

数据库检索系统SRS

罗静初

2011年10月14日

SRS — Sequence Retrieval System

- 90年代初 — 由EMBL Etzold和Argos开发，用于检索核酸和蛋白质序列数据库
- 90年代中 — EBI成立，Etzold主持SRS研究开发，并作为EBI主要数据库检索系统
- 90年代末 — LION公司成立，SRS为该公司主要生物信息产品，对学术单位免费
- 2007年 — 英国剑桥BioWisdom公司收购SRS，学术单位每年3000欧元（8.2版开始）

国际SRS服务器

单位	网 址	版本	库	程序
BIPS	bips.u-strasbg.fr/srs/	8.3	33	3
EMBL	srs.embl.de	8.3	85	6
AFFRC	srs.dna.affrc.go.jp/srs8/	8.1	46	2
SAS	www.embnet.sk:8080/srs81/	8.1	47	EMBOSS
CEINGE	bioinfo.ceinge.unina.it/srs7131/	7.1	59	无
EBI	srs.ebi.ac.uk	7.1	117	EMBOSS
NBIC	srs.bioinformatics.nl	7.1	53	EMBOSS
CBP	proteomics.leeds.ac.uk/srs71/	7.1	29	20
CABRI	srs71.cabri.org/	7.1	42	无
DKF	www.dkfz.de/srs/	7.1	902	无
IST	srs.o2i.it/srs71/	7.1	13	无
ICG	srs6.bionet.nsc.ru/srs6/	6.1	89	无
PBIL	srs-pbil.ibcp.fr/	6.1	4	2

国际SRS服务器提供单位

- BIPS** — Bioinformatics Platform of Strasbourg
- EMBL** — European Molecular Biology Laboratory
- AFFRC** — Agriculture, Forestry & Fisheries Research Council
- SAS** — Slovak Academy of Sciences
- CEINGE** — Biotechnologie Avanzate, Naples
- EBI** — European Bioinformatics Institute
- NBIC** — Netherlands Bioinformatics Center
- CBP** — Clinical and Biomedical Proteomics group, Univ of Leeds
- CABRI** — Common Access to Biological Resources and Information
- DKFZ** — German Cancer Research Center
- IST** — National Cancer Research Institute
- ICG** — Inst of Cytology and Genetics, Novosibirsk
- PBIL** — Pole Bio-Informatique Lyonnais, Lyon

SRS检索方法

- **快速检索 — Quick Search**
- **标准检索 — Standard Query**
- **扩展检索 — Extended Query**

SRS快速检索界面

SRS

Temporary Project

1nE_v1ehOng

Tips

★ *Want to know more about using SRS?*

- go to the [Help Center](#) for online searchable help.
- look in our [SRS@EBI FAQ](#) for answers to commonly asked questions

★ *Linking to SRS?*

- Please read our [Linking to SRS](#) guide for important information regarding linking to our SRS server.

★ [Public SRS servers worldwide](#)

Quick Text Search

[Search Tips](#)

Find : Nucleotides matching : Enter Text Here

Search: Nucleotides
Proteins
Structures
Protein Families
Literature
Genome
Mutations
Metabolic Pathways

Search

News

[Search Tips](#)

Important News:

11.10.11 Please Note: One of our data centres will be offline from Friday 21 October 2011 at 14.00 (GMT+1) to Monday 24 October, 12 noon (GMT+1). During this period this SRS server will not be available.

08.09.11 EMBL-Bank Release 109 is now on-line ([release notes](#), [data notes](#)). **Note:** since release 108 the Expressed Sequence Tag (EST) entries from the release have been moved out of [EMBLRELEASE](#) into [EMBLESTRELEASE](#).

03.06.11 EMBL-Bank Release 108 is now on-line ([release notes](#), [data notes](#)). **Note:** the Expressed Sequence Tag (EST) entries from the release have been moved out of [EMBLRELEASE](#) in to [EMBLESTRELEASE](#).

SRS快速检索数据库

- Nucleotides（核酸序列）
- Proteins（蛋白质序列）
- Structure（蛋白质结构）
- Protein Families（蛋白质家族）
- Literature（文献）
- Genomes（基因组）
- Mutations（突变体）
- Metabolic Pathways（代谢途径）

斑头雁血红蛋白—快速检索

操作方法

- 在浏览器地址栏中键入srs.ebi.ac.uk，进入SRS系统快速检索界面
- 将Quick Text Search栏下方Find下拉菜单中数据库Nucleotides改为Proteins
- 在Enter Text Here文本输入框中输入bar-headed goose hemoglobin
- 点击Search按钮

返回结果

斑头雁血红蛋白HBAD_ANSIN、HBA_ANSIN和HBB_ANSIN三个亚基，灰雁血红蛋白三个亚基，以及加拿大黑雁血红蛋白等几十个条目。

检索条件

```
((([swissprot-ALLTEXT:bar-headed*] &[swissprot-ALLTEXT:goose*]) &[swissprot-ALLTEXT:hemoglobin*]) |[swissprot-ALLTEXT:bar-headed goose hemoglobin*])
```

SRS数据库选择界面

Reset

Quick Search

Search Options

1. Select the **databanks** you want to search
2. Enter your **search terms** in the **Quick Search** box, or choose a **query form** from below

[Standard Query Form](#)

[Extended Query Form](#)

You can **browse** through all the **entries** in any **databanks**. First, **select** the **databanks** you want to browse, then click:

[Browse Entries](#)

Tips

- ▶ bookmark this [link](#) to return to your project
- ▶ [Linking to SRS?](#) - Please read our [Linking to SRS](#) guide for important information regarding linking to our SRS server.

BookMarkLets

[About BookmarkLets](#)

• Protein Seq

Available Databanks

[+ Expand all](#) [- Collapse all](#)

Show databanks tooltips:

Literature, Bibliography and Reference Databases

- [all](#) [MEDLINE](#) [Taxonomy](#) [OMIM](#)
 [OMIM Morbid Map](#) [Patent Abstracts](#) [Karyn's Genomes](#)
 [Patent Equivalents](#)

Literature, Bibliography and Reference Databases - subsections

- [all](#) [MEDLINE \(Updates\)](#) [MEDLINE \(Main Release 2011\)](#) [MED2PUB](#)

Gene Dictionaries and Ontologies

Nucleotide sequence databases

- [all](#) [EMBL](#) [Patent DNA](#) [EMBL \(Contig\)](#)
 [EMBL \(Contigs expanded\)](#) [EMBL \(Coding Sequences\)](#) [EMBL ID/Accession Mapping](#)
 [EMBL MGA](#) [IMG/HLA](#)
 [IPD-KIR](#) [IPD-MHC](#) [Genome Reviews](#)
 [GR Genes](#) [GR Transcripts](#) [RefSeq Genome](#)
 [LiveLists](#) [Patent DNA NRL1](#) [Patent DNA NRL2](#)

Nucleotide sequence databases - subsections

- [all](#) [EMBL \(Updates\)](#) [EMBL \(Release\)](#) [EMBL \(Whole Genome Shotgun\)](#)
 [EMBL \(Whole Genome Shotgun release\)](#) [EMBL \(Whole Genome Shotgun updates\)](#) [EMBL \(Contig release\)](#)
 [EMBL \(Contig updates\)](#) [EMBL \(Contigs expanded release\)](#) [EMBL \(Contigs expanded updates\)](#)
 [EMBL \(EST Release\)](#) [EMBL \(Release, Deleted\)](#) [EMBL \(Whole Genome Shotgun Masters\)](#)
 [ENA Project](#) [RefSeq Genome \(Release\)](#) [RefSeq Genome \(Updates\)](#)

Nucleotide related databases

UniProt Universal Protein Resource

- [all](#) [UniProtKB](#) [UniProtKB/Swiss-Prot](#) [UniProtKB/TrEMBL](#) [UniRef100](#) [UniRef90](#)
 [UniRef50](#) [UniParc](#)

Other protein sequence databases

SRS标准检索界面 (Swiss-Prot)

Quick Search Library Page **Query Form** Tools Results Projects Views Databanks

Reset search [UniProtKB/Swiss-Prot](#)

Search Options

Combine search terms

with:

Use wildcards

Get results of type:

Result Display Options

View results using:

or

Create a view

Fields you can search

Your search terms

In a single field, you can separate multiple values by: &, | or !

<input type="text" value="i"/> ID	<input type="text" value="HBA_HUMAN"/>
<input type="text" value="i"/> ID	<input type="text" value="HBA_MOUSE"/>
<input type="text" value="i"/> ID	<input type="text" value="HBA_RAT"/>
<input type="text" value="i"/> AllText	<input type="text" value=""/>

Create a view

Select the fields you want displayed in your view and choose the format

Choose 1 or more fields:

Display As: Table List

Sequence Format:

斑头雁血红蛋白—标准检索

操作方法

- 选择UniProt/Swiss-Prot蛋白质序列数据库
- 点击Search Options栏中Standard Query Form按钮
- 将Fields you can search第一个下拉菜单中All Text改为Species，在Your Search Term中输入bar-headed goose
- 将Fields you can search第二个下拉菜单中All Text改为Description，在Your Search Term中输入Hemoglobin
- 点击Search按钮

返回结果

返回斑头雁血红蛋白序列HBAD_ANSIN、HBA_ANSIN和HBB_ANSIN三个亚基。

检索条件

```
((([swissprot-Species:bar-headed*] &[swissprot-Species:goose*]) |[swissprot-Species:bar-headed goose*]) &[swissprot-Description:hemoglobin*])
```

人血红蛋白 α 亚基—标准检索

操作方法

- 选择UniProt/Swiss-Prot蛋白质序列数据库
- 点击Search Options栏中Standard Query Form按钮
- 将Fields you can search第一个下拉菜单中All Text改为Species，在Your search term输入框中输入homo sapiens
- 将Fields you can search第二个下拉菜单中All Text改为Description，在Your search term输入框中输入Hemoglobin alpha chain
- 点击Search按钮

返回结果

返回人血红蛋白亚基序列条目HBA_HUMAN。

检索条件

```
([swissprot-Species:human*] & ((([swissprot-Description:hemoglobin*] & [swissprot-Description:alpha*]) & [swissprot-Description:chain*]) | [swissprot-Description:hemoglobin alpha chain*]))
```

人血红蛋白所有亚基—标准检索

操作方法

- 选择UniProt/Swiss-Prot蛋白质序列数据库
- 点击Search Options栏中Standard Query Form按钮
- 将Fields you can search第一个下拉菜单中All Text改为Species，在Your Search Term中输入homo sapiens
- 将Fields you can search第二个下拉菜单中All Text改为Description，在Your Search Term中输入Hemoglobin !receptor
- 点击Search按钮

返回结果

返回人血红蛋白9个亚基序列条目，包括alpha珠蛋白基因家族4个亚基：HBA_HUMAN、HBT_HUMAN、HBZ_HUMAN和HBM_HUMAN，beta球蛋白基因家族5个亚基：HBB_HUMAN、HBD_HUMAN、HBE_HUMAN，以及HBAG1_HUMAN和HBAG2_HUMAN

检索条件

```
([swissprot-Species:human*] &([swissprot-Description:hemoglobin*] ! [swissprot-Description:receptor*]))
```

Swiss-Prot 常规注释信息 (一)

Fields you can search	Your search terms	Create a view
In a single field, you can separate multiple values by &, , !		
AllText	<input type="text"/>	
AllIDs	<input type="text"/>	
ID	<input type="text"/>	<input type="checkbox"/>
EntryName	<input type="text"/>	<input type="checkbox"/>
Data Class	<input type="text"/>	<input type="checkbox"/>
AccessionNumber	<input type="text"/>	<input type="checkbox"/>
Primary Accession Number	<input type="text"/>	<input type="checkbox"/>
Sequence Version	<input type="text"/>	<input type="checkbox"/>
Creation Date	select ▾ 1 ▾ Jan ▾ YYYY <input type="text"/> 1 ▾ Jan ▾ YYYY <input type="text"/>	<input type="checkbox"/>
Seq Mod Date	select ▾ 1 ▾ Jan ▾ YYYY <input type="text"/> 1 ▾ Jan ▾ YYYY <input type="text"/>	<input type="checkbox"/>
Annot Mod Date	select ▾ 1 ▾ Jan ▾ YYYY <input type="text"/> 1 ▾ Jan ▾ YYYY <input type="text"/>	<input type="checkbox"/>
Description	<input type="text"/>	<input type="checkbox"/>
Gene Name	<input type="text"/>	<input type="checkbox"/>
Synonym	<input type="text"/>	<input type="checkbox"/>
Ordered Locus Names	<input type="text"/>	<input type="checkbox"/>
ORF Names	<input type="text"/>	<input type="checkbox"/>
ECNumber	<input type="text"/>	<input type="checkbox"/>
Organism Name	<input type="text"/>	<input type="checkbox"/>
Species	<input type="text"/>	<input type="checkbox"/>
Taxonomy	<input type="text"/>	<input type="checkbox"/>
Organelle	<input type="text"/>	<input type="checkbox"/>
NCBI TaxId	<input type="text"/>	<input type="checkbox"/>
TaxCount	>= ▾ <input type="text"/> <= ▾ <input type="text"/>	<input type="checkbox"/>
Organism Host TaxId	<input type="text"/>	<input type="checkbox"/>
Organism Host Name	<input type="text"/>	<input type="checkbox"/>

Swiss-Prot 常规注释信息 (二)

[Organism Host Name](#)

[ProteinExistence](#)

- or
 and

[Keywords](#)

[ProteinID](#)

[Sequence Length](#)

[MolWeight](#)

[crc](#)

[DBxref](#)

[MedlineID](#)

[swProtName](#)

[Isoform](#)

[DBLink](#)

[SeqCrcDigest](#)

1: evidence at protein level
2: evidence at transcript level
3: inferred from homology
4: predicted
5: uncertain

>= <=

>= <=



Swiss-Prot 文献

References subentry fields References ▾

View results using: ▶▶▶ **Search**

Authors	<input type="text"/>	<input type="checkbox"/>
Title	<input type="text"/>	<input type="checkbox"/>
RefPosition	<input type="text"/>	<input type="checkbox"/>
RefGroup	<input type="text"/>	<input type="checkbox"/>
RefNumber	<input type="text" value=">="/> <input type="text"/> <input type="text" value="<="/> <input type="text"/>	<input type="checkbox"/>
RefCommentCode	<input type="text" value="plasmid strain tissue transposon"/>	<input type="checkbox"/>
<input checked="" type="radio"/> or <input type="radio"/> and		<input type="checkbox"/>
RefComment	<input type="text"/>	<input type="checkbox"/>
Journal	<input type="text"/>	<input type="checkbox"/>
VolumeNo	<input type="text" value=">="/> <input type="text"/> <input type="text" value="<="/> <input type="text"/>	<input type="checkbox"/>
FirstPage	<input type="text" value=">="/> <input type="text"/> <input type="text" value="<="/> <input type="text"/>	<input type="checkbox"/>
Year	<input type="text" value=">="/> <input type="text"/> <input type="text" value="<="/> <input type="text"/>	<input type="checkbox"/>
Citation	<input type="text"/>	<input type="checkbox"/>
SubmissionDate	<input type="text" value=">="/> <input type="text"/> <input type="text" value="<="/> <input type="text"/>	<input type="checkbox"/>
Patent	<input type="text"/>	<input type="checkbox"/>
PatentDate	<input type="text" value=">="/> <input type="text"/> <input type="text" value="<="/> <input type="text"/>	<input type="checkbox"/>
MedlineID	<input type="text"/>	<input type="checkbox"/>

Swiss-Prot 常规注释和交叉链接

Comments subentry fields Comments ▾

View results using: ▶▶▶ Search

CommentType

or and

allergen
alternative products
biophysicochemical properties
biotechnology
catalytic activity
caution
cofactor

Comment

Link subentry fields Link ▾

View results using: ▶▶▶ Search

DbName

DBxref

Swiss-Prot 序列注释和计数器

Features subentry fields Features ▾

View results using: ▶▶▶ Search

FtKey

- or
- and

- act_site ▲
- binding ≡
- ca_bind
- carbohyd
- chain
- coiled
- compbias ▼

FtLength

>= <=

FtDescription

Begin

>= <=

End

>= <=

FtResid

Counters subentry fields Counters ▾

View results using: ▶▶▶ Search

CountItem

CountN

>= <=

SRS扩展检索界面 (EMBL-1)

Quick Search	Library Page	Query Form	Tools	Results	Projects	Views	Databanks	HELP
------------------------------	------------------------------	----------------------------	-----------------------	-------------------------	--------------------------	-----------------------	---------------------------	----------------------

[Reset](#) search [EMBL](#)

Search Options

Combine search terms with:

Use wildcards

Get results of type:

Result Display Options

View results using:

or

Create a view using selected fields

Sequence Format:

Show results per page

Fields you can search	Your search terms	Create a view
In a single field, you can separate multiple values by &, , !		
AllText	<input type="text"/>	
AllIDs	<input type="text"/>	
ID	<input type="text"/>	<input type="checkbox"/>
Topology	<input type="checkbox"/> c <input type="checkbox"/> circular <input type="checkbox"/> l	<input type="checkbox"/>
	<input type="checkbox"/> linear	<input type="checkbox"/>
Molecule	<input type="checkbox"/> genomic dna <input type="checkbox"/> genomic ma	
	<input type="checkbox"/> mma <input type="checkbox"/> other dna	
	<input type="checkbox"/> other ma <input type="checkbox"/> rma	
	<input type="checkbox"/> transcribed ma <input type="checkbox"/> trna	<input type="checkbox"/>
	<input type="checkbox"/> unassigned dna <input type="checkbox"/> unassigned ma	
	<input type="checkbox"/> viral crna	
	<input type="checkbox"/> est <input type="checkbox"/> gss <input type="checkbox"/> htc <input type="checkbox"/> htg <input type="checkbox"/> pat <input type="checkbox"/> std	<input type="checkbox"/>
<input type="checkbox"/> sts <input type="checkbox"/> tsa <input type="checkbox"/> wgs		
Data Class	<input type="checkbox"/> env <input type="checkbox"/> fun <input type="checkbox"/> hum <input type="checkbox"/> inv <input type="checkbox"/> mam <input type="checkbox"/> mus	
	<input type="checkbox"/> phg <input type="checkbox"/> pln <input type="checkbox"/> pro <input type="checkbox"/> rod <input type="checkbox"/> syn <input type="checkbox"/> tgn	<input type="checkbox"/>
Division		

SRS扩展检索界面 (EMBL-2)

[Sequence Length](#)

>= <=

[Accession Number](#)

[Primary Accession Number](#)

[Sequence Version](#)

[Entry Creation Date](#)

select 1 Jan YYYY 1 Jan YYYY

[Entry Creation Release](#)

>= <=

[Last Update Date](#)

select 1 Jan YYYY 1 Jan YYYY

[Last Update Release](#)

>= <=

[Description](#)

[Genome Project ID](#)

[Keywords](#)

[Organism Name](#)

[Taxon](#)

[NCBI TaxId](#)

[Organelle](#)

[Link](#)

[UniLibID](#)

[SubAddress](#)

[Patent](#)



SRS扩展检索界面 (EMBL-3)

References subentry fields References ▾

View results using: * Names only * ▶▶▶ **Search**

RefPosition	<input type="text"/>	<input type="checkbox"/>
PubMedID	<input type="text"/>	<input type="checkbox"/>
RefGroup	<input type="text"/>	<input type="checkbox"/>
Authors	<input type="text"/>	<input type="checkbox"/>
Title	<input type="text"/>	<input type="checkbox"/>
Journal	<input type="text"/>	<input type="checkbox"/>
VolumeNo	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="checkbox"/>
FirstPage	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="checkbox"/>
Year	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="checkbox"/>

Features subentry fields Features ▾

View results using: * Names only * ▶▶▶ **Search**

[FtKey](#)

- or
 and

[FtQualifier](#)

- or

- 10_signal
- 35_signal
- 3'utr
- 5'utr
- c_region
- caat_signal
- cds

- allele
- anticodon
- artificial_location
- bio_material

SRS扩展检索界面 (EMBL-4)

	cas	
	allele	
	anticodon	
	artificial_location	
	bio_material	<input type="checkbox"/>
	bound_moiety	
	cell_line	
	cell_type	
FtQualifier		
<input checked="" type="radio"/> or		
<input type="radio"/> and		
ProteinID		<input type="checkbox"/>
Latitude	>= <input type="text"/> <= <input type="text"/>	<input type="checkbox"/>
Longitude	>= <input type="text"/> <= <input type="text"/>	<input type="checkbox"/>
Gene		<input type="checkbox"/>
FtDescription		<input type="checkbox"/>
FtBegin	>= <input type="text"/> <= <input type="text"/>	<input type="checkbox"/>
FtEnd	>= <input type="text"/> <= <input type="text"/>	<input type="checkbox"/>
FtLength	>= <input type="text"/> <= <input type="text"/>	<input type="checkbox"/>
Counters subentry fields		Counters ▾
View results using:	* Names only * ▾	Search
CountedItem		<input type="checkbox"/>
CountedN	>= <input type="text"/> <= <input type="text"/>	<input type="checkbox"/>

SRS高级检索

- 逻辑运算符 — &、|和!, 分别表示与、或、非, 默认为与
- 圆括号 — 改变检索优先次序, 默认由左至右
- 中括号 — 限定检索字段, 如检索作者Rice, 可用Rice[AU]
- 通配符 — 自动添加, 如输入peptide, 返回结果包括peptide和peptides
- 双引号 — 取消通配功能, 如输入“peptide”, 则不返回含peptides的条目; 限定多单词短语, 如“disulfide bridge”
- 冒号 — 限定数值型字段检索范围, 包括序列长度、日期等
- 正则表达式 — 限定检索范围, 如输入/^phos/, 返回含以“phos”起始的条目, 输入/ase\$/返回含以“ase”结尾的条目

SRS

Temporary Project

1nE_v1ehOng

Tips

★ *Want to know more about using SRS?*

- go to the [Help Center](#) for online searchable help.
- look in our [SRS@EBI FAQ](#) for answers to commonly asked questions

★ *Linking to SRS?*

- Please read our [Linking to SRS](#) guide for important information regarding linking to our SRS server.

★ [Public SRS servers worldwide](#)

Quick Text Search

[Search Tips](#)

Find : Nucleotides matching :

Search :

- Nucleotides
- Proteins
- Structures
- Protein Families
- Literature
- Genome
- Mutations
- Metabolic Pathways

News

[Search Tips](#)

Important News:

11.10.11 Please Note: One of our data centres will be offline from Friday 21 October 2011 at 14.00 (GMT+1) to Monday 24 October, 12 noon (GMT+1). During this period this SRS server will not be available.

08.09.11 EMBL-Bank Release 109 is now on-line ([release notes](#), [data notes](#)). **Note:** since release 108 the Expressed Sequence Tag (EST) entries from the release have been moved out of [EMBLRELEASE](#) into [EMBLESTRELEASE](#).

03.06.11 EMBL-Bank Release 108 is now on-line ([release notes](#), [data notes](#)). **Note:** the Expressed Sequence Tag (EST) entries from the release have been moved out of [EMBLRELEASE](#) in to [EMBLESTRELEASE](#).

biowisdom
SRS

SRS其它功能

- **进行数据分析 (Tools)**
EMBOSS、Phylip、HMMER和BLAST等
- **整合检索结果 (Results)**
查看检索过程，对检索结果进行组合、链接等处理
- **保存检索策略 (Project)**
可将检索过程保存到本地
- **定制输出界面 (Views)**
可根据不同数据库，定制检索结果显示方式和显示字段

SRS主要特点

- 统一的用户界面
- 高效的检索功能
- 灵活的交叉链接
- 方便的程序接口
- 开放的管理模式

参考文献

- Etzold T, Argos P. SRS--an indexing and retrieval tool for flat file data libraries. *Comput Appl Biosci*. 1993a Feb;9(1):49-57.
- Etzold T, Argos P. Transforming a set of biological flat file libraries to a fast access network. *Comput Appl Biosci*. 1993b Feb;9(1):59-64.
- Etzold T, Ulyanov A, Argos P. 1996. SRS: Information Retrieval System for Molecular Biology Data Banks. *Methods in Enzymology*, 266:114.
- Zdobnov EM, Lopez R, Apweiler R, Etzold T. The EBI SRS server--recent developments. *Bioinformatics*. 2002a Feb;18(2):368-73.
- Zdobnov EM, Lopez R, Apweiler R, Etzold T. The EBI SRS server--new features. *Bioinformatics*. 2002b Aug;18(8):1149-50.