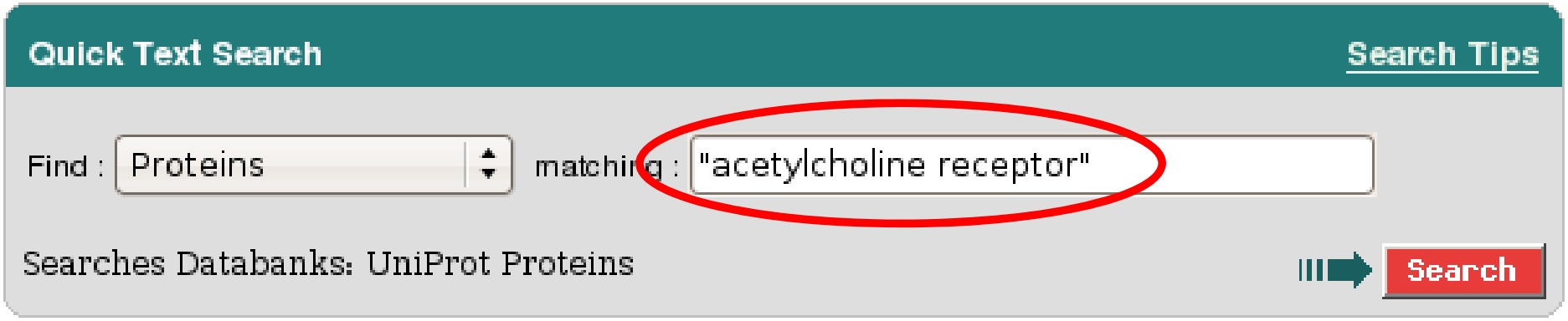


(([{UNIPROT UNIPARC}-alltext:acetylcholine*]
& [{UNIPROT UNIPARC}-alltext:receptor*])
| [{UNIPROT UNIPARC}-alltext:acetylcholine receptor*])

（1）alltext 则文字可能被包含于Reference中

（2）前两个条件，可能某条结果包含乙酰胆碱，
而受体却不是针对乙酰胆碱的受体

如何减少假阳性结果



添加引号后，查询语句变为：

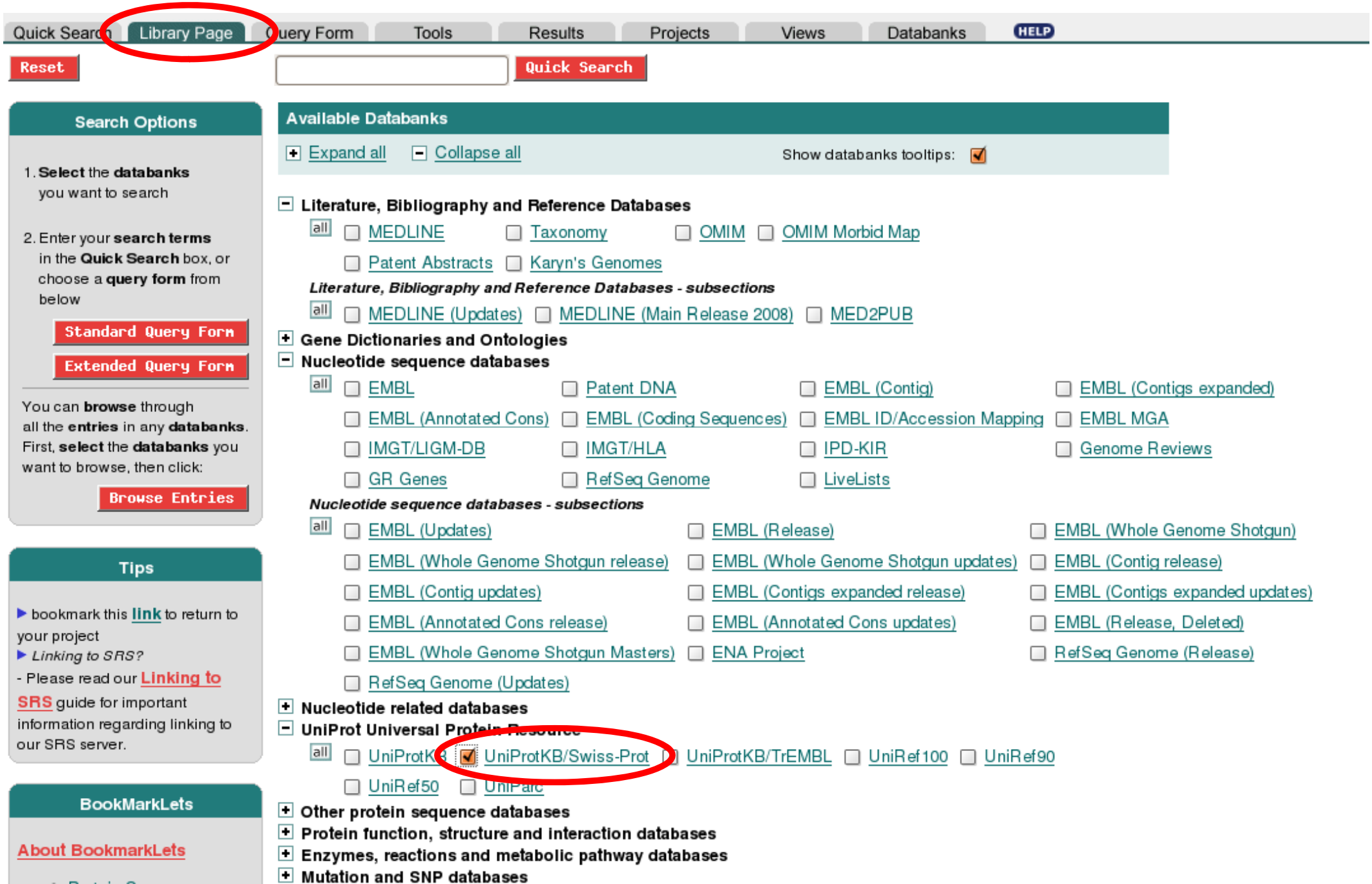
[{UNIPROT UNIPARC}-alltext:acetylcholine receptor]

从而有效避免了一部分假阳性结果的出现

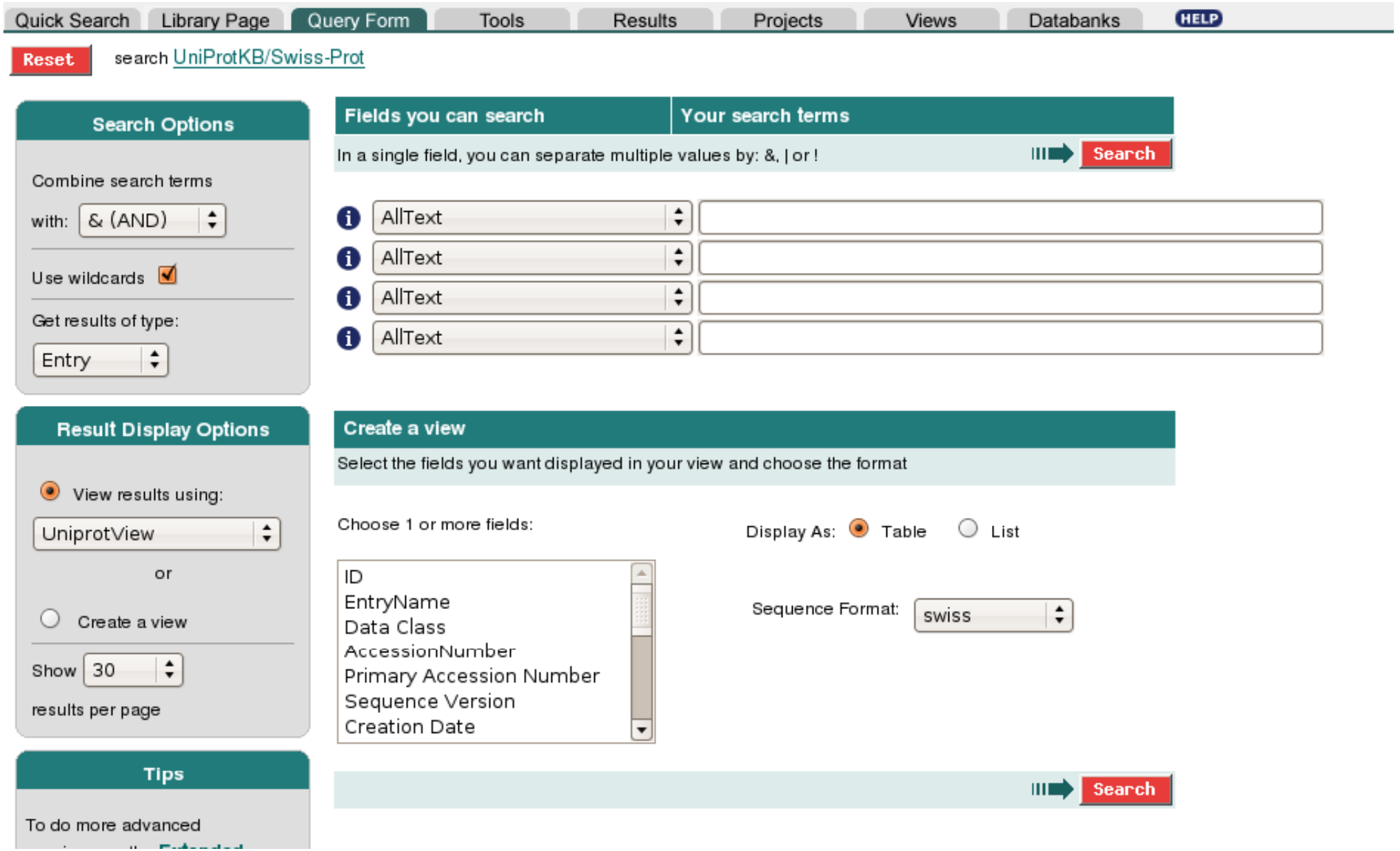
SRS查询语言

- SRS Query Language，为SRS系统实现的基础接口（所有查询操作都会转换成为SRS查询语言进行执行）
- 准确表示出查询语义
- 可任意灵活地组合各种条件，包括一些用网页操作界面很难完成的查询操作
- 但学习和掌握起来难度较大
- 这里不再深入介绍

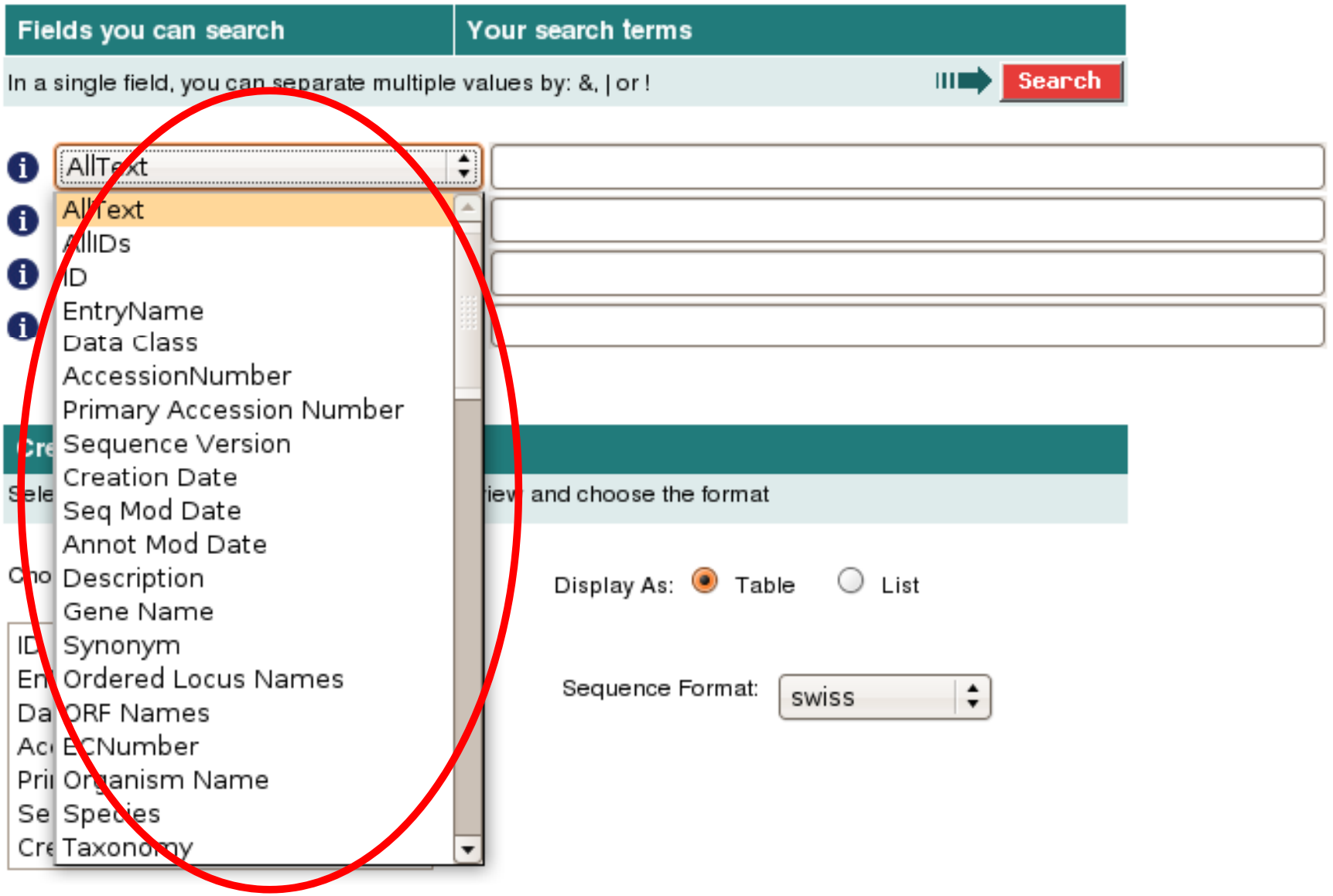
标准查询前需要先选择数据库



SRS标准查询界面

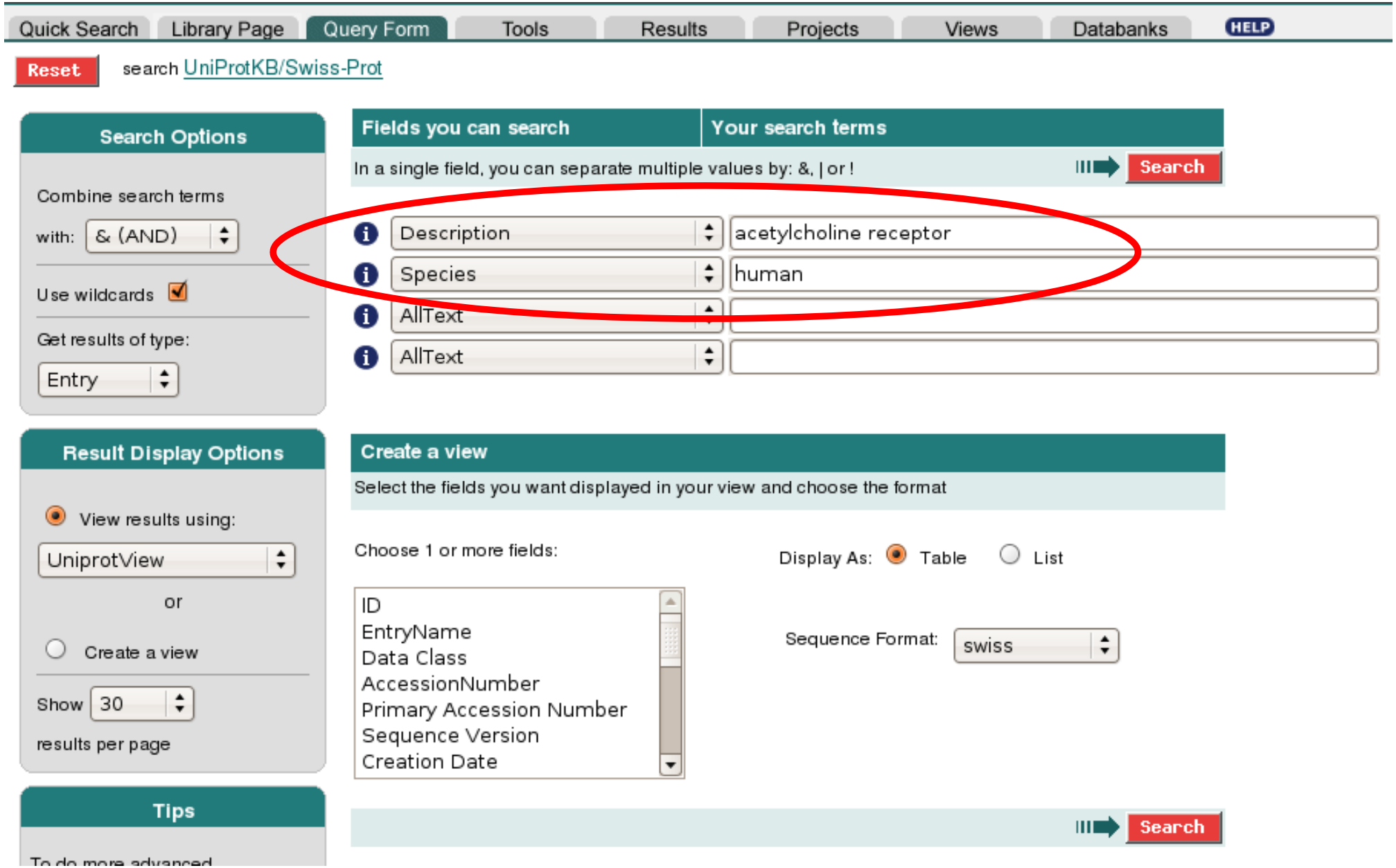

Quick Search
Library Page
Query Form
Tools
Results
Projects
Views
Databanks
HELP
Reset
search UniProtKB/Swiss-Prot
Search Options
Combine search terms
with: & (AND)
Use wildcards
Get results of type:
Entry
Fields you can search
Your search terms
In a single field, you can separate multiple values by: &, | or !
Search
AllText
AllText
AllText
AllText
Result Display Options
View results using:
UniprotView
or
Create a view
Show 30
results per page
Create a view
Select the fields you want displayed in your view and choose the format
Choose 1 or more fields:
ID
EntryName
Data Class
AccessionNumber
Primary Accession Number
Sequence Version
Creation Date
Display As: Table List
Sequence Format: swiss
Search
Tips
To do more advanced

SRS标准查询的字段选项



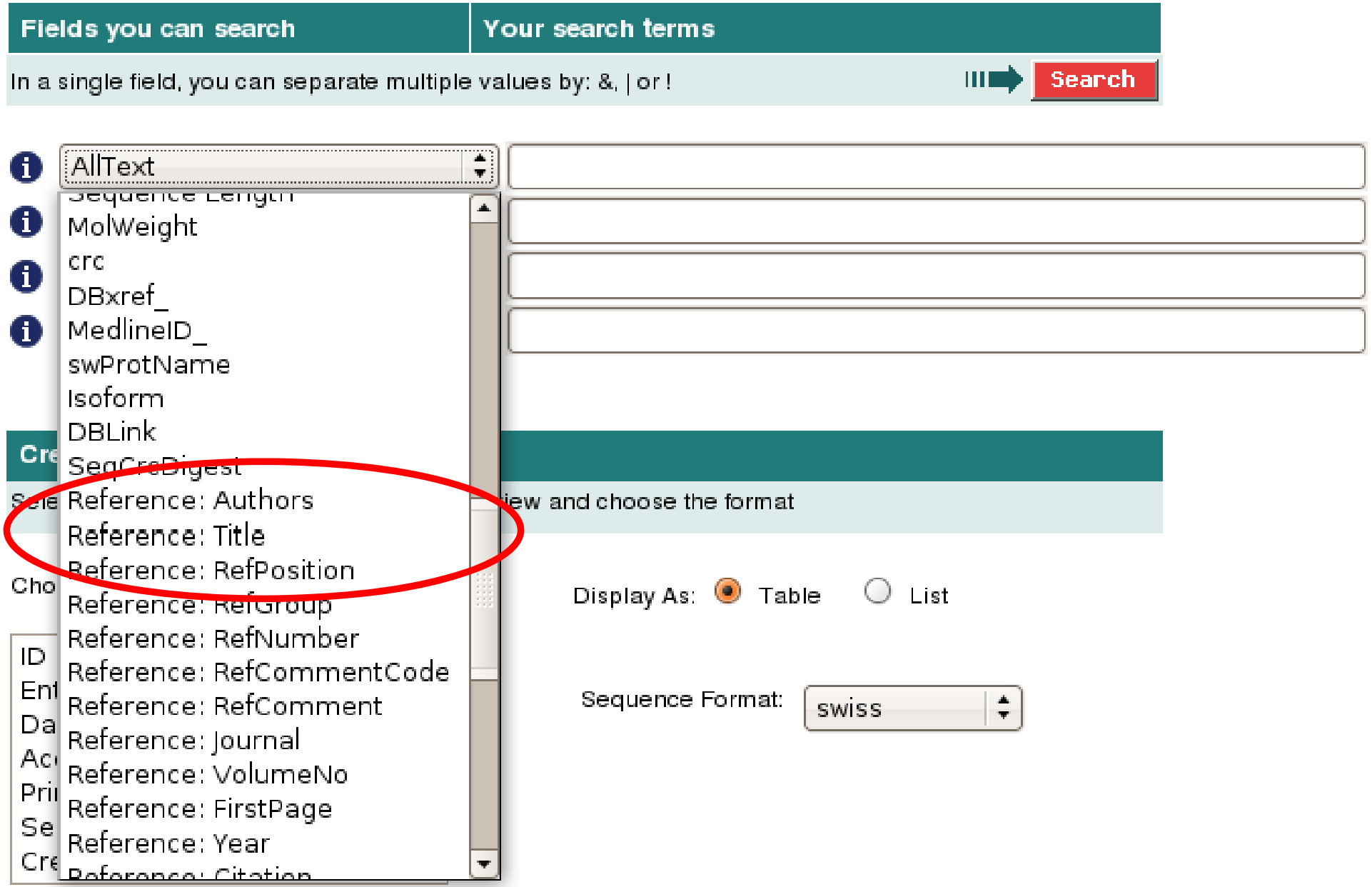
小技巧：当焦点位于该下拉框时，可以直接按字母键快速定位到某项

SRS标准查询实例



查询人的乙酰胆碱受体蛋白

子条目(Subentry)字段



General information

Entry name	**ACHB2_RAT**
Accession number	**P12390**, Q53YK1
Integrated	01-OCT-1989, UniProtKB/Swiss-Prot.
Sequence update	30-MAY-2000, sequence version 2
Annotation update	10-JUN-2008, entry version 77
UniSave	P12390, Q53YK1
UniRef100	UniRef100_P12390
UniParc	UPI000012526F

Description and origin of the Protein

Description	Neuronal acetylcholine receptor subunit beta-2 precursor (Neuronal acetylcholine receptor non-alpha-1 chain) (N-alpha 1).
Gene name(s)	Chrnb2
Synonym(s)	Acrb2
Organism source	Rattus norvegicus (Rat).
Taxonomy	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
NCBI TaxID	10116

References

[1] Deneris,E.S., Connolly,J.G., Boulter,J., Wada,E., Wada,K., Swanson,L.W., Patrick,J., Heinemann,S.F.,
Primary structure and expression of beta 2: a novel subunit of neuronal nicotinic acetylcholine receptors.
(1988) *Neuron* **1**:45-54

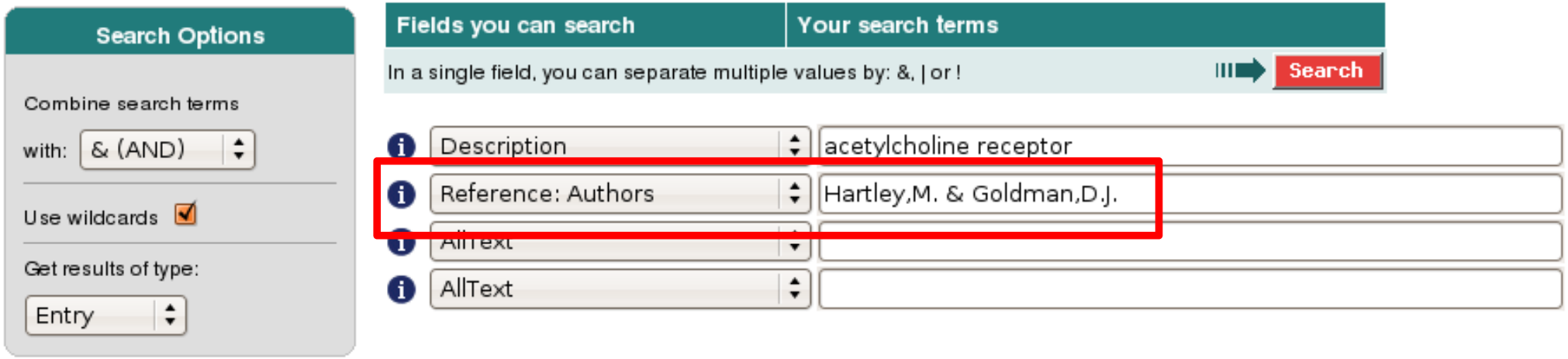
Position	NUCLEOTIDE SEQUENCE [MRNA].
Medline	90166479
DOI	10.1016/0896-6273(88)90208-5;
PubMed	3272154 CiteXplore

[2] Boulter,J., Connolly,J.G., Deneris,E.S., Goldman,D.J., Heinemann,S.F., Patrick,J.,
Functional expression of two neuronal nicotinic acetylcholine receptors from cDNA clones identifies a gene family.
(1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**:7763-7767

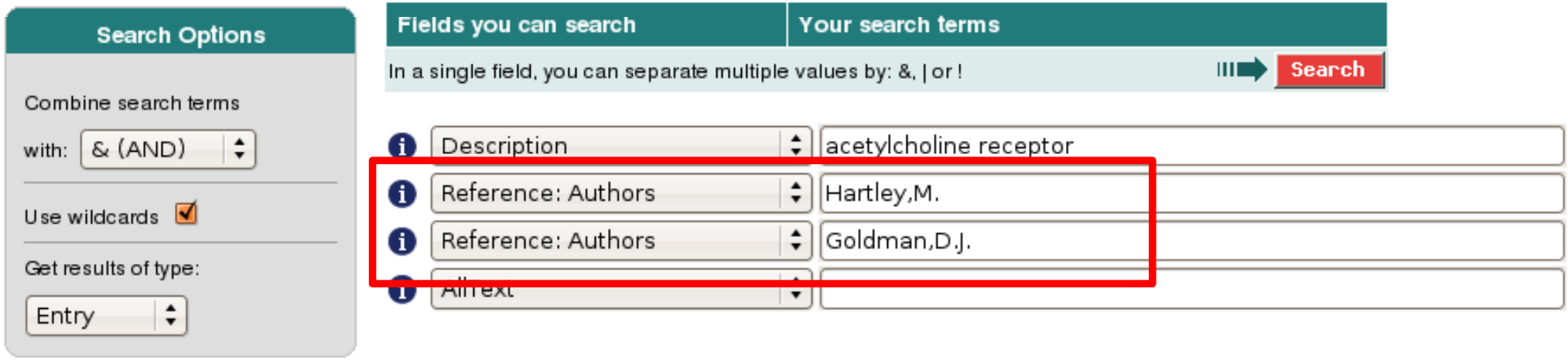
Position	NUCLEOTIDE SEQUENCE [MRNA].
Medline	88041184
PubMed	2444984 CiteXplore

关于子条目字段的查询

查询条件一：单个Reference:Authors字段



查询条件二：两个Reference:Authors字段



第一个查询条件的结果：

Query ((([swissprot-Description:acetylcholine*] & [swissp

found 1 entries

UniProtKB/Swiss-Prot	Accession	UniSave	Description	GeneName	Species	Keywords	SeqLength
UniProtKB/Swiss-Prot:ACHA4_RAT	P09483	P09483	Neuronal acetylcholine receptor subunit alpha-4 precursor.	Chrna4	Rattus norvegicus (Rat).	3D-structure Alternative splicing Cell junction Direct protein sequencing Glycoprotein Ion transport Ionic channel Membrane Postsynaptic cell membrane Signal Synapse Transmembrane Transport	630

第二个查询条件的结果：

Query (((([swissprot-Description:acetylcholine*] & [swiss

found 2 entries

UniProtKB/Swiss-Prot	Accession	UniSave	Description	GeneName	Species	Keywords	SeqLength
UniProtKB/Swiss-Prot:ACHA4_RAT	P09483	P09483	Neuronal acetylcholine receptor subunit alpha-4 precursor.	Chrna4	Rattus norvegicus (Rat).	3D-structure Alternative splicing Cell junction Direct protein sequencing Glycoprotein Ion transport Ionic channel Membrane Postsynaptic cell membrane Signal Synapse Transmembrane Transport	630
UniProtKB/Swiss-Prot:ACHB2_RAT	P12390	P12390	Neuronal acetylcholine receptor subunit beta-2 precursor (Neuronal acetylcholine receptor non-alpha-1 chain) (N-alpha 1).	Chrnb2	Rattus norvegicus (Rat).	3D-structure Cell junction Glycoprotein Ion transport Ionic channel Membrane Postsynaptic cell membrane Signal Synapse Transmembrane Transport	500

结果多出一条！

Annotation update	10-JUN-2008, entry version 77
UniSave	P12390, Q53YK1
UniRef100	UniRef100_P12390
UniParc	UPI000012526F

Description and origin of the Protein

Description	Neuronal acetylcholine receptor subunit beta-2 precursor (Neuronal acetylcholine receptor non-alpha-1 chain) (N-alpha 1).
Gene name(s)	Chrnb2
Synonym(s)	Acrb2
Organism source	Rattus norvegicus (Rat).
Taxonomy	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
NCBI TaxID	10116

References

[1] Deneris,E.S., Connolly,J.G., Boulter,J., Wada,E., Wada,K., Swanson,L.W., Patrick,J., Heinemann,S.F.,
Primary structure and expression of beta 2: a novel subunit of neuronal nicotinic acetylcholine receptors.
(1988) *Neuron* **1**:45-54

Position	NUCLEOTIDE SEQUENCE [MRNA].
Medline	90166479
DOI	10.1016/0896-6273(88)90208-5;
PubMed	3272154 CiteXplore

[2] Boulter,J., Connolly,J.G., Deneris,E.S., Goldman,D.J., Heinemann,S.F., Patrick,J.,
Functional expression of two neuronal nicotinic acetylcholine receptors from cDNA clones identifies a gene family.
(1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**:7763-7767

Position	NUCLEOTIDE SEQUENCE [MRNA].
Medline	88041184
PubMed	2444984 CiteXplore

[3] Hartley,M.,
Submitted NOV-1998 to the EMBL GenBank DDBJ databases

Position	SEQUENCE REVISION.

[4] Groot-Kormelink,P.J.,
Submitted MAR-2004 to the EMBL GenBank DDBJ databases

Position	NUCLEOTIDE SEQUENCE [MRNA].
Comments	STRAIN=Sprague-Dawley; TISSUE=Brain;

Comments

子条目字段查询条件解释

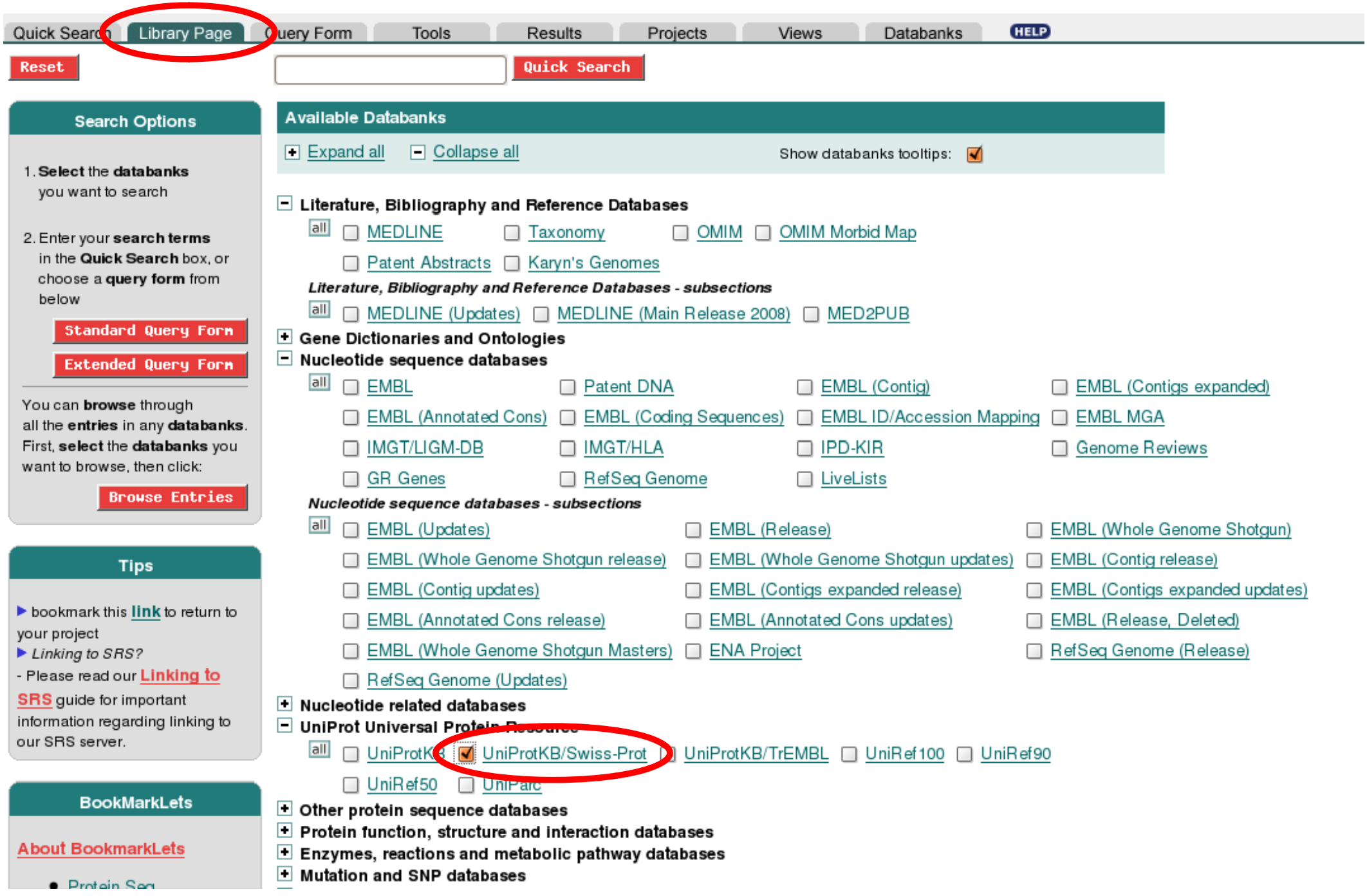
查询条件一：单个Reference:Authors字段

指被引用的同一篇文献的作者，需要同时包含两个人，
两个人为同一篇文献的共同作者

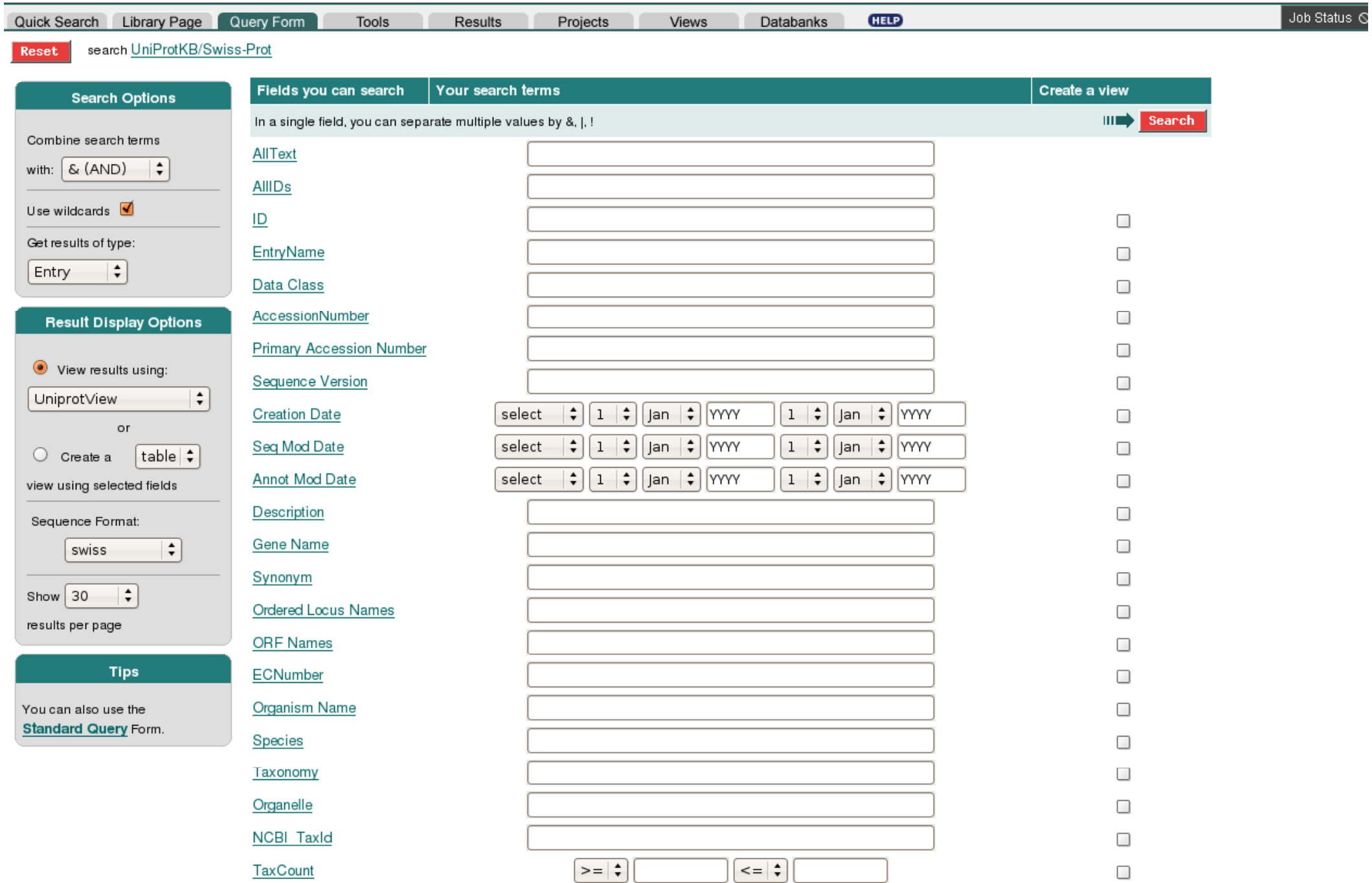
查询条件二：两个Reference:Authors字段

指被引用的多篇文献的作者，需要同时包含两个人，
而两个人不需要为同一篇文献的共同作者

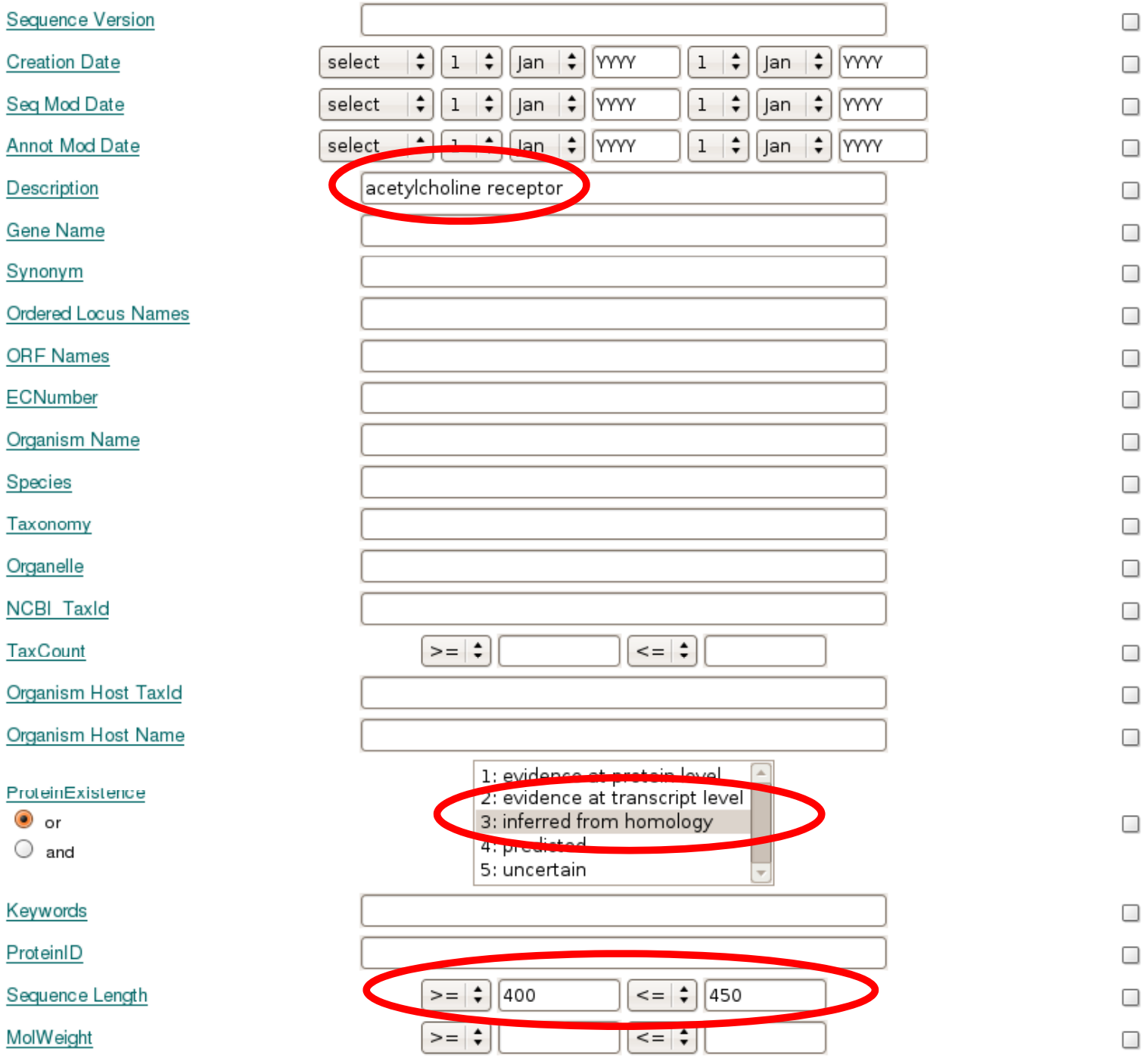
扩展查询前也需要先选择数据库



SRS扩展查询界面

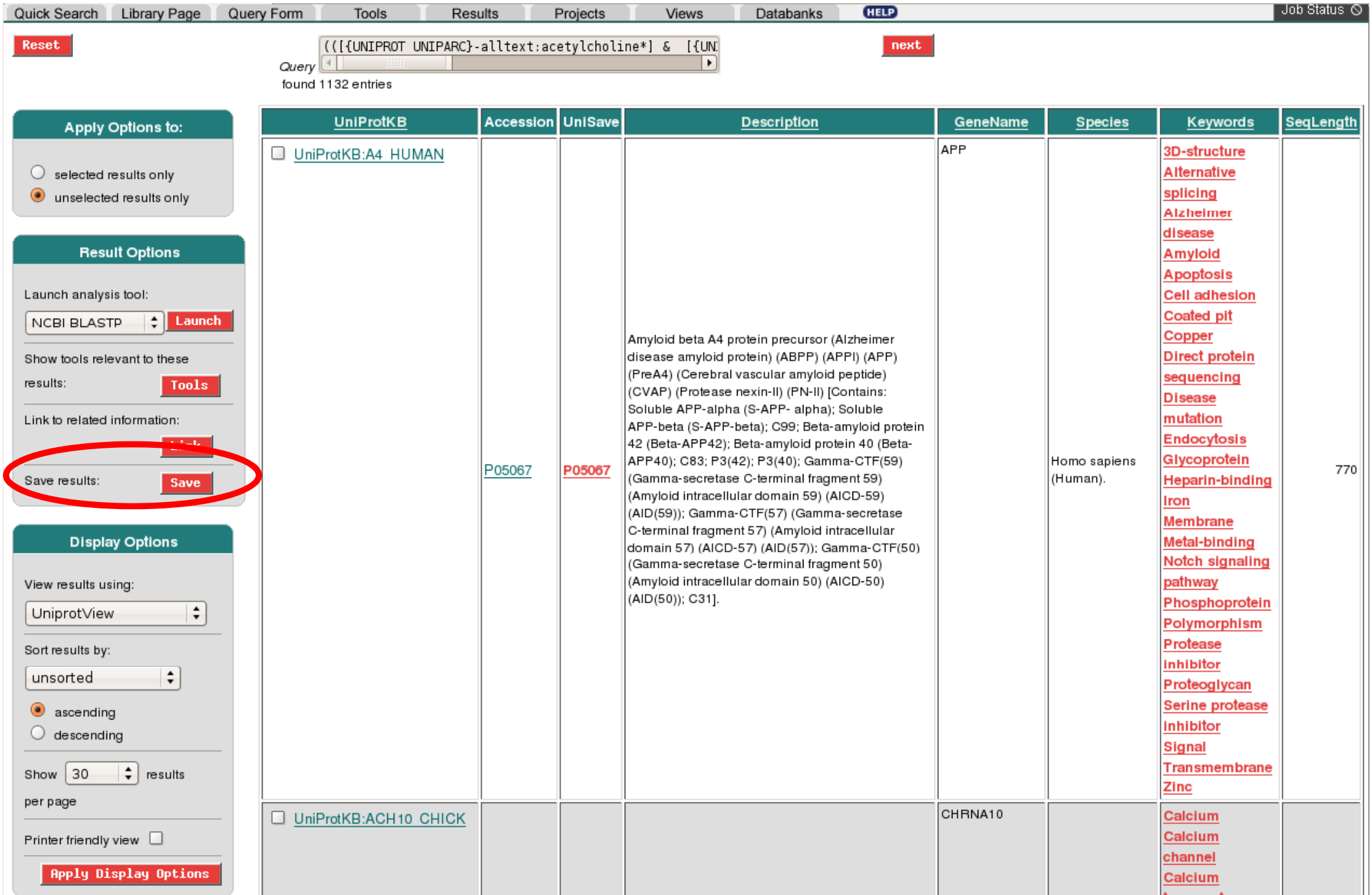

Quick Search
Library Page
Query Form
Tools
Results
Projects
Views
Databanks
HELP
Job Status
Reset
search UniProtKB/Swiss-Prot
Search Options
Combine search terms
with: & (AND)
Use wildcards
Get results of type:
Entry
Result Display Options
View results using:
UniprotView
or
Create a table
view using selected fields
Sequence Format:
swiss
Show 30
results per page
Tips
You can also use the
Standard Query Form.
Fields you can search
Your search terms
Create a view
In a single field, you can separate multiple values by &, |, !
Search
AllText
AllIDs
ID
EntryName
Data Class
AccessionNumber
Primary Accession Number
Sequence Version
Creation Date
Seq Mod Date
Annot Mod Date
Description
Gene Name
Synonym
Ordered Locus Names
ORF Names
ECNumber
Organism Name
Species
Taxonomy
Organelle
NCBI_TaxId
TaxCount
select
1
Jan
YYYY
>=
<=

扩展查询能够更详细地选择条件

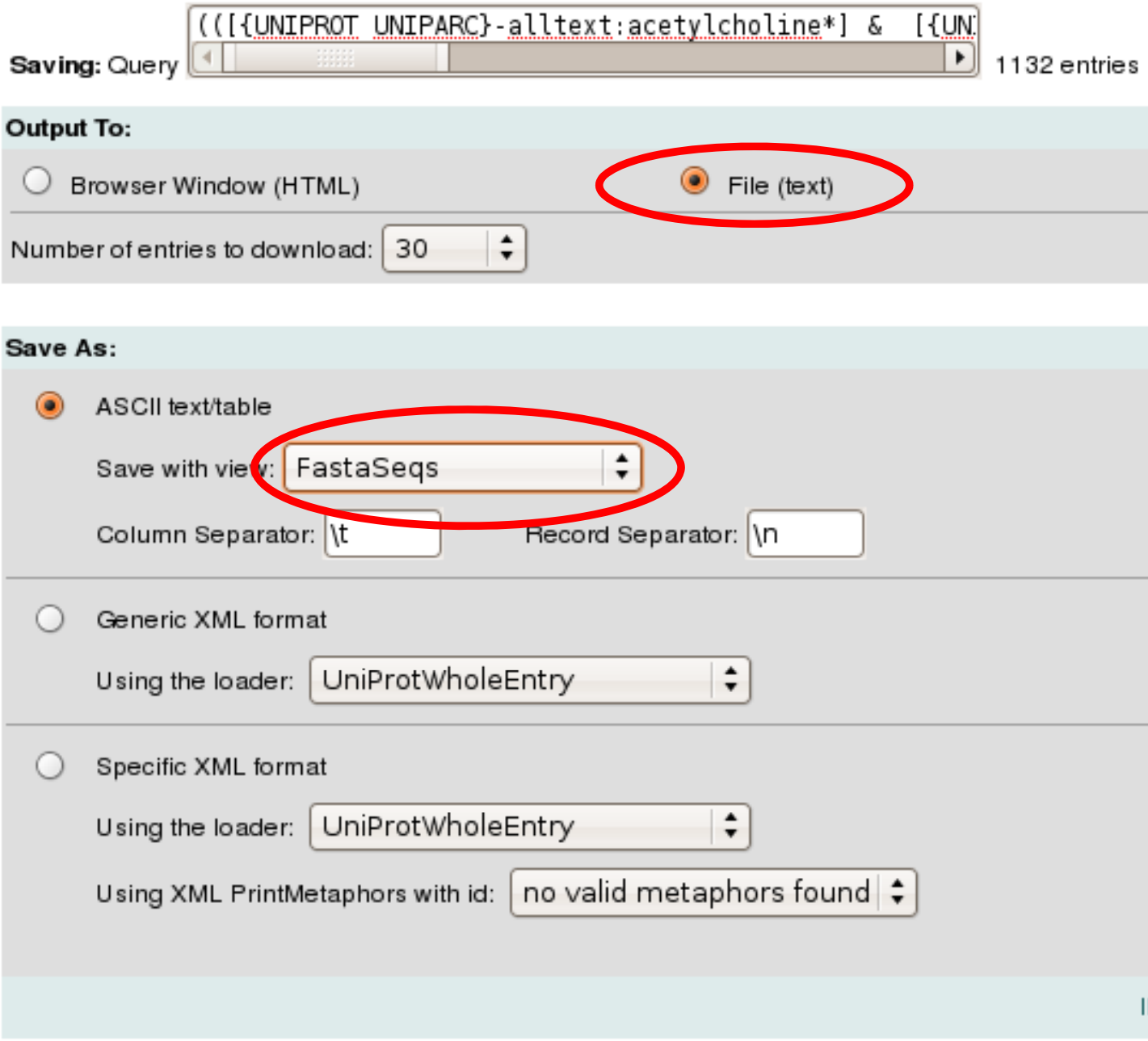


示例：
查询通过同源推断得知的序列长度在400到450之间的乙酰胆碱受体蛋白

保存查询结果


Quick Search
Library Page
Query Form
Tools
Results
Projects
Views
Databanks
HELP
Job Status
Reset
Query
([{UNIPROT_UNIPARC}-alltext:acetylcholine*] & [{UN
found 1132 entries
next
Apply Options to:
selected results only
unselected results only
Result Options
Launch analysis tool:
NCBI BLASTP
Launch
Show tools relevant to these results:
Tools
Link to related information:
Save results:
Save
Display Options
View results using:
UniprotView
Sort results by:
unsorted
ascending
descending
Show 30 results per page
Printer friendly view
Apply Display Options
UniProtKB
Accession
UniSave
Description
GeneName
Species
Keywords
SeqLength
UniProtKB:A4_HUMAN
P05067
P05067
Amyloid beta A4 protein precursor (Alzheimer disease amyloid protein) (ABPP) (APPI) (APP) (PreA4) (Cerebral vascular amyloid peptide) (CVAP) (Protease nexin-II) (PN-II) [Contains: Soluble APP-alpha (S-APP- alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59) (Amyloid intracellular domain 59) (AICD-59) (AID(59)); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57) (Amyloid intracellular domain 57) (AICD-57) (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain 50) (AICD-50) (AID(50)); C31].
APP
Homo sapiens (Human).
3D-structure
Alternative splicing
Alzheimer disease
Amyloid
Apoptosis
Cell adhesion
Coated pit
Copper
Direct protein sequencing
Disease mutation
Endocytosis
Glycoprotein
Heparin-binding
Iron
Membrane
Metal-binding
Notch signaling pathway
Phosphoprotein
Polymorphism
Protease inhibitor
Proteoglycan
Serine protease inhibitor
Signal
Transmembrane
Zinc
770
UniProtKB:ACH10_CHICK
CHRNA10
Calcium
Calcium channel
Calcium

保存查询结果的界面


Saving: Query
(([{UNIPROT UNIPARC}-alltext:acetylcholine*] & [{UN
1132 entries
Output To:
Browser Window (HTML)
File (text)
Number of entries to download: 30
Save As:
ASCII text/table
Save with view: FastaSeqs
Column Separator: \t
Record Separator: \n
Generic XML format
Using the loader: UniProtWholeEntry
Specific XML format
Using the loader: UniProtWholeEntry
Using XML PrintMetaphors with id: no valid metaphors found
Save

3、高级应用技巧

1.链接(Link)

2.结果视图(View)

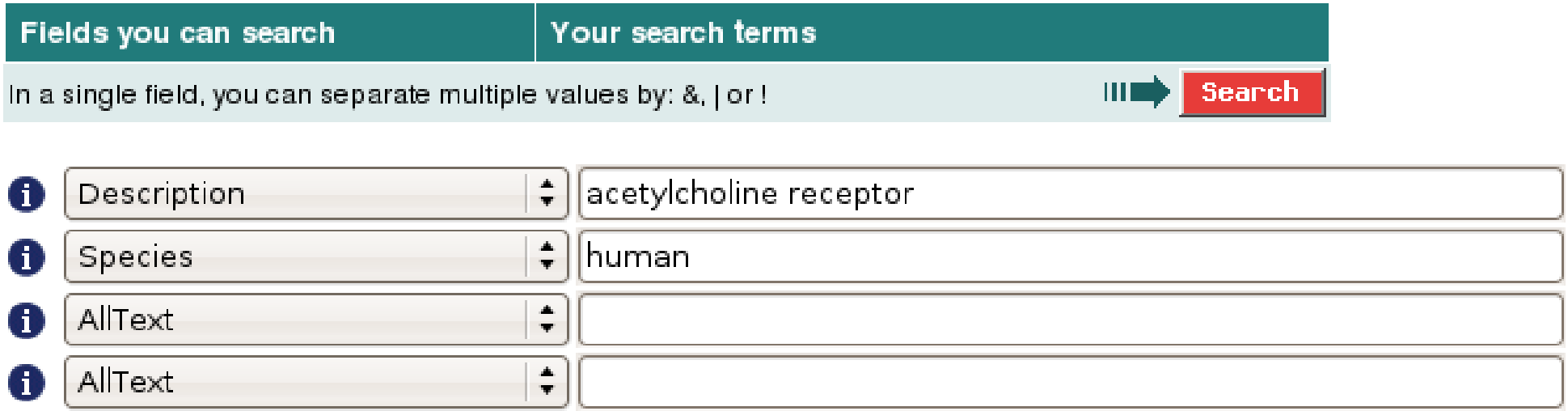
3.数据分析

4.SRS项目(Project)

SRS链接(Link)

- 数据库间的相互引用关系
- 对生物信息的理解与标注(annotation)
- 链接把数据库有意义地关联起来
  - 例如：embl 与 swissprot之间的链接，就能说明某核酸片段序列有相关的蛋白序列
- SRS对已存在的关联关系进行查询

SRS链接应用实例



仍然以“人的乙酰胆碱受体”为例，先进行标准查询

查询得到结果

Reset

Query ((([swissprot-Description:acetylcholine*] & [swissp

found 23 entries

Apply Options to:

- selected results only
- unselected results only

Result Options

Launch analysis tool: NCBI BLASTP Launch

Show tools relevant to these results: Tools

Link to related information: Link

Save results: Save

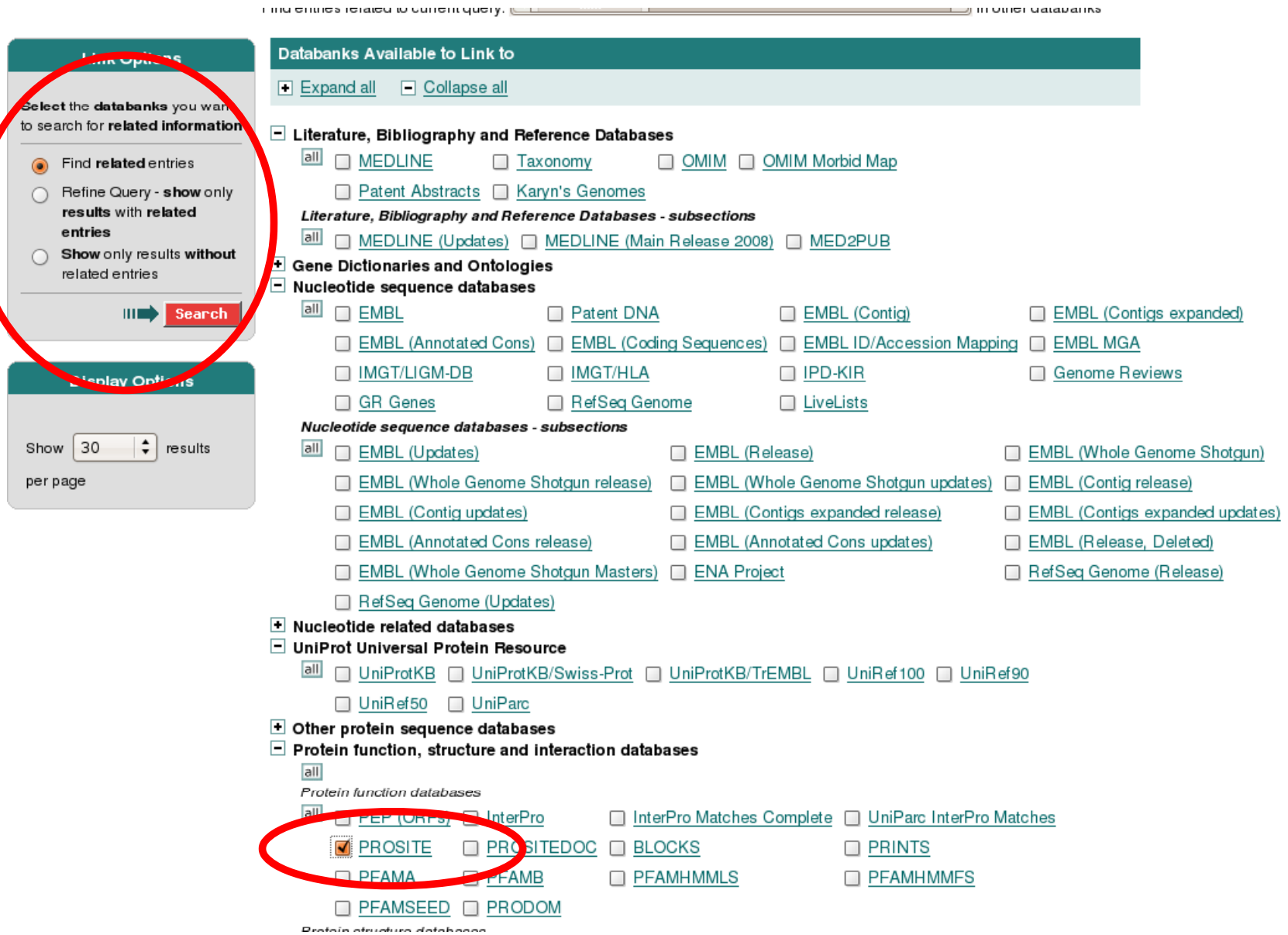
Display Options

View results using: UniprotView

Sort results by: unsorted

UniProtKB/Swiss-Prot	Accession	UniSave	Description	GeneName	Species	Keywords	SeqLength
UniProtKB/Swiss-Prot:ACH10_HUMAN	Q9GZZ6	Q9GZZ6	Neuronal acetylcholine receptor subunit alpha-10 precursor (Nicotinic acetylcholine receptor subunit alpha 10) (NACHR alpha 10).	CHRNA10	Homo sapiens (Human).	Calcium; Calcium channel; Calcium transport; Cell junction; Glycoprotein; Ion transport; Ionic channel; Membrane; Postsynaptic cell membrane; Receptor; Signal; Synapse; Transmembrane; Transport	450
UniProtKB/Swiss-Prot:ACHA2_HUMAN	Q15822	Q15822	Neuronal acetylcholine receptor subunit alpha-2 precursor.	CHRNA2	Homo sapiens (Human).	Cell junction; Disease mutation; Epilepsy; Glycoprotein; Ion transport; Ionic channel; Membrane; Polymorphism; Postsynaptic cell membrane; Signal; Synapse	529

选择需要链接的库



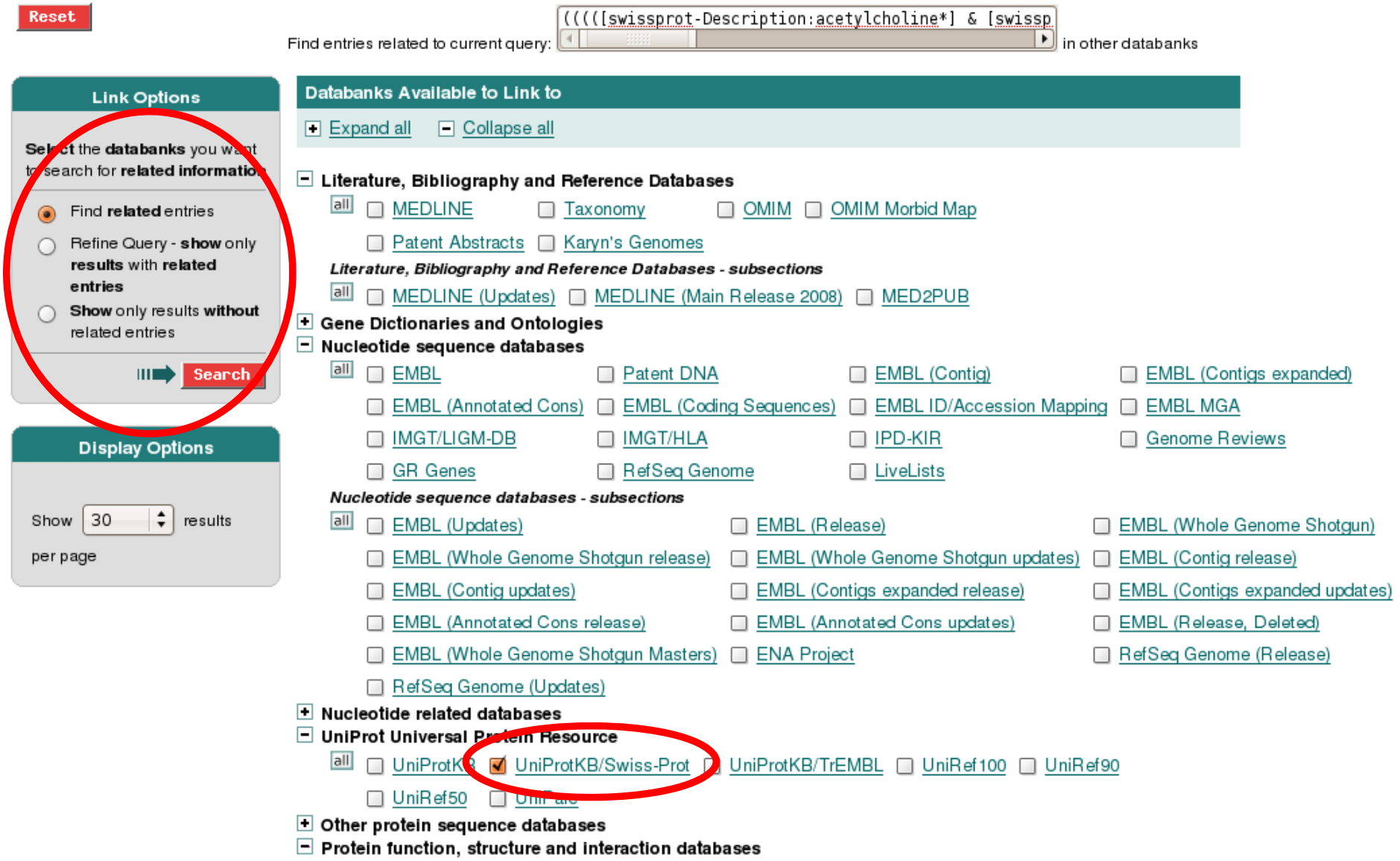
选择PROSITE库，表示结果相关蛋白家族

链接结果

所得结果为：
人乙酰胆碱受体蛋白相关家族

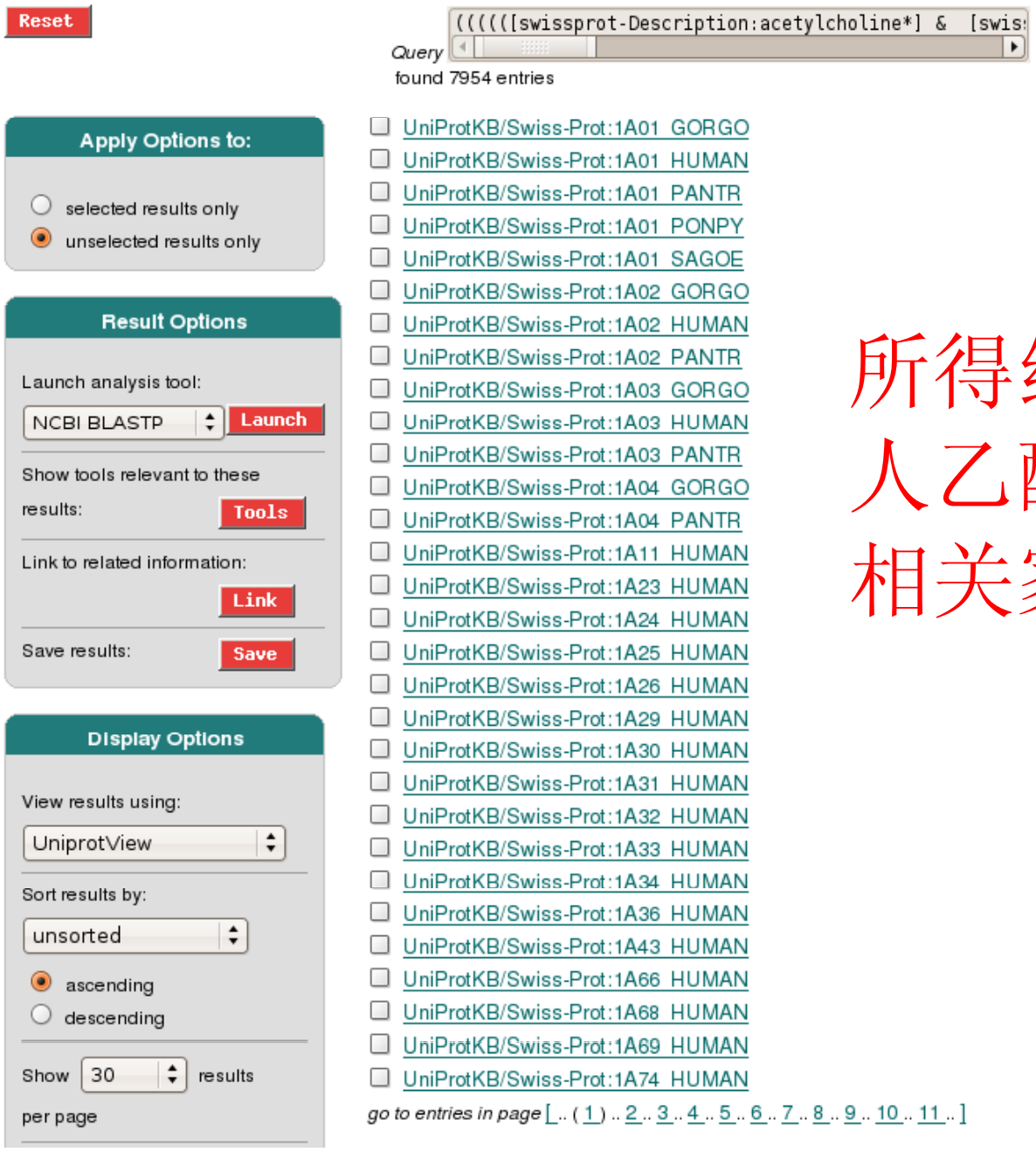
此时还可以继续进行链接

再次选择需要链接的库



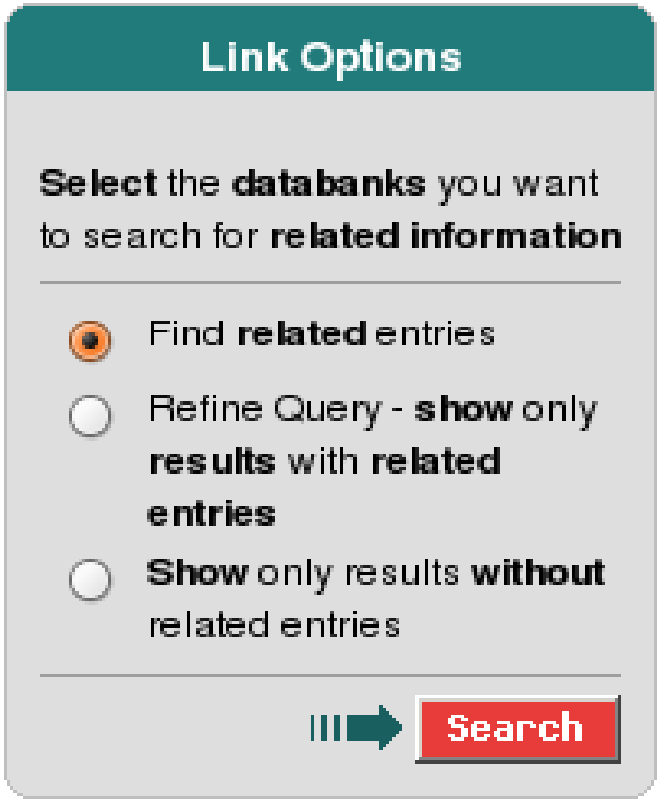
选择SwissProt库，表示结果相关的所有蛋白

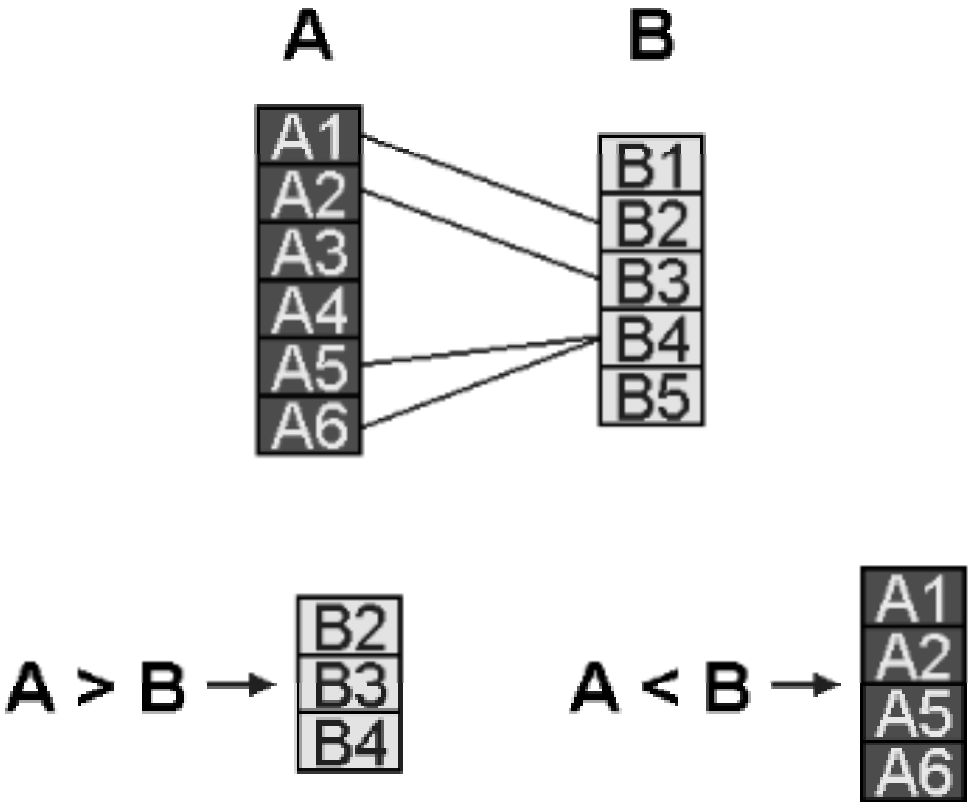
最终链接结果



所得结果为：
人乙酰胆碱受体蛋白
相关家族中的所有蛋白

链接类型





设A为前一次查询的结果，B为所链接的库。

于是三种方式的结果分别为：

选项1：B2, B3, B4（相当于A>B）

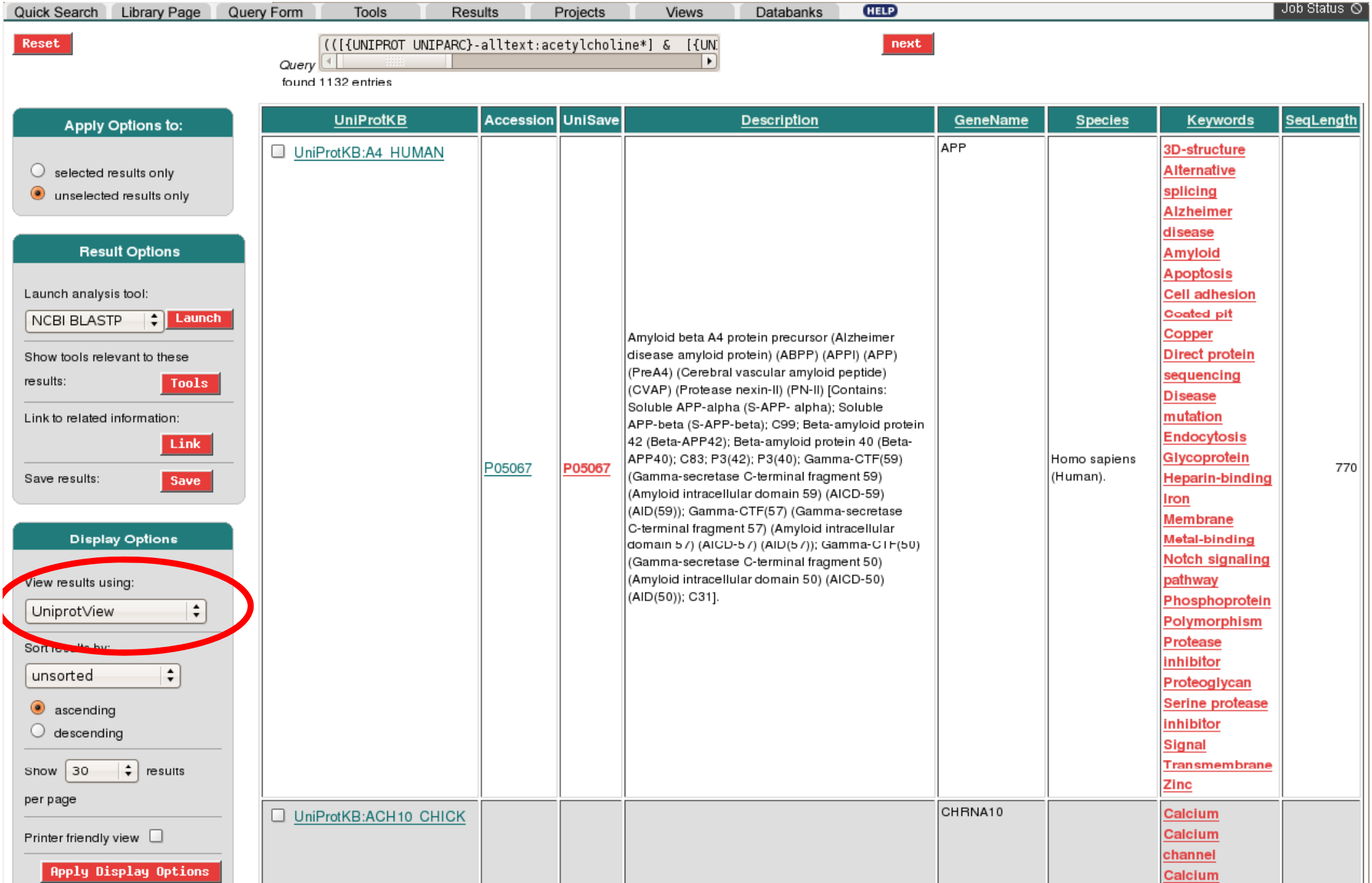
选项2：A1, A2, A5, A6（相当于A<B）

选项3：A3, A4（相当于A！(A<B)）

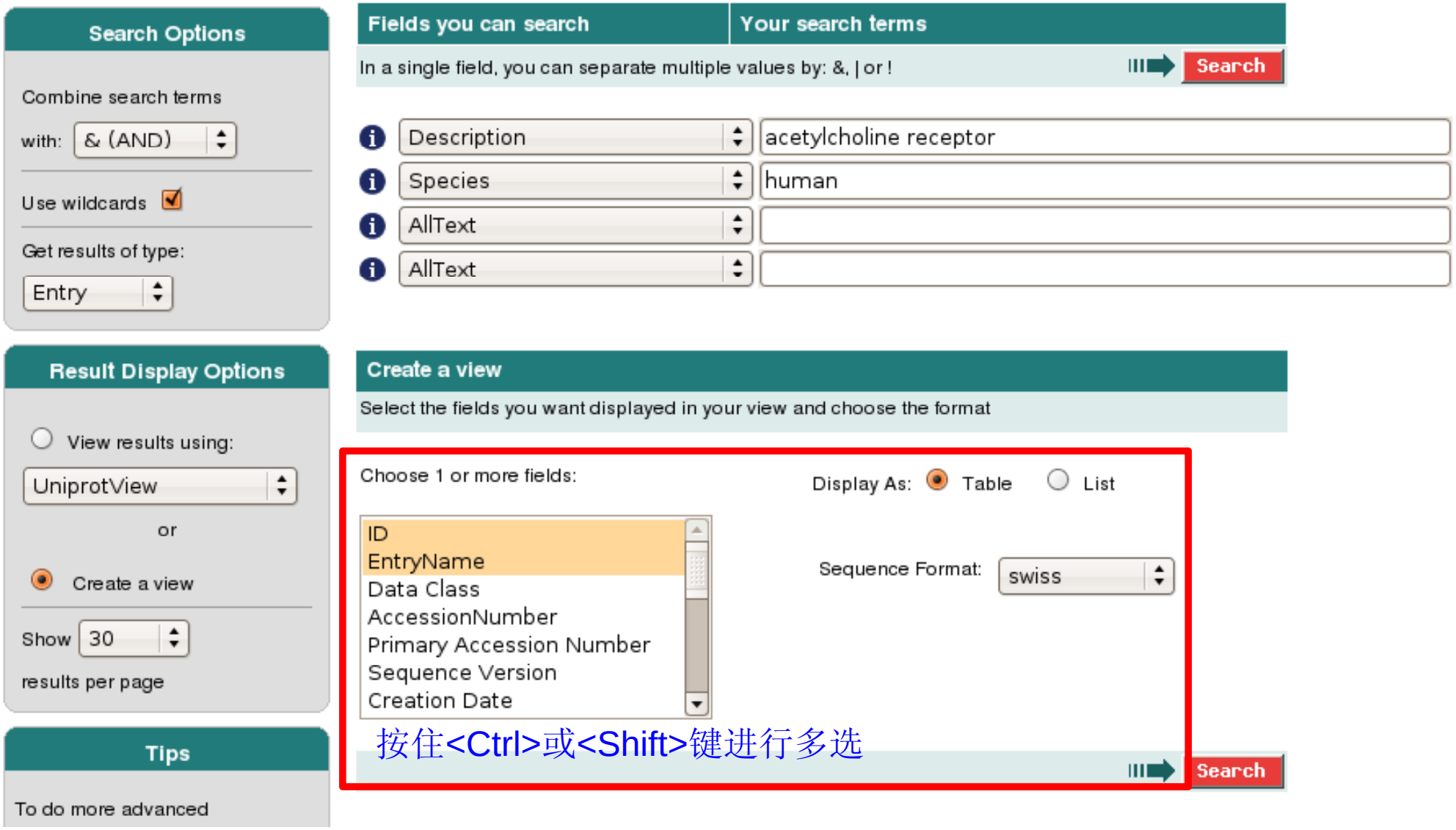
结果视图(View)

- 结果视图是结果页面的显示方式
- 可使用SRS系统默认视图
- 也可自定义视图（显示哪些特定的字段列）

通常的查询结果视图


Quick Search
Library Page
Query Form
Tools
Results
Projects
Views
Databanks
HELP
Job Status
Reset
([{UNIPROT UNIPARC}-alltext:acetylcholine*] & [{UN
Query
found 1132 entries
next
Apply Options to:
selected results only
unselected results only
Result Options
Launch analysis tool:
NCBI BLASTP
Launch
Show tools relevant to these results:
Tools
Link to related information:
Link
Save results:
Save
Display Options
View results using:
UniprotView
unsorted
ascending
descending
Show 30 results per page
Printer friendly view
Apply Display Options
UniProtKB
Accession
UniSave
Description
GeneName
Species
Keywords
SeqLength
UniProtKB:A4_HUMAN
P05067
P05067
Amyloid beta A4 protein precursor (Alzheimer disease amyloid protein) (ABPP) (APPI) (APP) (PreA4) (Cerebral vascular amyloid peptide) (CVAP) (Protease nexin-II) (PN-II) [Contains: Soluble APP-alpha (S-APP- alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59) (Amyloid intracellular domain 59) (AICD-59) (AID(59)); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57) (Amyloid intracellular domain 57) (AICD-57) (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain 50) (AICD-50) (AID(50)); C31].
APP
Homo sapiens (Human).
3D-structure
Alternative splicing
Alzheimer disease
Amyloid
Apoptosis
Cell adhesion
Coated pit
Copper
Direct protein sequencing
Disease mutation
Endocytosis
Glycoprotein
Heparin-binding
Iron
Membrane
Metal-binding
Notch signaling pathway
Phosphoprotein
Polymorphism
Protease inhibitor
Proteoglycan
Serine protease inhibitor
Signal
Transmembrane
Zinc
770
UniProtKB:ACH10_CHICK
CHRNA10
Calcium
Calcium
channel
Calcium

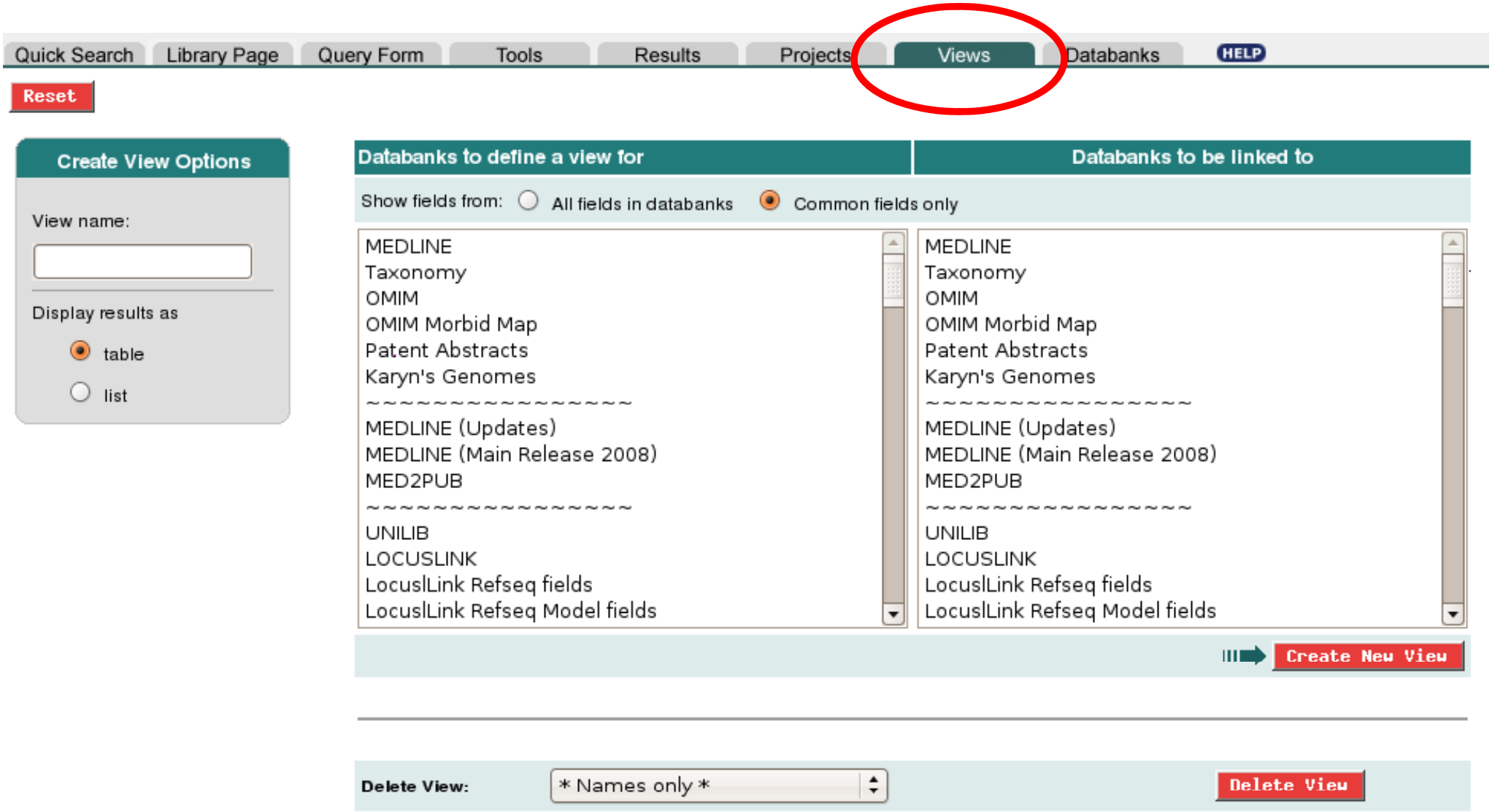
标准查询界面中直接定制视图



采用定制视图的结果页面

UniProtKB/Swiss-Prot	ID	EntryName	Description	Organism
UniProtKB/Swiss-Prot:ACH10_HUMAN	ACH10_HUMAN	ACH10_HUMAN	Neuronal acetylcholine receptor subunit alpha-10 precursor (Nicotinic acetylcholine receptor subunit alpha 10) (NACHR alpha 10).	Homo sapiens (Human). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.
UniProtKB/Swiss-Prot:ACHA2_HUMAN	ACHA2_HUMAN	ACHA2_HUMAN	Neuronal acetylcholine receptor subunit alpha-2 precursor.	Homo sapiens (Human). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria;

对定制结果视图的管理


Quick Search
Library Page
Query Form
Tools
Results
Projects
Views
Databanks
HELP
Reset
Create View Options
View name:
Display results as
table
list
Databanks to define a view for
Databanks to be linked to
Show fields from:
All fields in databanks
Common fields only
MEDLINE
Taxonomy
OMIM
OMIM Morbid Map
Patent Abstracts
Karyn's Genomes
~~~~~~~~~~~~~~~~~~
MEDLINE (Updates)
MEDLINE (Main Release 2008)
MED2PUB
~~~~~~~~~~~~~~~~~~
UNILIB
LOCUSLINK
LocuslLink Refseq fields
LocuslLink Refseq Model fields
Create New View
Delete View:
* Names only *
Delete View

SRS数据分析(以ClustalW为例)

Reset

Query ((([swissprot-Description:acetylcholine*] & [swissp

found 19 entries

Apply Options to:

- selected results only
- unselected results only

Result Options

Launch analysis tool:

ClustalW Launch

Show tools relevant to these results: Tools

Link to related information: Link

Save results: Save

Display Options

View results using: SwissView

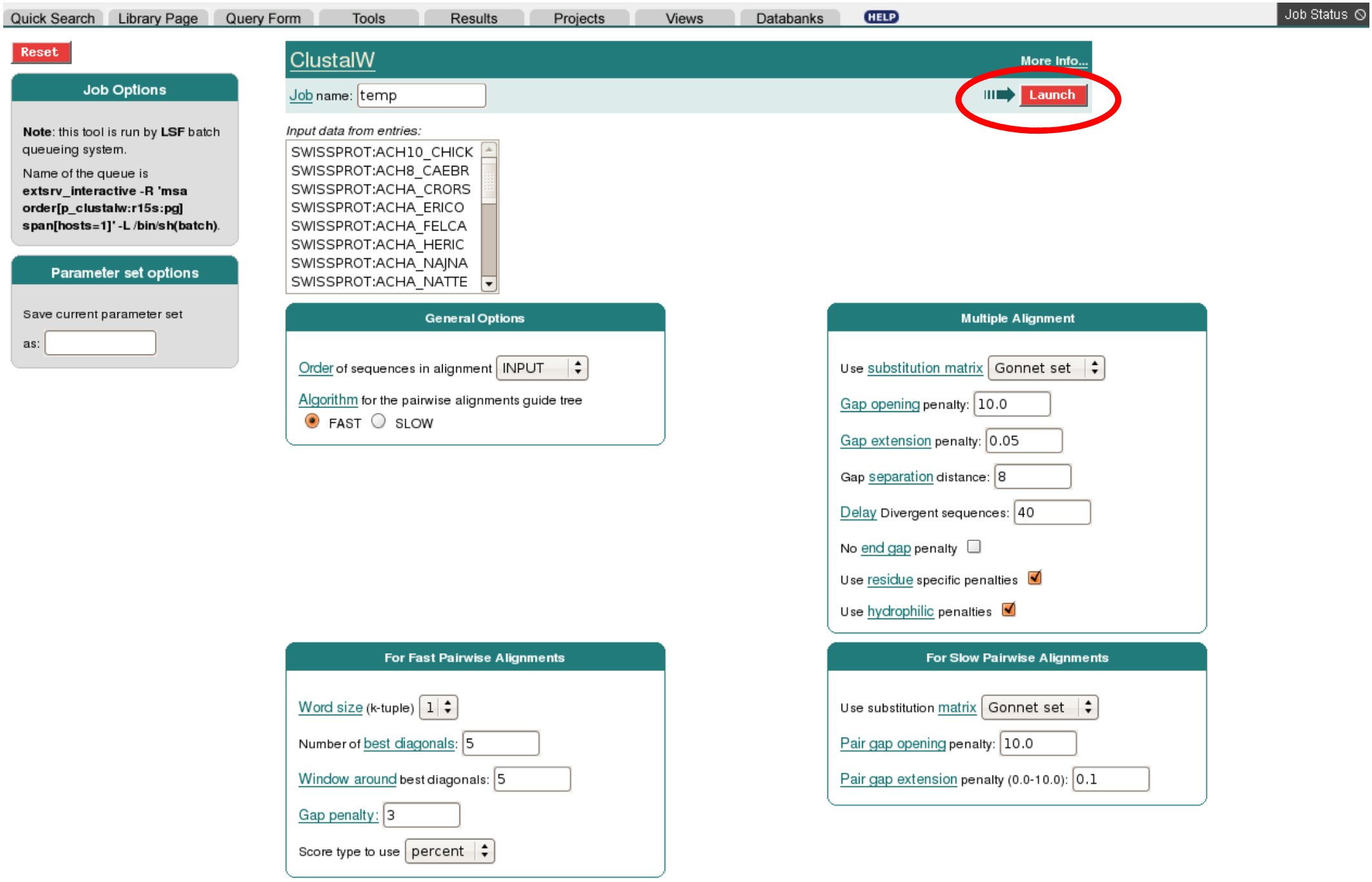
Sort results by: unsorted

- ascending
- descending

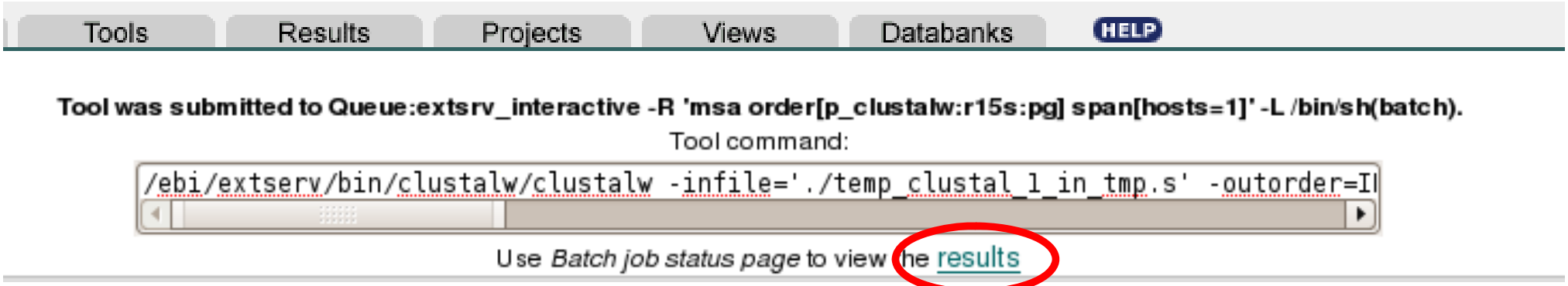
Show 30 results per page

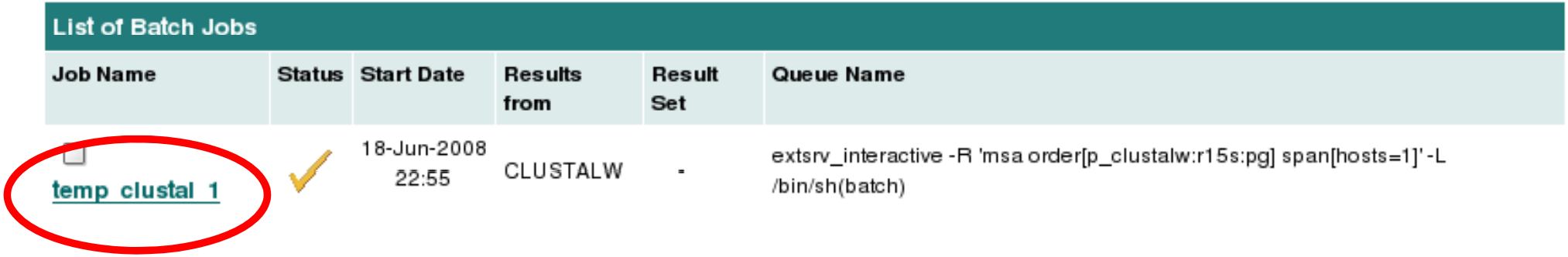
UniProtKB/Swiss-Prot	Accession	Description	GeneName	Organism	Authors
UniProtKB/Swiss-Prot:ACH10_CHICK	Q9I8C7	Neuronal acetylcholine receptor subunit alpha-10 precursor (Nicotinic acetylcholine receptor subunit alpha 10) (NACHR alpha 10) (Alpha 10 nAChR).	CHRNA10	Gallus gallus Chicken Eukaryota Metazoa Chordata Craniata Vertebrata Euteleostomi Archosauria Dinosauria Saurischia Theropoda Coelurosauria Aves Neognathae Galliformes Phasianidae Phasianinae Gallus	Barabino,B. Boulter,J. Alliod,C. Ballivet,M.
UniProtKB/Swiss-Prot:ACH8_CAEBR	Q60S81	Neuronal acetylcholine receptor subunit eat-2 precursor.	eat-2	Caenorhabditis briggsae Eukaryota Metazoa Nematoda Chromadorea Rhabditida Rhabditoidea Rhabditidae Peloderinae Caenorhabditis	Stein,L.D. Bao,Z. Blasiar,D. Blumenthal,T. Brent,M.R. Chen,N. Chinwalla,A. Clarke,L. Clee,C. Coghlan,A. Coulson,A. D'Eustachio,P. Fitch,D.H.A. Fulton,L.A. Fulton,R.E. Griffiths- Jones,S. Harris,T.W. Hillier,L.W.

选择分析工具参数


Quick Search
Library Page
Query Form
Tools
Results
Projects
Views
Databanks
HELP
Job Status
Reset
Job Options
Note: this tool is run by LSF batch queueing system.
Name of the queue is extsrv_interactive -R 'msa order[p_clustalw:r15s:pg] span[hosts=1]' -L /bin/sh(batch).
Parameter set options
Save current parameter set as:
ClustalW
More Info...
Job name: temp
Launch
Input data from entries:
SWISSPROT:ACH10_CHICK
SWISSPROT:ACH8_CAEBR
SWISSPROT:ACHA_CRORS
SWISSPROT:ACHA_ERICO
SWISSPROT:ACHA_FELCA
SWISSPROT:ACHA_HERIC
SWISSPROT:ACHA_NAJNA
SWISSPROT:ACHA_NATTE
General Options
Order of sequences in alignment INPUT
Algorithm for the pairwise alignments guide tree
FAST SLOW
Multiple Alignment
Use substitution matrix Gonnet set
Gap opening penalty: 10.0
Gap extension penalty: 0.05
Gap separation distance: 8
Delay Divergent sequences: 40
No end gap penalty
Use residue specific penalties
Use hydrophilic penalties
For Fast Pairwise Alignments
Word size (k-tuple) 1
Number of best diagonals: 5
Window around best diagonals: 5
Gap penalty: 3
Score type to use percent
For Slow Pairwise Alignments
Use substitution matrix Gonnet set
Pair gap opening penalty: 10.0
Pair gap extension penalty (0.0-10.0): 0.1

执行数据分析和查看结果





File Name: ./temp_job1.clustalw

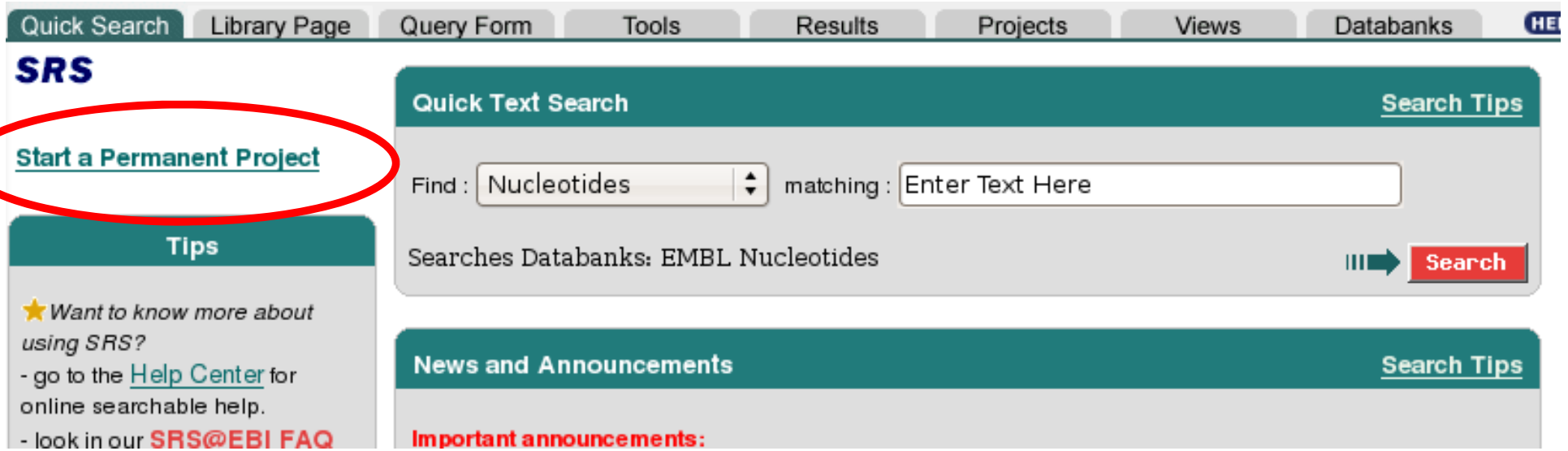
Content of the tool output file:

```
! Sequence: SWISSPROT:ACH10_CHICK Neuronal acetylcholine receptor subunit alpha-10 p...
! Sequence: SWISSPROT:ACH8_CAEBR Neuronal acetylcholine receptor subunit eat-2 prec...
! Sequence: SWISSPROT:ACHA_CRORS Acetylcholine receptor subunit alpha (Fragment).
! Sequence: SWISSPROT:ACHA_ERICO Acetylcholine receptor subunit alpha (Fragment).
! Sequence: SWISSPROT:ACHA_FELCA Acetylcholine receptor subunit alpha (Fragment).
! Sequence: SWISSPROT:ACHA_HERIC Acetylcholine receptor subunit alpha (Fragment).
! Sequence: SWISSPROT:ACHA_NAJNA Acetylcholine receptor subunit alpha (Fragment).
! Sequence: SWISSPROT:ACHA_NATTE Acetylcholine receptor subunit alpha (Fragment).
! Sequence: SWISSPROT:ACHB_CHICK Acetylcholine receptor subunit beta precursor (Fra...
! Sequence: SWISSPROT:ACHD_CHICK Acetylcholine receptor subunit delta precursor.
! Sequence: SWISSPROT:ACHG_CHICK Acetylcholine receptor subunit gamma precursor.
! Sequence: SWISSPROT:ACM2_CHICK Muscarinic acetylcholine receptor M2.
! Sequence: SWISSPROT:ACM2_PANTR Muscarinic acetylcholine receptor M2 (Fragment).
! Sequence: SWISSPROT:ACM3_GORGO Muscarinic acetylcholine receptor M3.
! Sequence: SWISSPROT:ACM3_PANTR Muscarinic acetylcholine receptor M3.
! Sequence: SWISSPROT:ACM3_PIG Muscarinic acetylcholine receptor M3.
! Sequence: SWISSPROT:ACM4_XENLA Muscarinic acetylcholine receptor M4.
! Sequence: SWISSPROT:ACR5_CAEEL Acetylcholine receptor subunit alpha-type acr-5 pr...
! Sequence: SWISSPROT:CUP4_CAEBR Acetylcholine receptor-like protein cup-4 precurso...
CLUSTAL W (1.83) multiple sequence alignment


ACH10_CHICK      ------------------------------------------------------------
ACH8_CAEBR       ------------------------------------------------------------
ACHA_CRORS       ------------------------------------------------------------
ACHA_ERICO       ------------------------------------------------------------
ACHA_FELCA       ------------------------------------------------------------
ACHA_HERIC       ------------------------------------------------------------
ACHA_NAJNA       ------------------------------------------------------------
ACHA_NATTE       ------------------------------------------------------------
ACHB_CHICK       ------------------------------------------------------------
ACHD_CHICK       ------------------------------------------------------------
ACHG_CHICK       ------------------------------------------------------------
ACM2_CHICK       ------------------------------------------------------------
ACM2_PANTR       ------------------------------------------------------------
ACM3_GORGO       ------------------------------------------------------------
ACM3_PANTR       ------------------------------------------------------------
ACM3_PIG         ------------------------------------------------------------
ACM4_XENLA       ------------------------------------------------------------
ACR5_CAEEL       MLPNIILILLLIRYCSCGAGSRVYEKYGKQVQLSPATTSSYREWYDSNREHSTRNNTNVDD
CUP4_CAEBR       ------------------------------------------------------------MR


ACH10_CHICK      --------------MGSAPLPACLLALSLAGTVLAPGCGAAQGRLAHKLLHDLFANYSSA
```

SRS项目

- 项目是指一组查询、分析等操作的总和
- 在没有登录(创建项目)时进行查询，SRS会自动创建一个临时项目，并将查询放入其中
- 临时项目在用户离开SRS网站时会被自动清除
- 创建(非临时)项目可以将查询操作永久保存，可以在不同时间、不同终端上来继续操作
- 一些分析工具的配置参数，也能够被保存到项目中来

创建项目


Quick Search
Library Page
Query Form
Tools
Results
Projects
Views
Databanks
SRS
Start a Permanent Project
Quick Text Search
Search Tips
Find :
Nucleotides
matching :
Enter Text Here
Searches Databanks: EMBL Nucleotides
Search
Tips
Want to know more about using SRS?
- go to the Help Center for online searchable help.
- look in our SRS@EBI FAQ
News and Announcements
Search Tips
Important announcements:

The page at http://srs.ebi.ac.uk says:
Enter your SRS user name:
yanlinlin82
Cancel
OK

SRS项目管理界面

所有查询、分析的结果，以及创建的视图，都能在这里列举出来

本次内容总结

- SRS查询（快速、标准、扩展）
- SRS链接(Link)功能
- 对查询结果的分析
- SRS项目的使用

这些都只是SRS的基本功能。
还有更多的功能尚待进一步发掘使用。

进一步发掘方向

- 可进一步了解SRS查询语言，它能帮助我们更准确地查找到所需数据
- 多了解SRS中包含的各个数据库的含义，了解这些数据库之间链接的含义
- 尝试SRS自带的各种分析工具
- 还可以使用其他网站的SRS系统（因为各个网站所关联的库、包含的工具，以及数据库更新速度都有所不同）
- ……

展望SRS

- SRS最早开发目的是一个以序列提取为目的的工具
- 在整合了各种数据库和工具之后，如今已逐步发展成为一个全面的生物信息研究平台
- SRS作为Biowisdom公司的一个重要产品，其功能肯定也会更进一步得到改进和提升

SRS最新情况

截至2008年6月21日：

- 关联了1119个不同的数据库
- 集成了181个不同的数据分析工具
- 最新版本8.3，新增功能：
  - 查询的提醒功能
  - 为数据添加标注(annotation)
  - 定制工作流(Workflow)

（信息来自：http://www.biowisdom.com/）

谢谢！