

# 生物序列获取策略

刘萌萌

2022年3月22日



ACS  
International



A division of the  
American Chemical Society

# CAS SciFinder<sup>n</sup>中覆盖的内容合集



来源: <https://www.cas.org/cas-data>和<https://www.cas.org/about/cas-content>

# CAS SciFinder<sup>n</sup> 的生物序列解决方案

## -- Biosequences

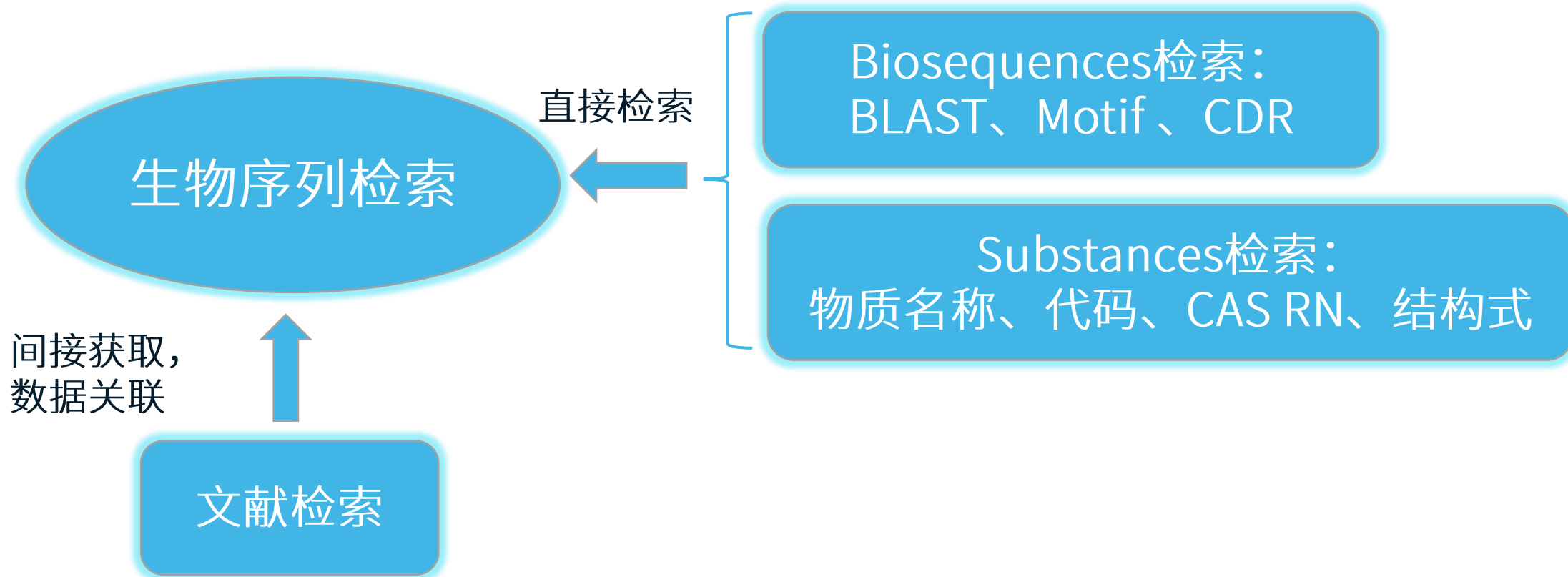
- 超过10亿条序列
- 涵盖专利、非专利文献披露的序列
- 专有的CAS人工标引的化学修饰序列
- NCBI中的序列
- 可实现新颖性、创造性检索
- 简洁直观的检索界面 & 便捷的筛选、可视化分析、相关信息获取和下载功能

# 大纲

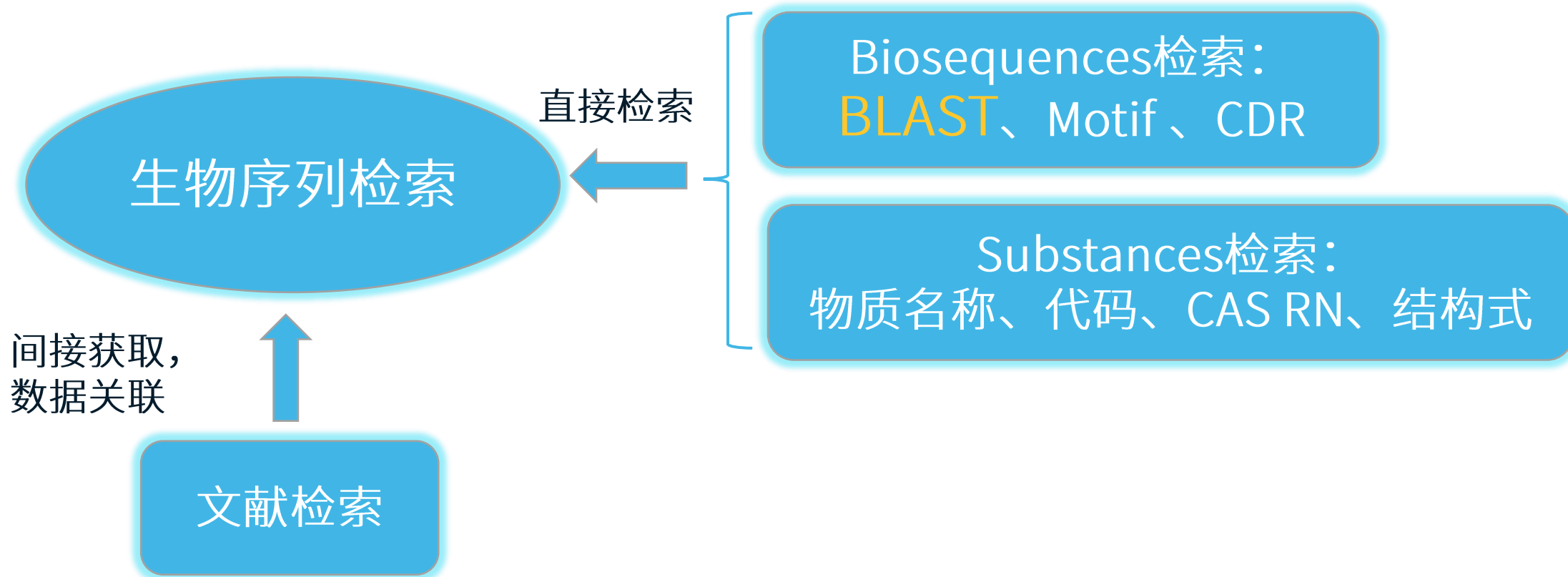
- CAS和CAS SciFinder<sup>n</sup>简介
- 生物序列及相关信息的获取策略
  - 序列获取方法
  - 序列检索结果的精炼和分析 & 序列相关信息的获取
  - 序列相关抗体偶联药物 (ADC) 信息的获取
- 在线演示以及Q&A



# 生物序列获取方法



# 生物序列获取方法



# BLAST检索

CAS SciFinder<sup>®</sup> ★ Saved 🕒 History 👤 Account

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

## Biosequences

Enter a protein or nucleotide string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST | CDR | Motif | Upload Sequence | Clear Search

```
ATGCAGATCCCACAGGCGCCTGGCCAGTCGTCTGGCCGGTGTACAACCTGGGCTGGCGCCAGGATGGTTCTTAGACTC
CCCAGACAGGCCCTGGAACCCCCACCTTCTCCCCAGCCCTGCTCGTGGTGACCGAAGGGGACACGCCACCTTACCT
GCAGCTTCTCCAACACATCGGAGAGCTTCGTGCTAACTGGTACCGCATGAGCCCAAGCAGCAGGACAAGCTGGCC
GCCTTCCCCGAGGACCGAGCCAGCCGGCCAGGACTGCCGCTCCGTGTACACAACCTGCCAACGGGCGTGACTTCCA
CATGAGCGTGGTCAGGGCCGGCGCAATGACAGCGGCACCTACCTCTGTGGGGCCATCTCCCTGGCCCCAAGGCGCAGA
TCAAAGAGAGCTGCGGGCAGAGCTCAGGGTGACAGAGAGAAGGGCAGAAGTGCCACAGCCACCCAGCCCTCACC
AGGCCAGCCGGCCAGTTCAAACCTGGTGGTTGGTGTCTGTGGCCGGCTGCTGGGCAGCTGGTGTCTAGTCTGGGT
CCTGGCCGTATCTGCTCCCGGGCCGCACGAGGGACAATAGGAGCCAGGCGCACCGCCAGCCCTGAAGGAGGACCCCT
CAGCCGTGCTGTGTTCTCTGTGGACTATGGGGAGCTGGATTTCCAGTGGCGAGAGAAGACCCGGAGCCCCCGTGCC
TGTGTCCCTGAGCAGACGGAGTATGCCACCATTTGCTTTCTAGCGGAATGGGCACCTCATCCCCGCCCGCAGGGGCTC
AGCTGACGGCCCTCGGAGTGCCAGCCACTGAGGCCTGAGGATGGACACTGCTCTTGGCCCTC
```

Sequence Type: **Nucleotide** | Protein

Search Within:  Nucleotides  Proteins

Include NCBI Sequences

Limit Total Sequence Results to: 20000

**Start Biosequence Search**

[Advanced Biosequence Search](#) | [Adjust Parameters for Short Sequences](#) | [Reset All](#)

四种检索选择：  
Protein-Proteins  
Protein-Nucleotides  
Nucleotide-Nucleotides  
Nucleotide-Proteins  
可选择包含NCBI中的序列

### Recent Search History

View All Search History

March 21, 2022

**Biosequences**  
3:19 PM

Sequence Type: **Nucleotide**  
Search Within: **Nucleotides**  
NCBI Included: **Yes**  
BLAST Algorithm: **BLASTn**  
Alignment Identity: **50%**  
Query Coverage: **50%**

Results will expire on  
Apr 21, 2022.

```
ATGCAGATCCCACAGGCGCCTGGCCAGTCGTCTGGCCGGTGTACAACCTGGGCTGGCGCCAGGATGGTTCTTAGACTC
TGGTCTTAGACTCCCAGACAGGCCCTGGAACCCCCACCTTCTCCCCAGCCCTGCTCGTGGTG
ACCGAAGGGGACACGCCACCTTACCTGCGACTTCTCCAACACATCGGAGAGCTTCGTGCTAAAC
TGGTACCGCATGAGCCCGAGCAACAGAGGGACAAGCTGGCCGCTTCCCCGAGGACCGCAGCCAG
CCCGGCCAGGACTGCCGCTTCCGTGTACACAACCTGCCAACGGGCGTGACTTCCACATGAGCGTG
GTCAGGGCCCGGCGCAATGACAGCGGCACCTACCTCTGTGGGGCCATCTCCCTGGCCCCAAGGCG
CAGATCAAAGAGAGCTGCGGGCAGAGCTCAGGGTGACAGAGAGAAGGGCAGAAGTGCCACAGCC
CACCCAGCCCTCACCAGGCGAGCCGGCCAGTTCCAAACCTGGTGGTTGGTGTCTGTGGCCGGC
CTGCTGGGAGCCCTGGTGTGCTAGTCTGGGTCCTGGCCGTCTGCTCCCGGGCCGACAGGG
ACAATAGGAGCCAGGCGCACCGGCCAGCCCTGAAGGAGGACCCCTCAGCCGTGCTGTGTTCTC...
```

**View Results** | Edit Search | Complete

# BLAST检索结果页面

按需求选择不同方式排序

BLAST Search Details

Sequence Type: Nucleotide  
Search Within: Nucleotides  
BLAST Algorithm: BLASTn  
NCBI Included: Yes  
Alignment Identity: 50%  
Query Coverage: 50%  
E-Value: 10  
Match with Gaps?: No  
Gap Costs: Existence 5  
Extension 2  
Word Size: 11

Bioscape Analysis

Visually explore sequence similarity with a new tool.  
Learn more about Bioscape.

Create Bioscape Analysis

Filter by

^ E-Value

0 to  $10^6$

^ Query Coverage %

0 to 100

^ Subject Coverage %

0 to 100

^ Alignment Identity %

0 to 100

Biosequences (926)

Sort: Alignment Identity View: Expanded

References

Query Details ATGCAGATCCCACAGGCGCCCTGGCCAGTCGCTGGGCGGTGCTACAACCTGGGCTGGCGCCAGGATGGTTCTTAGACTCCCCA... View More

1 Alignment Identity: 100%

Query 1 864

Subject 1 1,747

Matches: 864  
Mismatches: 0

View Less

Alignment Subject References

Alignment Data  
BLAST Score: 1728  
E-Value: 0

```
S 1194 GAGGACCCCT CAGCCGTGCC TGTGTTCTCT GTGGACTATG GGGAGCTGGA TTTCCAGTGG CGAGAGAAGA 1263
Q 701 CCCC GGAGCC CCCC GTGCC TGTGTCCTG AGCAGACGGA GTATGCCACC ATTGTCTTTC CTAGCGGAAT 770
S 1264 CCCC GGAGCC CCCC GTGCC TGTGTCCTG AGCAGACGGA GTATGCCACC ATTGTCTTTC CTAGCGGAAT 1333
Q 771 GGGCACCTCA TCCCCGCC GCAGGGGCTC AGCTGACGGC CCTCGGAGTG CCCAGCCACT GAGGCCTGAG 840
S 1334 GGGCACCTCA TCCCCGCC GCAGGGGCTC AGCTGACGGC CCTCGGAGTG CCCAGCCACT GAGGCCTGAG 1403
Q 841 GATGGACACT GCTCTTGCC CCTC 864
```

2 Alignment Identity: 100%

Query 1 864

Matches: 864  
Mismatches: 0

- Alignment Identity
- E-Value
- Query Coverage
- Subject Coverage

直观呈现比对结果  
筛选和精炼序列结果  
浏览相关信息

# 通过筛选选项精炼序列结果

Filter by

^ E-Value

0 to  $10^6$

^ Query Coverage %

0 to 100

^ Subject Coverage %

0 to 100

^ Alignment Identity %

0 to 100

Apply Reset Filters

E-Value:

- 随机情况下，数据库中存在的比当前匹配分数更高的比对的数目。
- 匹配分数越高，E值越小。
- 因此选择较低的 E-Value 范围会显示更严格的匹配结果。

例子:

Query Sequence QQLLVVEEGG  
                  | | | | | | | | | |  
Subject Sequence QQLLVVEEIGS


Alignment

$$\text{Query Coverage (100\%)} = \frac{\text{Alignment Length}}{\text{Query Length}} = \frac{10}{10}$$

$$\text{Subject Coverage (91\%)} = \frac{\text{Alignment Length}}{\text{Subject Length}} = \frac{10}{11}$$

$$\text{Alignment Identity (90\%)} = \frac{\text{Number of Matches}}{\text{Alignment Length}} = \frac{9}{10}$$

# BLAST检索

★ Saved🕒 History👤 Account

## Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

## Biosequences

Enter a protein or nucleotide string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST | CDR | Motif |  |

```
ATGCAGATCCCACAGGCGCCCTGGCCAGTCGTCTGGGCGGTGCTACAACCTGGGCTGGCGCCAGGATGGTTCTTAGACTC
CCCAGACAGGCCCTGGAACCCCCACCTTCTCCCCAGCCCTGCTCGTGGTGACCGAAGGGGACAACGCCACCTTCACCT
GCAGCTTCTCCAACACATCGGAGAGCTTCGTGCTAAACTGGTACCGCATGAGCCCCAGCAACCAGACGGACAAGCTGGCC
GCCTTCCCCGAGGACCGCAGCCAGCCCGGCCAGGACTGCCGCTTCCGTGTCACACAACCTGCCAACGGGCGTGACTTCCA
CATGAGCGTGGTCAGGGCCCGGCGCAATGACAGCGGCACCTACCTCTGTGGGGCCATCTCCCTGGCCCCAAGGCGCAGA
TCAAAGAGAGCCTGCGGGCAGAGCTCAGGGTGACAGAGAGAAGGGCAGAAGTGCCACAGCCACCCCAAGCCCTCACCC
AGGCCAGCCGGCCAGTTCCAACCCCTGGTGGTGGTGTGCTGGGCGGCCCTGCTGGGCGAGCCTGGTGTGCTAGTCTGGGT
CCTGGCCGTCATCTGCTCCCGGGCCGCACGAGGGACAATAGGAGCCAGGGCGACCGGCCAGCCCTGAAGGAGGACCCCT
CAGCCGTGCCTGTGTTCTCTGTGGACTATGGGGAGCTGGATTTCCAGTGGCGAGAGAAGACCCCGAGCCCCCGTGCCC
TGTGTCCCTGAGCAGACGGAGTATGCCACCATTTGCTTTCTAGCGGAATGGGCACCTATCCCCCGCCCGAGGGGCTC
AGCTGACGGCCCTCGGAGTGCCAGCCACTGAGGCCTGAGGATGGACTGCTCTTGGCCCCCTC
```

Sequence Type:  Nucleotide  Protein

Search Within:  Nucleotides  Proteins

Include NCBI Sequences

Limit Total Sequence Results to:

[Adjust Parameters for Short Sequences](#) | [Reset All](#)

# BLAST高级检索参数设置

例子：空位

```
A T G T - - - T A T A C
| | | | | | | |
T A T G T G C G T A T A
```

Advanced Biosequence Search ^

Adjust Parameters for Short Sequences | Reset All

Alignment Identity % ?

50

值越高，比对上的序列的一致性越高。

Query Coverage % ?

50

值越高，检索序列比对上的比例越高。

BLAST Algorithm

BLASTn

MegaBLAST 查找非常相似的序列

BLASTn 支持相似度较低的序列

BLASTn-short 允许更短序列的比对（如短到7个碱基的序列）

Match with Gaps?

Yes  No

是否允许比对结果中出现空位

Word Size ?

11

起始相似性检索的最低匹配长度  
word size越大，检索越严格。

E-Value ?

10

E值越小，检索越严格。

Gap Costs ?

Existence 5 Extension 2 ▾

较高的gap costs导致与含有较少空位的序列匹配。

Reward for Match

Penalty for Mismatch ?

2, -3

匹配奖励，错配罚分。Reward/Penalty比值越高，检出序列越离散

Exclude Low

Complexity Regions ?

Yes  No

LCR如AAATAAAAAAAAAATAAAAAT，多个A会导致比对分数偏高，序列中有LCR时可考虑排除



# BLAST高级检索参数设置--氨基酸序列

Advanced Biosequence Search ^ Adjust Parameters for Short Sequences | Reset All

Alignment Identity % ?  
-

Match with Gaps?  
 Yes  No

Gap Costs ?  
Existence 11 Extension 1 ▾

Query Coverage % ?  
90

Word Size ?  
3 ▾

Scoring Matrix ?  
BLOSUM62 ▾

BLAST Algorithm  
BLASTp ▾

E-Value ?  
10 ▾

Exclude Low Complexity Regions ?  
 Yes  No

Sequence Type:  
 Nucleotide  Protein

Search Within:  
 Nucleotides  Proteins

# BLAST高级检索参数设置--氨基酸序列

## BLAST Algorithm

Alignment Identity % ⓘ -

Match with Gaps?  Yes  No

Gap Costs ⓘ Existence 11 Extension 1 ▾

Query Coverage % ⓘ 90

Word Size ⓘ 6

Scoring Matrix ⓘ BLOSUM62 ▾

BLAST Algorithm ▾ BLASTp-fast

E-Value ⓘ 10

Exclude Low Complexity Regions ⓘ  Yes  No

Alignment Identity % ⓘ -

Match with Gaps?  Yes  No

Gap Costs ⓘ Existence 11 Extension 1 ▾

Query Coverage % ⓘ 90

Word Size ⓘ 3

Scoring Matrix ⓘ BLOSUM62 ▾

BLAST Algorithm ▾ BLASTp

E-Value ⓘ 10

Exclude Low Complexity Regions ⓘ  Yes  No

Advanced Biosequence Search ^ Adjust Parameters for Short Sequences Reset All

Alignment Identity % ⓘ -

Match with Gaps?  Yes  No

Gap Costs ⓘ Existence 9 Extension 1

Query Coverage % ⓘ 90

Word Size ⓘ 2

Scoring Matrix ⓘ PAM30

BLAST Algorithm ▾ BLASTp-short

E-Value ⓘ 10

Exclude Low Complexity Regions ⓘ  Yes  No

# BLAST高级检索参数设置--氨基酸序列

## BLAST Algorithm

BLAST Algorithm: BLASTp-fast

BLAST Search Details

Sequence Type: Protein  
Search Within: Proteins  
BLAST Algorithm: BLASTp-fast  
NCBI Included: No  
Alignment Identity: -  
Query Coverage: 90%  
E-Value: 10  
Match with Gaps?: No  
Gap Costs: Existence 11  
Extension 1  
Word Size: 6

Biosequences (15)

References

Query Details > AYH97874.1 Sequenc

1

Query ①

BLAST Algorithm: BLASTp

BLAST Search Details

Sequence Type: Protein  
Search Within: Proteins  
BLAST Algorithm: BLASTp  
NCBI Included: No  
Alignment Identity: -  
Query Coverage: 90%  
E-Value: 10  
Match with Gaps?: No  
Gap Costs: Existence 11  
Extension 1  
Word Size: 3

Biosequences (424)

References

Query Details > AYH97874.1 Sequenc

1

Query ①

Alignment Identity %

Match with Gaps?  Yes  No

Gap Costs

Query Coverage %

Word Size

Scoring Matrix

BLAST Algorithm: BLASTp-short

E-Value

Exclude Low Complexity Regions  Yes  No

BLAST Search Details

Sequence Type: Protein  
Search Within: Proteins  
BLAST Algorithm: BLASTp-short  
NCBI Included: No  
Alignment Identity: -  
Query Coverage: 90%  
E-Value: 10  
Match with Gaps?: No  
Gap Costs: Existence 9  
Extension 1  
Word Size: 2

Biosequences (15)

References

Query Details > AYH97874.1 Sequenc

1

Query ①

# 通过筛选选项精炼序列结果

找与Query序列长度相同且完全匹配的序列：E值不限，其它各参数都设为100 to 100

BLAST Search Details

Sequence Type: Nucleotide  
Search Within: Nucleotides  
BLAST Algorithm: BLASTn  
NCBI Included: Yes  
Alignment Identity: 50%  
Query Coverage: 50%  
E-Value: 10  
Match with Gaps?: No  
Gap Costs: Existence 5  
Extension 2  
Word Size: 11

Bioscape Analysis

Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)

Create Bioscape Analysis

Filter by

E-Value

0 to 10<sup>6</sup>

Query Coverage %

100 to 100

Subject Coverage %

100 to 100

Alignment Identity %

100 to 100

## Biosequences (1)

View: Expanded

References

Query Details ATGCAGATCCCACAGGCCCTGGCCAGTCGTCTGGGCGGTGCTACAACCTGGGCTGGCGCCAGGATGGTTCTTAGACTCCCCA... [View More](#)

1 Alignment Identity: 100%

Query 1 864

Subject 1 864

Matches: 864  
Mismatches: 0

View Less

Alignment Subject References

Alignment Data

BLAST Score: 1728  
E-Value: 0

|   |     |  |     |
|---|-----|--|-----|
| Q | 1   | ATGCAGATCC CACAGGCCCT GGCCAGTC GTCTGGGCGG TGCTACAACCT GGGCTGGCGG CCAGGATGGT  | 70  |
| S | 1   | ATGCAGATCC CACAGGCCCT GGCCAGTC GTCTGGGCGG TGCTACAACCT GGGCTGGCGG CCAGGATGGT  | 70  |
| Q | 71  | TCTTAGACTC CCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCAGACC CTGCTCGTGG TGACCGAAGG | 140 |
| S | 71  | TCTTAGACTC CCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCAGACC CTGCTCGTGG TGACCGAAGG | 140 |
| Q | 141 | GGACAACGCC ACCTTCACCT GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG | 210 |
| S | 141 | GGACAACGCC ACCTTCACCT GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG | 210 |

# 通过筛选选项精炼序列结果

对Subject序列长度无要求，但完全匹配的序列：将Query Coverage和Alignment Identity设为100 to 100

BLAST Search Details

Sequence Type: Nucleotide  
Search Within: Nucleotides  
BLAST Algorithm: BLASTn  
NCBI Included: Yes  
Alignment Identity: 50%  
Query Coverage: 50%  
E-Value: 10  
Match with Gaps?: No  
Gap Costs: Existence 5  
Extension 2  
Word Size: 11

Bioscape Analysis

Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)

Create Bioscape Analysis

Filter by

E-Value

0 to 10<sup>6</sup>

Query Coverage %

100 to 100

Subject Coverage %

0 to 100

Alignment Identity %

100 to 100

Organisms

Homo sapiens (2)

Biosequences (16)

Sort: Alignment Identity View: Expanded

References

Query Details ATGCAGATCCCACAGGCGCCCTGGCCAGTCGTCTGGGCGGTGCTACAACCTGGGCTGGCGCCAGGATGGTTCTTAGACTCCCCA... View More

1 Alignment Identity: 100%

Query 1 864

Subject 1 1,747

Matches: 864  
Mismatches: 0

View Less

Alignment Subject References

Alignment Data

BLAST Score: 1728  
E-Value: 0

```
Q 1 ATGCAGATCC CACAGGCGCC CTGGCCAGTC GTCTGGGCGG TGCTACAACCT GGGCTGGCGG CCAGGATGGT 70
S 564 ATGCAGATCC CACAGGCGCC CTGGCCAGTC GTCTGGGCGG TGCTACAACCT GGGCTGGCGG CCAGGATGGT 633
Q 71 TCTTAGACTC CCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCCAGCC CTGCTCGTGG TGACCGAAGG 140
S 634 TCTTAGACTC CCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCCAGCC CTGCTCGTGG TGACCGAAGG 703
Q 141 GGACAACGCC ACCTTCACCT GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG 210
S 704 GGACAACGCC ACCTTCACCT GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG 773
```

2 Alignment Identity: 100%

Query 1 864

Subject 1 888

Matches: 864  
Mismatches: 0

View Less

# 通过筛选选项精炼序列结果

BLAST Search Details

Sequence Type: Nucleotide  
Search Within: Nucleotides  
BLAST Algorithm: BLASTn  
NCBI Included: Yes  
Alignment Identity: 50%  
Query Coverage: 50%  
E-Value: 10  
Match with Gaps?: No  
Gap Costs: Existence 5  
Extension 2  
Word Size: 11

Bioscape Analysis

Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)

Create Bioscape Analysis

Filter by

^ E-Value  
0 to  $10^6$

^ Query Coverage %  
85 to 100

^ Subject Coverage %  
0 to 100

^ Alignment Identity %  
90 to 99

^ Organisms

Biosequences (41) Sort: Alignment Identity View: Expanded

References

Query Details ATGCAGATCCCACAGGCGCCCTGGCCAGTCGTCTGGGCGGTGCTACAACCTGGGCTGGCGCCAGGATGGTTCTTAGACTCCCCA... View More

1 Alignment Identity: 98.84%

Query 1 864

Subject 1 2,187

Matches: 854  
Mismatches: 10

View Less

Alignment Subject

Alignment Data  
BLAST Score: 1678  
E-Value: 0

```
Q 1 ATGCAGATCC CACAGGCGCC CTGGCCAGTC GTCTGGGCGG TGCTACAACCT GGGCTGGCGG CCAGGATGGT 70
  |||
S 150 ATGCAGATCC CACAGGCGCC CTGGCCAGTC GTCTGGGCGG TGCTACAACCT GGGCTGGCGG CCAGGATGGT 219
Q 71 TCTTAGACTC CCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCCAGCC CTGCTCGTGG TGACCGAAGG 140
  |||
S 220 TCTTAGATT CCCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCCAGCC CTGCTCGTGG TGACCGAAGG 289
Q 141 GGACAACGCC ACCTTCACCT GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG 210
  |||
S 290 GGACAACGCC ACCTTCACCT GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG 359
```

2 Alignment Identity: 98.84%

Query 1 864

Subject 1 2,103

Matches: 854  
Mismatches: 10

对Query序列覆盖度高

有一定程度错配,但错配不多的

# 通过筛选选项精炼序列结果

The screenshot displays a BLAST search interface with the following components:

- BLAST Search Details:** Sequence Type: Nucleotide, Search Within: Nucleotides, BLAST Algorithm: BLASTn, NCBI Included: Yes, Alignment Identity: 50%, Query Coverage: 50%, E-Value: 10, Match with Gaps?: No, Gap Costs: Existence 5 Extension 2, Word Size: 11.
- Bioscape Analysis:** A section for visualizing sequence similarity, with a "Create Bioscape Analysis" button.
- Filter by:** E-Value (0 to 10<sup>6</sup>), Query Coverage % (85 to 100), **Subject Coverage % (0 to 25, highlighted in blue)**, Alignment Identity % (0 to 99).
- Biosequences (6):** A list of results, with the first one selected. It shows "Alignment Identity: 95.72%", "Matches: 760", and "Mismatches: 34".
- Alignment Data:** A detailed view of the alignment between the query (864 bp) and subject (5,306 bp) sequences. The BLAST Score is 1418 and the E-Value is 0. The alignment shows matches (vertical bars) and mismatches (vertical bars with red highlights).

Query序列对Subject序列占比小，即找相对长的序列，了解其上下游序列情况

比如：  
想了解某个输入的蛋白质编码序列上游的启动子序列



# 通过物种来筛选包含来自NCBI的序列结果

Filter by

^ E-Value

0 to  $10^6$

^ Query Coverage %

0 to 100

^ Subject Coverage %

0 to 100

^ Alignment Identity %

0 to 99

^ Organisms

- Homo sapiens (9)
- Chlorocebus sabaesus (6)
- Pan paniscus (6)
- Rousettus aegyptiacus (6)
- Bos taurus (5)

[View All](#)

Organisms ×

By Count  Alphanumeric [按相关结果数量排序](#)

1 Selected [筛选出特定物种的来自NCBI的序列](#)

|  |  |  |
|--|--|--|
| <input checked="" type="checkbox"/> Homo sapiens (9) | <input type="checkbox"/> synthetic construct (3)                 | <input type="checkbox"/> Molossus molossus (2)         |
| <input type="checkbox"/> Chlorocebus sabaesus (6)    | <input type="checkbox"/> Ursus maritimus (3)                     | <input type="checkbox"/> Mustela erminea (2)           |
| <input type="checkbox"/> Pan paniscus (6)            | <input type="checkbox"/> Ailuropoda melanoleuca (2)              | <input type="checkbox"/> Mustela putorius furo (2)     |
| <input type="checkbox"/> Rousettus aegyptiacus (6)   | <input type="checkbox"/> Balaenoptera acutorostrata scammoni (2) | <input type="checkbox"/> Orcinus orca (2)              |
| <input type="checkbox"/> Bos taurus (5)              | <input type="checkbox"/> Bubalus bubalis (2)                     | <input type="checkbox"/> Pan troglodytes (2)           |
| <input type="checkbox"/> Equus przewalskii (5)       | <input type="checkbox"/> Callithrix jacchus (2)                  | <input type="checkbox"/> Papio anubis (2)              |
| <input type="checkbox"/> Macaca nemestrina (5)       | <input type="checkbox"/> Camelus ferus (2)                       | <input type="checkbox"/> Phocoena sinus (2)            |
| <input type="checkbox"/> Phyllostomus discolor (5)   | <input type="checkbox"/> Cricetulus griseus (2)                  | <input type="checkbox"/> Ptilocolobus tephrosceles (2) |
| <input type="checkbox"/> Rattus norvegicus (5)       | <input type="checkbox"/> Felis catus (2)                         | <input type="checkbox"/> Pteropus vampyrus (2)         |
| <input type="checkbox"/> Canis lupus dingo (4)       | <input type="checkbox"/> Globicephala melas (2)                  | <input type="checkbox"/> Rhinolophus ferrumequinum (2) |
| <input type="checkbox"/> Cercocebus atys (4)         | <input type="checkbox"/> Halichoerus grypus (2)                  | <input type="checkbox"/> Rhinopithecus bieti (2)       |
| <input type="checkbox"/> Equus asinus (4)            | <input type="checkbox"/> Hipposideros armiger (2)                | <input type="checkbox"/> Rhinopithecus roxellana (2)   |
| <input type="checkbox"/> Equus caballus (4)          | <input type="checkbox"/> Lagenorhynchus obliquidens (2)          | <input type="checkbox"/> Suricata suricatta (2)        |
| <input type="checkbox"/> Monodon monoceros (4)       | <input type="checkbox"/> Lontra canadensis (2)                   | <input type="checkbox"/> Trachypithecus francoisi (2)  |
| <input type="checkbox"/> Canis lupus familiaris (3)  | <input type="checkbox"/> Macaca fascicularis (2)                 | <input type="checkbox"/> Tursiops truncatus (2)        |
| <input type="checkbox"/> Macaca mulatta (3)          | <input type="checkbox"/> Microcebus murinus (2)                  | <input type="checkbox"/> Ursus arctos horribilis (2)   |
| <input type="checkbox"/> Sus scrofa (3)              |  | <input type="checkbox"/> Vicugna pacos (2)             |

Organisms [按字母顺序排序](#)

By Count  Alphanumeric

1 Selected

- Ailuropoda melanoleuca (2)
- Balaenoptera acutorostrata scammoni (2)
- Bos taurus (5)
- Bubalus bubalis (2)
- Callithrix jacchus (2)

# 获取目标序列的更多信息

The image displays a BLAST search result interface. On the left, the 'Alignment' tab shows a 100% identity match between a query sequence (length 864) and a subject sequence (length 867). The alignment data shows a perfect match (BLAST Score: 1728, E-Value: 0). Below the alignment, the 'Subject' tab is selected, displaying '目标序列信息' (Target Sequence Information). This section includes CAS Registry Numbers (2306097-89-0, 503752-44-1), NCBI Identifier (AY238517, KJ865859.1), Length (867 nt), and Organisms (Homo sapiens). The full sequence is listed below. The 'References' tab is also selected, displaying '目标序列相关文献' (Target Sequence Related Literature), which lists three references related to PD-1 inhibitors and cancer treatment.

Alignment Identity: 100%

Query 1 864

Subject 1 867

Matches: 864  
Mismatches: 0

View Less ▾

Alignment Subject References

Alignment Data  
BLAST Score: 1728  
E-Value: 0

序列比对详情

Q 1 ATGCAGATCC CACAGGCGCC CTGGCCAGTC GTCTGGGCGG TGCTACAAC TGGGCTGGCGG CCAGGATGGT 70  
S 1 ATGCAGATCC CACAGGCGCC CTGGCCAGTC GTCTGGGCGG TGCTACAAC TGGGCTGGCGG CCAGGATGGT 70

Q 71 TCTTAGACTC CCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCCAGCC CTGCTCGTGG TGACCGAAGG 140  
S 71 TCTTAGACTC CCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCCAGCC CTGCTCGTGG TGACCGAAGG 140

Q 141 GGACAACGCC ACCTTCACCT GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG 210  
S 141 GGACAACGCC ACCTTCACCT GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG 210

Alignment Subject References

目标序列信息

CAS Registry Numbers: 2306097-89-0, 503752-44-1  
NCBI Identifier: AY238517, KJ865859.1  
Length: 867 nt  
Organisms: Homo sapiens

Sequence

1 ATGCAGATCC CACAGGCGCC CTGGCCAGTC GTCTGGGCGG TGCTACAAC TGGGCTGGCGG CCAGGATGGT TCTTAGACTC  
81 CCCAGACAGG CCCTGGAACC CCCCCACCTT CTCCCCAGCC CTGCTCGTGG TGACCGAAGG GGACAACGCC ACCTTCACCT  
161 GCAGCTTCTC CAACACATCG GAGAGCTTCG TGCTAAACTG GTACCGCATG AGCCCCAGCA ACCAGACGGA CAAGCTGGCC  
241 GCCTTCCCCG AGGACCGCAG CCAGCCCGGC CAGGACTGCC GCTTCCGTGT CACACAAC TCCCAACGGGC GTGACTTCCA  
321 CATGAGCGTG GTCAGGGCCC GGC GCAATGA CAGGGCACC TACCTCTGTG GGGCCATCTC CCTGGCCCCC AAGGCGCAGA  
401 TCAAAGAGAG CCTGCGGGCA GAGCTCAGGG TGACAGAGAG AAGGGCAGAA GTGCCACAG CCCACCCAG CCCCCTACCC  
481 AGGCCAGCCG GCCAGTTCCA AACCTG6TG GTTGGTGTG TGGGCGGCT GCTGGGCGC CTGGTGTCTG TAGTCTGGT  
561 CCTGGCCGTC ATCTGCTCCC GGGCCGACG AGGGACAATA GGAGCCAGGC GCACCGGCCA GCCCTGAAG GAGGACCCCT  
641 CAGCCGTGCC TGTGTTCTCT GTGGACTATG GGGAGTGG TTTCCAGTGG CGAGAGAAGA CCCCAGAGCC CCCCCTGCC  
721 TGTGTCCTG AGCAGACGGA GTATGCCACC ATTGTCTTCT CTAGCGGAAT GGGCACCTCA TCCCCGCC GCAGGGGCTC

Alignment Subject References

目标序列相关文献

使用PD-1轴拮抗剂和HPK1拮抗剂用于治疗癌症的方法和组合物  
Assignee: 豪夫迈·罗氏有限公司  
CN107206088 A | Seq ID No: 1

Compositions and Methods for Identification, Assessment, Prevention, and Treatment of Cancer Using PD-L1 Isoforms  
Assignee: DANA-FARBER CANCER INSTITUTE, INC.  
US20160122829 A1 | Seq ID No: 1

Novel pd1 isoforms, and uses thereof for potentiating immune responses  
Assignee: VERSITECH LIMITED  
US20140302070 A1 | Seq ID No: 24

# 获取目标序列的相关文献

与序列关联的专利、  
非专利文献

Filter Behavior

**Filter by** Exclude


Document Type

- Journal (832)
- Patent (24)
- Review (55)
- Clinical Trial (36)
- Editorial (1)
- Letter (6)
- Preprint (1)
- Report (2)

[View Fewer](#)

Language

Publication Year



No Min to No Max [Apply](#)

[View Larger](#)

Available at My Institution

Author

Organization

Publication Name

Concept

CA Section

## References (857)

Sort: Publication Date: Newest View: Partial Abstract

Substances  Reactions  Citing

1

**Genomic and transcriptomic correlates of immunotherapy response within the tumor microenvironment of leptomeningeal metastases**

By: Prakadan, Sanjay M.; Alvarez-Breckenridge, Christopher A.; Markson, Samuel C.; Kim, Albert E.; Klein, Robert H.; Nayyar, Naema; Navia, Andrew W.; Kuter, Benjamin M.; Kolb, Kellie E.; Bihun, Ivanna; et al  
Nature Communications (2021), 12(1), 5955 | Language: English, Database: C+plus and MEDLINE

Leptomeningeal disease (LMD) is a devastating complication of solid tumor malignancies, with dire prognosis and no effective systemic treatment options. Over the past decade, the incidence of LMD has steadily increased due to therapeutics that have extended the survival of cancer patients, highlighting the need for new interventions. To examine the efficacy of immune checkpoint inhibitors (ICI) in patients with LMD, we completed two phase II clin. trials. Here, we investigate the cellular and mol. features underpinning observed patient trajectories in these trials by applying single-cell RNA a...

[View More](#)

[Full Text](#)  Substances (3)  Reactions (0)  Citing (0)  Citation Map

2

**Programmed cell death protein-1 (PD-1) protects liver damage by suppressing IFN- $\gamma$  expression in T cells in infants and neonatal mice**

By: Guo, Xiangjie; Xu, Yiping; Luo, Wei; Fang, Rongli; Cai, Li; Wang, Ping; Zhang, Yuxia; Wen, Zhe; Xu, Yanhui  
BMC Pediatrics (2021), 21(1), 317 | Language: English, Database: C+plus and MEDLINE

Biliary atresia (BA) is a severe cholangiopathy possibly resulting from virus-induced and immune-mediated injury of the biliary system. IFN- $\gamma$ , secreted from CD4+ Th1 cells and CD8+ cytotoxic T cells, is a major mediator of liver pathol. Programmed death protein-1 (PD-1) signaling suppresses T cell function. However, how PD-1 modify T cell function in BA remains incompletely understood. Frequencies of PD-1 expressing CD4+ and CD8+ T cells were analyzed in the liver and blood from BA and control subjects. Associations of PD-1+CD4+/CD8+T cell abundances with liver function indexes were measured. ...

[View More](#)

[Full Text](#)  Substances (7)  Reactions (0)  Citing (1)  Citation Map

3

# 直接链接NCBI，获取序列相关信息

Alignment Subject References

CAS Registry Numbers: 2306097-89-0, 503752-44-1  
NCBI Identifier: [KJ865859.1](#), [AY238517.1](#)  
Length: 867 nt

Sequence

```
1 ATGCAGATCC CACAGGCGCC CTGGCCAGTC GTCTGGGCGG TGCTACAAC TGGGCTGGCGG CCAGGATGGT TCTTAGACTC
81 CCCAGACAGG CCCTGGAACC CCCCACCTT CTCCCAGCC CTGCTCGTGG TGACCGAAGG GGACAACGCC ACCTTCACCT
```

2个NCBI编号分别来自2个不同的文献来源

[ncbi.nlm.nih.gov/nuccore/KJ865859.1](#)

GenBank

### Homo sapiens cell-line A375 programmed cell death 1 protein (PDCD1) mRNA, complete cds

GenBank: KJ865859.1  
[FASTA](#) [Graphics](#) [PopSet](#)

Go to:

LOCUS KJ865859 867 bp mRNA linear PRI 22-SEP-2015  
DEFINITION Homo sapiens cell-line A375 programmed cell death 1 protein (PDCD1) mRNA, complete cds.  
ACCESSION KJ865859  
VERSION KJ865859.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 867)  
AUTHORS Kleffel,S., Posch,C., Barthel,S.R., Mueller,H., Schlapbach,C., Guenova,E., Elco,C.P., Lee,N., Juneja,V.R., Zhan,Q., Lian,C.G., Thomi,R., Hoetzenecker,W., Cozzio,A., Dummer,R., Mihm,M.C. Jr., Flaherty,K.T., Frank,M.H., Murphy,G.F., Sharpe,A.H., Kupper,T.S. and Schatton,T.  
TITLE Melanoma Cell-Intrinsic PD-1 Receptor Functions Promote Tumor Growth  
JOURNAL Cell 162 (6), 1242-1256 (2015)

[ncbi.nlm.nih.gov/nuccore/AY238517.1](#)

GenBank

### Homo sapiens programmed cell death 1 (PDCD1) mRNA, complete cds

GenBank: AY238517.1  
[FASTA](#) [Graphics](#)

Go to:

LOCUS AY238517 867 bp mRNA linear PRI 19-MAR-2003  
DEFINITION Homo sapiens programmed cell death 1 (PDCD1) mRNA, complete cds.  
ACCESSION AY238517  
VERSION AY238517.1  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 867)  
AUTHORS He,X., Xu,L., Liu,Y. and Zeng,Y.  
TITLE Cloning of PD-1 cDNA from activated peripheral leukocytes  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 867)

# 序列信息下载，便于进一步整理、分析、保存

**Biosequences** (919) Sort: Alignment Identity View: Expanded

References ↓

Query Details ATGCAGATCCCACAGGCGCCCTGGCCAGTCGCTCTGGGCGGTGCTACAAC... View More

1 Alignment Identity: 100%

Query ① ⑧64

Subject ① ⑧67

Matches: 864  
Mismatches: 0

**Download Biosequence Results** ×

File Type

Excel (.xlsx)

Select Quantity

All Results

Range (ex. 2 to 20)

to

| Alignment Image                                  | Alignment Text  | Sequence Length | CAS Registry Number          |
|--|---|-----------------|------------------------------|
| Query ① <span style="float: right;">⑧64</span>   | Q: 1 ATGCAGATCCCACAGGCGCCCTGGCCAGTCGCTCTGGGCGGTGCTACAAC 50<br> <br>S: 1 ATGCAGATCCCACAGGCGCCCTGGCCAGTCGCTCTGGGCGGTGCTACAAC 50     | 867             | 2306097-89-0,<br>503752-44-1 |
| Subject ① <span style="float: right;">⑧67</span> | Q: 51 GGGCTGGCGGCCAGGATGGTTCTTAGACTCCCCAGACAGGCCCTGGAACC 100<br> <br>S: 51 GGGCTGGCGGCCAGGATGGTTCTTAGACTCCCCAGACAGGCCCTGGAACC 100 |                 |                              |

| Accession Number | PubMed ID | NCBI Identifier | Number of Patents | Patent No.    | Sequence ID | E-Value | BLAST Score | Matches | Mismatches | Alignment Identity |
|------------------|-----------|-----------------|-------------------|---------------|-------------|---------|-------------|---------|------------|--------------------|
| 2011:368996      | 21237493  | AY238517        | 113               | US20170296676 | 3           | 0.0     | 1728.0      | 864     | 0          | 100.0              |
| 2011:609092      | 21346771  |                 |                   | US86042       | 19          |         |             |         |            |                    |
| 2013:320307      | 23291409  |                 |                   | US20160158360 | 1           |         |             |         |            |                    |
| 2017:339010      | 28238417  |                 |                   | US20160340407 | 1           |         |             |         |            |                    |
| 2020:78256       | 31838881  |                 |                   | EP3004877     | 1           |         |             |         |            |                    |
| 2008:643337      | 18440040  |                 |                   | WO2017079703  | 47          |         |             |         |            |                    |
| 2015:548093      | 25707450  |                 |                   | WO2017059168  | 19          |         |             |         |            |                    |
| 2008:55676       | 17942371  |                 |                   | CN105431449   |             |         |             |         |            |                    |
| 2014:1763855     | 24928021  |                 |                   | EP3094736     | 1           |         |             |         |            |                    |





# Bioscape : 快捷获取披露感兴趣序列的专利

Filter Behavior

Filter by Exclude

Document Type

Patent (121)

Language

Publication Year

No Min to No Max Apply

View Larger

Available at My Institution

Author

Organization

Publication Name

Concept

CA Section

CAS Solutions

## References (121)

Sort: Relevance View: Partial Abstract

Substances Reactions Citing

Download Email Save

1

### Human CD33-binding domain-containing chimeric antigen receptor or CAR-expressing T cells for cancer therapy

By: Brogdon, Jennifer; Ebersbach, Hilmar; Gill, Saar; Glass, David; Huber, Thomas; Jascur, Julia; Kenderian, Saad; Mannick, Joan; Milone, Michael C.; Murphy, Leon; et al

World Intellectual Property Organization, WO2016014576 A1 2016-01-28 | Language: English, Database: CAplus

The disclosure provides compositions and methods for treating diseases associated with expression of CD33. The disclosure also relates to chimeric antigen receptor (CAR) specific to CD33, vectors encoding the same, and recombinant T cells comprising the CD33 CAR. The disclosure also includes methods of administering a genetically modified T cell expressing a CAR that comprises a CD33 binding domain.

PatentPak Full Text Substances (162) Reactions (0) Citing (5) Citation Map

2

### B-cell inhibitors for combination therapy with CD19 chimeric antigen receptor-expressing T-cells

By: Bitter, Hans; Bordeaux, Jennifer Mary; Brannetti, Barbara; Brogdon, Jennifer; Dakappagari, Naveen Kumar; Gill, Saar; Highfill, Steven; Huang, Lu; June, Carl H.; Kim, Ju Young; et al

World Intellectual Property Organization, WO2016164731 A2 2016-10-13 | Language: English, Database: CAplus

The authors disclose a method for treating diseases associated with expression of CD19. The method comprises the administration of a recombinant T-cell expressing a chimeric antigen receptor (CAR) targeting CD19 in combination with one or more B-cell inhibitors. In one example, chronic lymphocytic leukemia (CLL) patients that received CD19 CAR T-cell therapy were examined for the expression of immune checkpoint inhibitor mols. The results indicated that non-responding patients exhibited increased expression of PD-1, LAG3, and TIM3. In a second example, CD123 was expressed in CD19-neg. B-ALL re...



# 通过Concept快速了解序列相关文献主题

Concept ✕

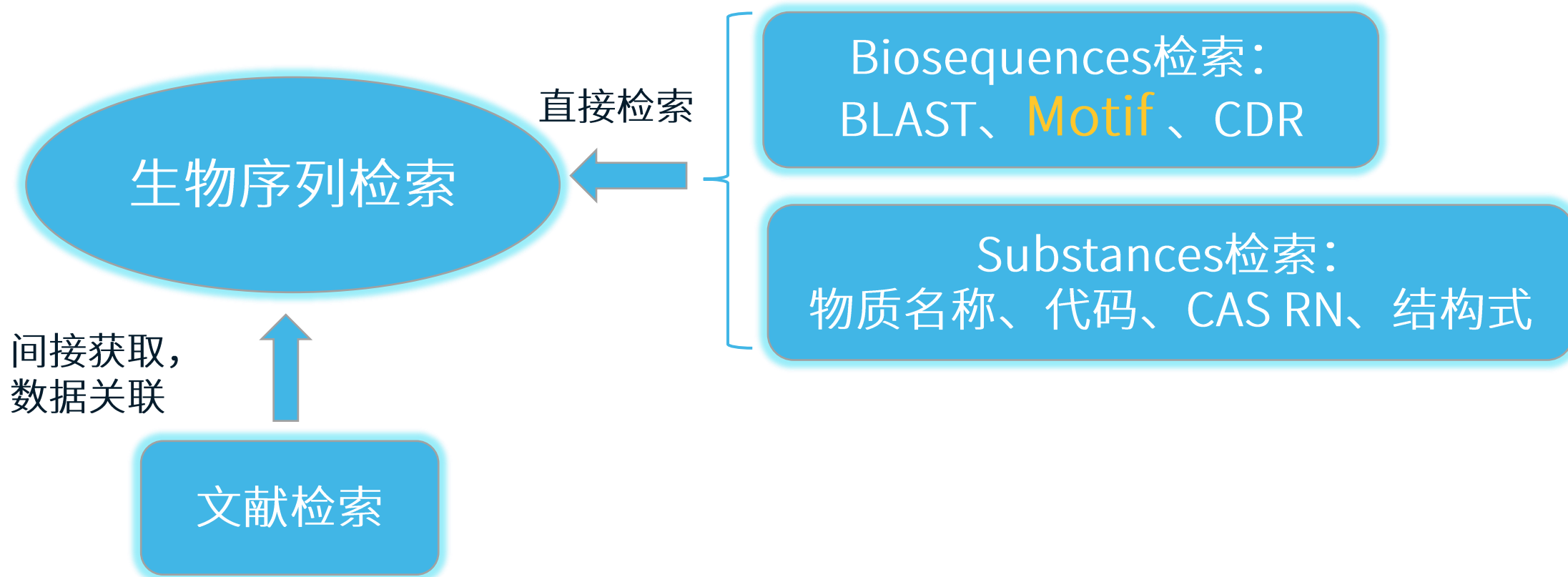
Top Count | Alphanumeric | Search

0 Selected

|  |   |  |
|--|---|--|
| <input type="checkbox"/> Homo sapiens (88)                         | <input type="checkbox"/> Protein motifs (35)  | <input type="checkbox"/> B cell-activating factor receptors (27)     |
| <input type="checkbox"/> Human (88)                                | <input type="checkbox"/> Leukemia (34)  | <input type="checkbox"/> Cancer immunotherapy (27)                   |
| <input type="checkbox"/> Chimeric antigen receptors (77)           | <input type="checkbox"/> T cell receptors (34)  | <input type="checkbox"/> CAR T-cell immunotherapy (27)               |
| <input type="checkbox"/> Protein sequences (68)                    | <input type="checkbox"/> Tumor necrosis factor receptor superfamily protein TNFRSF17 (34) | <input type="checkbox"/> CD226 antigens (27)                         |
| <input type="checkbox"/> CD19 antigens (67)                        | <input type="checkbox"/> CD160 antigens (33)  | <input type="checkbox"/> CD4-positive T cell (27)                    |
| <input type="checkbox"/> Antitumor agents (56)                     | <input type="checkbox"/> CD80 antigens (33)   | <input type="checkbox"/> CD7 antigens (27)                           |
| <input type="checkbox"/> Neoplasm (52)                             | <input type="checkbox"/> CD86 antigens (33)   | <input type="checkbox"/> CD9 antigens (27)                           |
| <input type="checkbox"/> CD28 antigens (50)                        | <input type="checkbox"/> Mesothelins (33)   | <input type="checkbox"/> Chronic myeloid leukemia (27)               |
| <input type="checkbox"/> Natural killer cell (50)                  | <input type="checkbox"/> CD20 antigens (32)   | <input type="checkbox"/> FcγRI receptors (27)                        |
| <input type="checkbox"/> T cell (48)                               | <input type="checkbox"/> Hodgkin disease (32)   | <input type="checkbox"/> Interleukin 3 receptors (27)                |
| <input type="checkbox"/> Hematopoietic neoplasm (47)               | <input type="checkbox"/> Intercellular adhesion molecule 1 (32)                           | <input type="checkbox"/> Killer cell lectin-like receptor KLRK1 (27) |
| <input type="checkbox"/> Sialic acid-binding Ig-like lectin 2 (47) | <input type="checkbox"/> Mantle cell lymphoma (32)  | <input type="checkbox"/> Pancreatic neoplasm (27)                    |
| <input type="checkbox"/> CD27 antigens (46)                        | <input type="checkbox"/> Proteins (32)  | <input type="checkbox"/> Waldenstrom                                 |

Apply Cancel

# 生物序列获取方法



# Motif检索

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

## Biosequences

Enter a protein or nucleotide string. [Learn more about Biosequence Search.](#)

BLAST CDR **Motif** Clear Search

[SG]X{4}GK[DT]

X代表未指定氨基酸，可以是常见氨基酸、不常见氨基酸

Sequence Type:  
 Nucleotide  Protein

Include NCBI Sequences

Limit Total Sequence Results to:  
1000

Start Biosequence Search

Advanced Biosequence Search ^ Reset All


Query Coverage %  E-Value   Combine Motif Results

Motif中有可变部分，可借助符号来表示：

“[]” 中括号：代表或者，表示出现在该位置的氨基酸可以是被括起来的氨基酸中的任意一个

“{}” 大括号：代表氨基酸的重复次数。其中字段可用逗号开，{2,6}表示在大括号左边紧密相连的氨基酸可重复2-6次

# 浏览用于检索的序列

| Recent Search History  |  |   | <a href="#">View All Search History</a>                              |
|--|--|---|--|
| March 15, 2022   |  |   |  |
|  <b>Biosequences</b><br>9:15 PM | Sequence Type: Protein<br>NCBI Included: No<br>Query Coverage: 90%<br>E-value: 10<br><br>Results will expire on<br>Apr 15, 2022. | [SG]X{4}GK[DT]<br>4 Biosequences Submitted<br><a href="#">Hide Individual Sequences</a> ^ | <a href="#">View Results (4 of 4)</a><br><a href="#">Edit Search</a> |
| 1 of 4   | SXXXXGKD   | <i>Complete</i>   | <a href="#">View Results</a>   |
| 2 of 4   | SXXXXGKT   | <i>Complete</i>   | <a href="#">View Results</a>   |
| 3 of 4   | GXXXXGKD   | <i>Complete</i>   | <a href="#">View Results</a>   |
| 4 of 4   | GXXXXGKT   | <i>Complete</i>   | <a href="#">View Results</a>   |

# 获取Motif检索结果

**Motif Search Details**

Sequence Type: Protein  
NCBI Included: No  
Query Coverage: 90%  
E-Value: 10

**Bioscape Analysis**

Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)

[Create Bioscape Analysis](#)

**Filter by**

**E-Value**  
0 to 10<sup>6</sup>

**Query Coverage %**  
0 to 100

**Subject Coverage %**  
0 to 100

**Alignment Identity %**  
0 to 100

**Biosequences (1,000)** Sort: Alignment Identity View: Expanded

References

Query Details [View More](#)

> Seq 2: 1 SXXXXGKT 8

1 Alignment Identity: 87.5%

Query 1 8

Subject 1 3,556

Matches: 7  
Mismatches: 1

[View Less](#)

Alignment Subject References

Alignment Data  
BLAST Score: 77  
E-Value: 0.000257318

```
Q      1  SXXXXGKT  8
      +| | | | |
S     1953 XXXXXGKT 1960
```

2 Alignment Identity: 87.5%

Query 1 8

Query Details [View More](#)

> Seq 2: 1 SXXXXGKT 8

> Seq 1: 1 SXXXXGKD 8

> Seq 2: 1 SXXXXGKT 8

> Seq 3: 1 GXXXXGKD 8

> Seq 4: 1 GXXXXGKT 8

# Motif检索—Combine Motif Results

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

## Biosequences

Enter a protein or nucleotide string. [Learn more about Biosequence Search.](#)

BLAST CDR Motif Clear Search

[SG]X{4}GK[DT]

Sequence Type:  
Nucleotide **Protein**

Include NCBI Sequences

Limit Total Sequence Results to:  
1000

**Start Biosequence Search**

Advanced Biosequence Search Reset All

Query Coverage %  E-Value

**Combine Motif Results**

将通式代表的多条序列的检索结果整合在同一结果集

**Biosequences**  
9:15 PM

Sequence Type: Protein  
NCBI Included: No  
Query Coverage: 90%  
E-value: 10

Results will expire on  
Apr 15, 2022.

[SG]X{4}GK[DT]

**View Results**

Edit Search

Complete

# Motif检索结果—Combine Motif Results

Motif Search Details

Sequence Type: Protein  
NCBI Included: No  
Query Coverage: 90%  
E-Value: 10

Bioscape Analysis

Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)

[Create Bioscape Analysis](#)

Filter by

^ E-Value

0 to 10<sup>6</sup>

^ Query Coverage %

0 to 100

^ Subject Coverage %

0 to 100

^ Alignment Identity %

0 to 100

**Biosequences** (1,000) Sort: Alignment Identity View: Expanded

References

Query Details [SG]X{4}GK[DT] [View Less](#)

[SG]X{4}GK[DT]

1 Alignment Identity: 100%

Query 1 14

Subject 1 40

Matches: 8  
Mismatches: 0

[View Less](#)

Alignment Subject References

Alignment Data  
BLAST Score: 78  
E-Value: 0.0000957867

|   |   |          |   |
|---|---|----------|---|
| Q | 1 | GXXXXGKT | 8 |
|   |   |          |   |
| S | 1 | GXXXXGKT | 8 |

2

6 Alignment Identity: 87.5%

Query 1 14

Subject 1 118

Matches: 7  
Mismatches: 1

[View Less](#)

Alignment Subject References

Alignment Data  
BLAST Score: 78  
E-Value: 0.000404075

|   |    |          |    |
|---|----|----------|----|
| Q | 1  | GXXXXGKD | 8  |
|   |    | +        |    |
| S | 77 | GXXXXGXD | 84 |

查看整合了通式代表的多条序列检索结果的结果集



# 结果精炼&获取目标序列的更多信息

Filter by

^ E-Value

0 to  $10^6$

^ Query Coverage %

100 to 100

^ Subject Coverage %

0 to 100

^ Alignment Identity %

100 to 100

**Biosequences** (4) Sort: Alignment Identity View: Expanded

References [Download] [Email]

Query Details [SG]X{4}GK[DT] [View More](#)

2 Alignment Identity: 100%

Query 1

Subject 1 8

Alignment Subject References

CAS Registry Numbers: 852985-75-2, 2244730-61-6, 851346-82-2, 312910-16-0, 850390-40-8  
Length: 8 aa  
Sequence  
1 GXXXXGKT

View Less

Alignment Subject References

Alignment Data  
BLAST Score: 78  
E-Value: 0.0000957885

|   |   |          |   |
|---|---|----------|---|
| Q | 1 | GXXXXGKT | 8 |
|   |   |          |   |
| S | 1 | GXXXXGKT | 8 |

Alignment Subject References

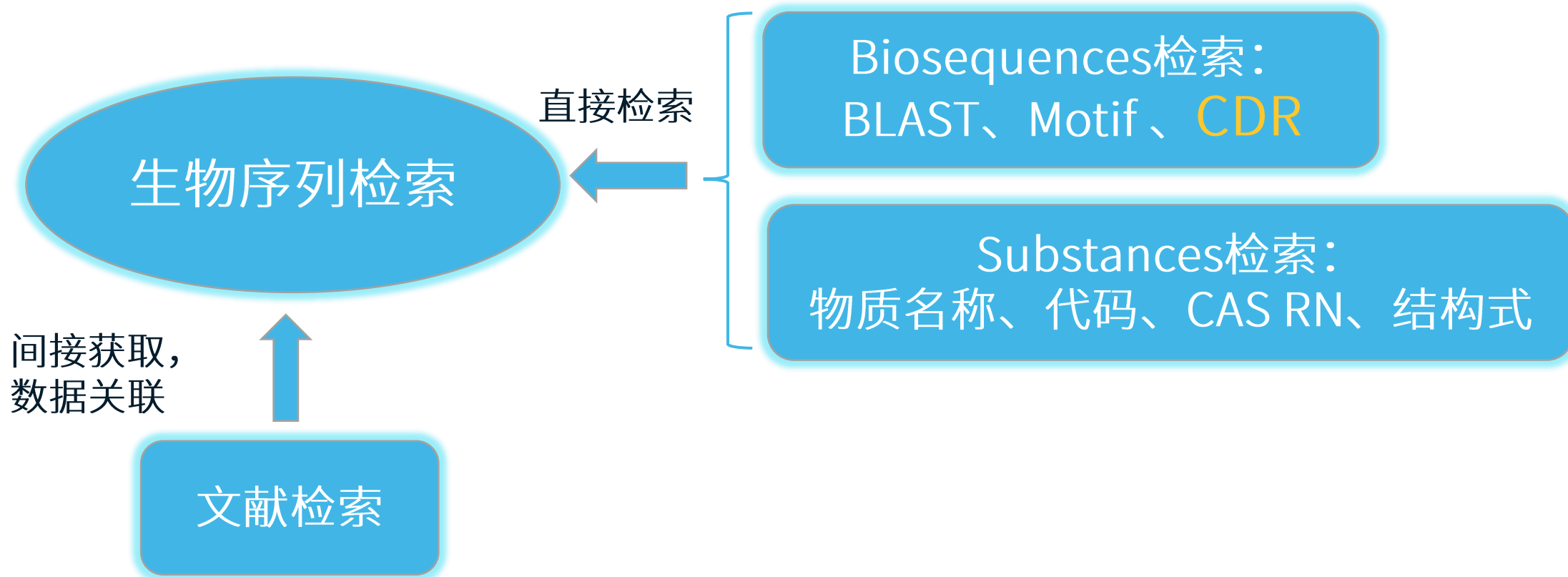
References

**Tumor suppressor genes, proteins encoded thereby and use of said genes and proteins**  
Assignee: YEDA RESEARCH & DEVELOPMENT COMPANY, LTD.  
EP0804613 A1 | Seq ID No: 8

**一种多肽——人RNA解螺旋酶44和编码这种多肽的多核苷酸**  
Assignee: 上海博德基因开发有限公司  
CN1477198 A | Seq ID No: 26

**Rapid detection of antibiotic resistance in mycobacterium tuberculosis**  
Assignees: INSTITUT PASTEUR; MEDICAL RESEARCH COUNCIL; ASSISTANCE PUBLIQUE; UNIVERSITE PIERRE ET MARIE CURIE (PARIS VI); UNIVERSITE DE BERNE  
US6124098 A | Seq ID No: 29

# 生物序列获取方法



# 运用CDR检索生物序列

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

## Biosequences

Enter a protein string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST **CDR** Motif Upload Sequence Clear Search

|      |                 |   |
|------|-----------------|---|
| CDR1 | RASQSVSGSRFTYMH | × |
| CDR2 | YASILES         | × |
| CDR3 | QHSWEIPPWT      | × |

Include NCBI Sequences

Limit Total Sequence Results to:

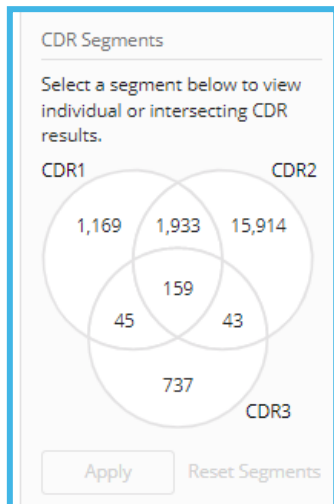
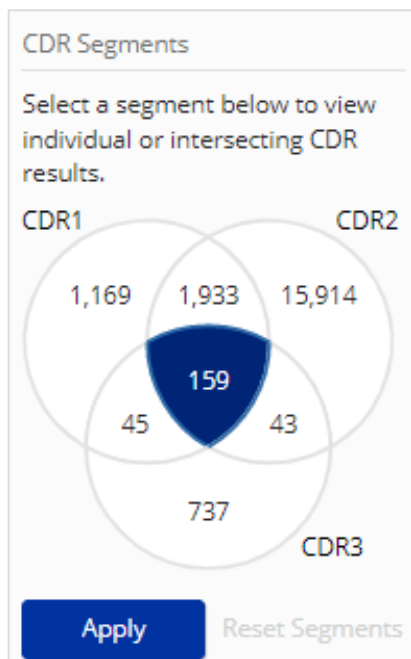
Start Biosequence Search

支持单个或多个CDR序列检索并用

## CDR Segments:

匹配到一个或者多个CDR的  
subject序列的序列数量。

查看3个CDR都被包含的159个序列结果



Bioscape Analysis

Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)

Create Bioscape Analysis

Filter by

^ E-Value  
0 to 10<sup>6</sup>

^ Query Coverage %  
0 to 100

^ Subject Coverage %  
0 to 100

^ Alignment Identity %  
0 to 100

### Biosequences (20,000)

Sort: Alignment Identity View: Expanded

References

Query Details View Less

> CDR1  
RASQSVSGSRFTYMH

> CDR2  
YASILES

> CDR3  
QHSWEIPPWT

1 Alignment Identity: 100%

Subject 1 468

CDR1 1 15  
CDR3 1 10  
CDR2 1 7

Matches: 32  
Mismatches: 0

View Less

| Alignment                  | Subject                  | References |
|----------------------------|--------------------------|------------|
| Alignment Data             |                          |            |
| BLAST Score: 84            |                          |            |
| E-Value: 10.51305880353855 |                          |            |
| CDR1                       | 1 RASQSVSGSR FTYMH 15    |            |
| S                          | 273 RASQSVSGSR FTYMH 287 |            |
| CDR2                       | 1 YASILES 7              |            |
| S                          | 303 YASILES 309          |            |
| CDR3                       | 1 QHSWEIPPWT 10          |            |
| S                          | 342 QHSWEIPPWT 351       |            |

# 获取序列检索结果集

**Biosequences (159)** Sort: Alignment Identity View: Collapsed

References

Query Details View Less

> CDR1  
RASQSVSGSRFTYMH

> CDR2  
YASILES

> CDR3  
QHSWEIIPPWT

Alignment Identity: 100%

Subject 1 112

Alignment Identity: 100%

Subject 1 219

Alignment Identity: 62.07%

Subject 1 111

Matches: 18 Mismatches: 11

CDR Segments

Select a segment below to view individual or intersecting CDR results.

CDR1 CDR2 CDR3

159

Apply Reset Segments

Bioscape Analysis

Visually explore sequence similarity with a new tool. [Learn more about Bioscape.](#)

Create Bioscape Analysis

Filter by

E-Value

0 to 10<sup>6</sup>

Query Coverage %

0 to 100

Subject Coverage %

0 to 100

Alignment Identity %

0 to 100

Query Coverage %

100 to 100

获取每个Query序列全长皆比对上的结果

**Biosequences (11)** Sort: Alignment Identity View: Collapsed

Alignment Identity: 100%

Subject 1 219

Matches: 32 Mismatches: 0

Alignment Identity: 75%

Subject 1 113

Matches: 24 Mismatches: 8

View Less

| Alignment             | Subject                | References |
|-----------------------|------------------------|------------|
| Alignment Data        |                        |            |
| BLAST Score: 55       |                        |            |
| E-Value: 1041.8646405 |                        |            |
| CDR1                  | 1 RASQSVSGSR FTYMH 15  | +          |
| S                     | 24 RASKSVSTSS YSYMH 38 |            |
| CDR2                  | 1 YASILES 7            |            |
| S                     | 54 YASYLE 60           |            |
| CDR3                  | 1 QHSWEIIPPWT 10       |            |
| S                     | 93 QHSREFPPWT 102      |            |

Query序列未全长比对上

# 获取目标序列的更多信息

5 Alignment Identity: 100%

Matches: 32  
Mismatches: 0

View Less ▾

Alignment Subject References

Alignment Data  
BLAST Score: 84  
E-Value: 10.51305880353855

|      |    |            |       |     |
|------|----|------------|-------|-----|
| CDR1 | 1  | RASQSVSGSR | FTYMH | 15  |
| S    | 24 | RASQSVSGSR | FTYMH | 38  |
| CDR2 | 1  | YASILES    |       | 7   |
| S    | 54 | YASILES    |       | 60  |
| CDR3 | 1  | QHSWEIPPWT |       | 10  |
| S    | 93 | QHSWEIPPWT |       | 102 |

Alignment Subject References

CAS Registry Numbers: 1884717-60-5, 2055950-84-8  
Length: 112 aa

Sequence

```
1 DIVMTQSPDS LAVSLGERAT INCRASQSVS GSRFTYMHVY QQKPGQPPKL LIKYASILES GVPDRFSGSG SGTDFTLTIS
81 SLQAEDVAVY YCQHSWEIPP WTFGQGTKVE IK
```

## Substance Detail

Reference (1) Reactions (0)

CAS Registry Number

1884717-60-5

Unspecified

Immunoglobulin G1, anti-(human tyrosine kinase receptor ErbB2 domain I) (human-Mus musculus clone hu7C2 κ-chain V-J region) (ACI)

Protein/Peptide Sequence

Sequence Length: 112

Related Sequences (2)

CAS科学家标引的信息

更多相关序列

### Other Names and Identifiers

1 Other Name for this Substance

2: PN: WO2016040723 SEQID: 10 claimed protein 来源&专利中的序列编号

### Bioactivity Indicators

Antitumor agents (1)

生物活性信息

### Target Indicators

Enzymes

Epidermal growth factor receptor HER2 (1)

靶点信息

Signaling proteins

Epidermal growth factor receptor HER2 (1)



ACS  
International



# 获取目标序列的更多信息

## Substance Detail

Reference (1) Reactions (0)

CAS Registry Number  
**2055950-84-8**

Unspecified

Immunoglobulin, humanized, anti-(human tyrosine kinase receptor ErbB2) (synthetic clone hu7C2 light chain VL fragment) (ACI)

Protein/Peptide Sequence  
Sequence Length: 112

Related Sequences (2) 更多相关序列

### Other Names and Identifiers

1 Other Name for this Substance

61: PN: WO2016205176 SEQID: 63 claimed protein 相同序列的不同来源

### Sequence Details

Sequence: linear

|     |            |            |            |            |            |
|-----|------------|------------|------------|------------|------------|
| 1   | DIVMTQSPDS | LAVSLGERAT | INCRASQSVS | GSRFTYMHWY | QQKPGQPPKL |
| 51  | LIKYASILES | GVPDRFSGSG | SGTDFTLTIS | SLQAEDVAVY | YCQHSWEIPP |
| 101 | WTFGQGTKVE | IK         | -          | -          | -          |

### Substances (2)

References Reactions Suppliers

|           |  |  |
|-----------|--|--|
| 1         | 2055950-84-8   | 1884717-60-5   |
|           | Image Not Available  | Image Not Available  |
|           | Unspecified<br>Immunoglobulin, humanized, anti-(human tyrosine kinase receptor ErbB2) (syntheti... | Unspecified<br>Immunoglobulin G1, anti-(human tyrosine kinase receptor ErbB2 domain I) (human-M... |
|           | Protein/Peptide Sequence<br>Sequence Length: 112   | Protein/Peptide Sequence<br>Sequence Length: 112   |
| 1         | 0  | 0  |
| Reference | Reactions  | Suppliers  |

### Bioactivity Indicators 生物活性信息

- Antiproliferative agents (1)
- Antitumor agents (1)
- Biopharmaceuticals (1)
- Peptide analogs
  - Peptidomimetics (1)

### Target Indicators 靶点信息

- Enzymes
  - Epidermal growth factor receptor HER2 (1)
  - Metalloreductase STEAP1 (1)
- Glycoproteins
  - Cancer antigen 125 (1)
- Proteinaceous antigens
  - B cell receptor-associated protein CD79b (1)
  - Cancer antigen 125 (1)
  - Lymphocyte antigen Ly6E (1)
  - Metalloreductase STEAP1 (1)
  - Tumor antigens (1)
- Signaling proteins
  - B cell receptor-associated protein CD79b (1)
  - Epidermal growth factor receptor HER2 (1)

# 获取披露目标序列的文献

5 Alignment Identity: 100%

Subject 1 112

View Less ▾

Alignment Subject **References**

**Antibodies and immunoconjugates**  
Assignee: GENENTECH, INC.  
WO2016205176 A1 | Seq ID No: 63

**Antibodies and immunoconjugates**  
Assignee: GENENTECH, INC.  
US20170002086 A1 | Seq ID No: 63

**Anti-her2 antibodies and immunoconjugates**  
Assignee: GENENTECH, INC.  
WO2016040723 A1 | Seq ID No: 10

Filter Behavior

**Filter by** Exclude

Document Type

Patent (16)

Language

Publication Year

| Year | Count |
|------|-------|
| 2012 | 1     |
| 2013 | 2     |
| 2014 | 2     |
| 2015 | 1     |
| 2016 | 2     |
| 2017 | 2     |
| 2018 | 1     |
| 2019 | 2     |
| 2020 | 2     |
| 2021 | 1     |
| 2022 | 1     |

No Min to No Max Apply

[View Larger](#)

Available at My Institution

Author

Organization

Publication Name

Concept

CA Section

CAS Solutions

**References** (16) Sort: Times Cited ▾ View: Partial Abstract ▾

Substances ▾  Reactions ▾  Citing ▾

1

**Phosphodiesterase PDE5 compositions and methods for immunotherapy**  
By: Suri, Vipin; Delabarre, Byron; Gladstone, Michael N.; Richardson, Celeste; Dolinski, Brian; Kulkarni, Abhishek; Inniss, Mara Christine; Sun, Dexue; Li, Dan Jun; Olinger, Grace Y.; et al  
World Intellectual Property Organization, WO2018231759 A1 2018-12-20 | Language: English, Database: CAPlus

The present invention relates to compositions and methods for the regulated and controlled expression of proteins. Methods for inducing anti-cancer immune responses in a subject are also provided.

2

**Anti-HER2 antibodies and immunoconjugates and therapeutic and diagnostic uses thereof**  
By: Chen, Xiaocheng; Dennis, Mark; Junutula, Jagath Reddy; Phillips, Gail Lewis; Pillow, Thomas Harden; Sliwkowski, Mark X.  
World Intellectual Property Organization, WO2016040723 A1 2016-03-17 | Language: English, Database: CAPlus

The disclosure provides anti-HER2 antibodies and immunoconjugates and methods of using the same for treating a HER2-pos. cancer, particularly a breast cancer or gastric cancer.



# PatentPak获取目标序列的相关专利信息

**Reference Detail** (2 of 16) ← Prev Next →

Substances (114) Reactions (203) Citing (4) Citation Map Download Email Save

**PATENT**

Patent Number  
WO2016040723

Publication Date  
2016-03-17

Application Number  
WO2015-US49549

Application Date  
2015-09-11

Kind Code  
A1

Assignee  
Genentech, Inc., United States

Source  
World Intellectual Property Organization  
CODEN: PIXXD2

Database Information  
AN: 2016:431027  
CAN: 164:392451  
CAplus

Language  
English

CAS Formulus®, the comprehensive formulations database and workflow solution, is now available for all SciFinder® users. [View content from CAS Formulus®](#) in this document. [Learn more about Formulus®.](#)

**Anti-HER2 antibodies and immunoconjugates and therapeutic and diagnostic uses thereof**

By: Chen, Xiaocheng; Dennis, Mark; Junutula, Jagath Reddy; Phillips, Gail Lewis; Pillow, Thomas Harden; Sliwkowski, Mark X.

The disclosure provides anti-HER2 antibodies and immunoconjugates and methods of using the same for treating a HER2-pos. cancer, particularly a breast cancer or gastric cancer.

**Keywords:** HER2 antibody immunoconjugate anticancer cancer diagnostic

**PatentPak Viewer** Get Prior Art Analysis Full Text ▾

**Patent Family**

| Patent       | Language | Kind Code | PatentPak Options   | Publication Date | Application Number | Application Date |
|--------------|----------|-----------|---|------------------|--------------------|------------------|
| WO2016040723 | English  | A1        | <a href="#">PDF</a>   <a href="#">PDF+</a>   <a href="#">Viewer</a> | 2016-03-17       | WO2015-US49549     | 2015-09-11       |

| Patent         | Language   | Kind Code | PatentPak Options   | Publication Date | Application Number  | Application Date |
|----------------|------------|-----------|---|------------------|---------------------|------------------|
| WO2016040723   | English    | A1        | <a href="#">PDF</a>   <a href="#">PDF+</a>   <a href="#">Viewer</a> | 2016-03-17       | WO2015-US49549      | 2015-09-11       |
|                |            |           |   |                  | US2014-6262049594   | 2014-09-12       |
| CA2957238      | English    | A1        | <a href="#">PDF</a>   | 2016-03-17       | CA2015-2957238      | 2015-09-11       |
| US20160096893  | English    | A1        | <a href="#">PDF</a>   | 2016-04-07       | US2015-1414851001   | 2015-09-11       |
| US9518118      | English    | B2        | <a href="#">PDF</a>   | 2016-12-13       | US2015-1414851001   | 2015-09-11       |
| AR101845       | Spanish    | A1        | <a href="#">PDF</a>   | 2017-01-18       | AR2015-102906       | 2015-09-11       |
| AU2015314954   | English    | A1        | <a href="#">PDF</a>   | 2017-03-02       | AU2015-314954       | 2015-09-11       |
| IL250440       | English    | A         |   | 2017-03-30       | IL2015-250440       | 2015-09-11       |
| KR2017057280   | Korean     | A         | <a href="#">PDF</a>   | 2017-05-24       | KR2017-7007335      | 2015-09-11       |
| CR2017000131   | Spanish    | A         |   | 2017-07-19       | CR2017-131          | 2015-09-11       |
| EP3191135      | English    | A1        |   | 2017-07-19       | EP2015-772090       | 2015-09-11       |
| CN107001479    | Chinese    | A         | <a href="#">PDF</a>   | 2017-08-01       | CN2015-80048008     | 2015-09-11       |
| JP2017534253   | Japanese   | T         | <a href="#">PDF</a>   | 2017-11-24       | JP2017-513637       | 2015-09-11       |
| BR112017004631 | Portuguese | A2        |   | 2018-01-30       | BR2017-112017004631 | 2015-09-11       |
| SG10201809668  | English    | A1        |   | 2018-11-29       | SG2018-10201809668  | 2015-09-11       |
| EP3191135      | English    | B1        | <a href="#">PDF</a>   | 2020-08-19       | EP2015-772090       | 2015-09-11       |
| TWI702231      | Chinese    | B         |   | 2020-08-21       | TW2015-130172       | 2015-09-11       |
| EP3782654      | English    | A1        | <a href="#">PDF</a>   | 2021-02-24       | EP2020-185412       | 2015-09-11       |
| AU2015314954   | English    | B2        |   | 2021-05-13       | AU2015-314954       | 2015-09-11       |
| E52830385      | Spanish    | T3        |   | 2021-06-03       | E52015-772090       | 2015-09-11       |

View All ▾

Expand All | Collapse All

- ✓ IPC Data 快捷浏览IPC分类信息，定位相关技术领域，精准查找现有技术
- ✓ Concepts 科学家标引的核心技术概念词
- ✓ Substances 科学家标引的物质信息
- ✓ Formulations 科学家标引的配方信息：成分/功能/含量/作用/适应症/全文中位置等
- ✓ Cited Documents

# PatentPak快捷浏览并高效获取物质信息

浏览获取感兴趣的序列信息

The screenshot displays the CAS PatentPak interface. At the top, there are navigation controls for 'PAGE' (178 / 204), 'ZOOM' (minus and plus icons), and 'DOWNLOAD' (PDF and PDF+). The main content area shows a list of patent claims, with three blue location pins marking specific points in the text. A blue arrow points from the sidebar to the first pin. The sidebar on the left is titled 'Key Substances in Patent' and lists three substances with their CAS RN numbers and descriptions. The first substance is highlighted in yellow, and a blue arrow points to its CAS RN number. The second substance is also highlighted in yellow, and a blue arrow points to its CAS RN number. The third substance is highlighted in yellow, and a blue arrow points to its CAS RN number.

**CAS** PatentPak

PAGE 178 / 204 ZOOM DOWNLOAD PDF PDF+

Key Substances in Patent

Analyst Markup Locations (1)  
Page 106

**CAS RN**  
**180288-69-1**  
Immunoglobulin G1, anti-  
(human p185<sup>neu</sup> receptor)  
(human-mouse monoclonal  
rhuMab HER2  $\gamma_1$ -chain),  
disulfide with human-mouse  
monoclonal rhuMab HER2 light  
chain, dimer

Analyst Markup Locations (2)  
Page 126  
Page 177

**CAS RN**  
**380610-27-5**  
Immunoglobulin G1, anti-  
(human neu (receptor))  
(human-mouse monoclonal 2C4  
heavy chain), disulfide with

44. The pharmaceutical formulation of claim 40, wherein the additional therapeutic agent is selected from trastuzumab, trastuzumab-MCC-DM1 (T-DM1), and pertuzumab.

45. The pharmaceutical formulation of claim 40, further comprising (1) trastuzumab or T-DM1, and (2) pertuzumab.

46. A method of treating an individual having a HER2-positive cancer, the method comprising administering to the individual an effective amount of the immunoconjugate of any one of claims 20 to 38 or the pharmaceutical formulation of any one of claims 39 to 45.

47. The method of claim 46, wherein the HER2-positive cancer is breast cancer or gastric cancer.

48. The method of claim 47, wherein the HER2-positive breast cancer is early-stage breast cancer.

49. The method of claim 47, wherein the HER2-positive breast cancer is metastatic breast cancer.

50. The method of any one of claims 46 to 49, further comprising administering an additional therapeutic agent to the individual.

CAS科学家标引了物质在专利中的位置，同时提供物质结构、CAS登记号等信息，有助于提高阅读、理解专利的效率。

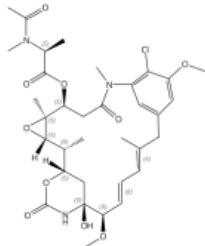
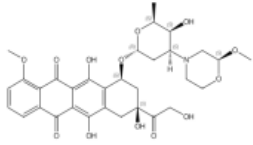
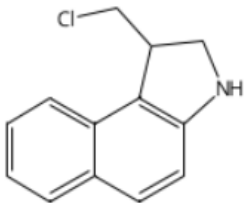
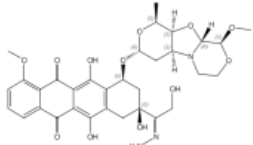
# PatentPak 高效获取专利中重要的物质信息

PatentPak Viewer   Get Prior Art Analysis   Full Text ▾

Patent Family

| Patent       | Language | Kind Code | PatentPak Options   |
|--------------|----------|-----------|---------------------|
| WO2018140778 | English  | A1        | PDF   PDF+   Viewer |

PDF+ 中CAS 科学家标引的物质表格

| Mark | Page # | CAS RN        | Name   | Structure   |
|------|--------|---------------|--|---|
| 502  | p.169  | 35846-53-8D   | Maytansine<br>derivs., antibody conjugates<br>与抗体偶联的小分子  |    |
| 503  | p.169  | 12794-10-4D   | Benzodiazepine<br>pyrrolo derivs., antibody conjugates   |   |
| 504  | p.169  | 108852-90-0D  | 5,12-Naphthacenedione, 7,8,9,10-tetrahydro-6,8,11-trihydroxy-8-(2-hydroxyacetyl)-1-methoxy-10-[[2,3,6-trideoxy-3-[(2S)-2-methoxy-4-morpholinyl]-α-L-lyxo-hexopyranosyl]oxy]-, (8S,10S)-<br>derivs., antibody conjugates  |    |
| 505  | p.169  | 1214986-71-6D | 1H-Benz[e]indole, 1-(chloromethyl)-2,3-dihydro-<br>derivs., antibody conjugates<br>与抗体偶联的小分子   |   |
| 506  | p.170  | 1463439-37-3D | 5,12-Naphthacenedione, 8-(1-hydrazinylidene-2-hydroxyethyl)-7,8,9,10-tetrahydro-6,8,11-trihydroxy-1-methoxy-10-[[[1S,3R,4aS,9S,9aR,10aS]-octahydro-9-methoxy-1-methyl-1H-pyrano[4',3':4,5]oxazolo[2,3-c][1,4]oxazin-3-yl]oxy]-, (8S,10S)-<br>antibody conjugates |  |

# 获取相关序列信息

专利全文中的序列

**Table of Sequences**

| NAME  | SEQUENCE  | SEQ ID NO |
|---|---|-----------|
| Humanized 7C2.v2.2.LA ("hu7C2") light chain variable region | DIVMTQSPDS LAVSLGERAT INCRASQSVS GSRFTYMHY QQKPGQPPKL LIKYASILES GVPDRFSGSG SGTDFTLTIS SLQAEDVAVY YCQHSWEIPP WTFQGGTKVE IK        | 10        |
| Hu7C2 heavy chain variable region                           | EVQLVQSGAE VKKPGASVKV SCKASGYSFT GYWMNWVRQA PGQGLEWIGM IHPLDAEIRA NQKFRDRVTI TVDTSTSTAY LELSSLRSED TAVYYCARGT YDGGFEYWGQ GTLVTVSS | 11        |
| Hu7C2 HVR-L1  | RASQSVSGSRFTYMH   | 12        |
| Hu7C2 HVR-L2  | YASILES   | 13        |
| Hu7C2 HVR-L3  | QHSWEIPPWT  | 14        |
| Hu7C2 HVR-H1  | GYWMN   | 15        |
| Hu7C2 HVR-H2 (Hu7C2.v2.1.S53L, S55A HVR-H2)                 | MIHPLDAEIRANQKFRD   | 16        |
| Hu7C2 HVR-H3  | GYDGGFEY  | 17        |



# Prior Art Analysis 分析目标专利的现有技术

Reference Detail (2 of 16) [详情页面](#)

Substances (114) Reactions (203) Citing (4) Citation Map

PatentPak Viewer **Get Prior Art Analysis** Full Text ▾

### Prior Art Analysis



SciFinder<sup>®</sup> will search for all relevant information and send you a Prior Art Analysis based on this document. You can view the status of the request in your Recent Search History and on the History Page. You may continue to search while the analysis is processing.

Email Address

**Submit** Cancel [发送报告邮件](#)

利用机器学习深度分析目标专利  
相关的现有技术

## SciFinder-n Prior Art Analysis Complete

 SciFinder-n Prior Art Analysis Results <no-replies@cas.org>  
To  Liu, Mengmeng Y

[↩ Reply](#) [↩ Reply All](#) [→ Forward](#) [⋮](#)

X Thu 3/17/2022 8:05 AM

 Some of the content in this message couldn't be downloaded because you're working offline or aren't connected to a network.

### Your Prior Art Analysis is Complete

This analysis is based on the reference:

[Anti-HER2 antibodies and immunoconjugates and therapeutic and diagnostic uses thereof](#) [当前文献](#)

By: Chen, Xiaocheng; Dennis, Mark; Junutula, Jagath Reddy; Phillips, Gail Lewis; Pillow, Thomas Harden; Sliwkowski, Mark X.  
World Intellectual Property Organization | English | WO2016040723

Here are the top similar documents based on your request. You can view all references in SciFinder<sup>®</sup>. [相关文献](#)

[Preparation of anti-human LgR5 antibodies and immunoconjugates for cancer diagnosis and therapy](#)

By: Hongo, Jo-Anne; Mao, Weiguang; Polakis, Paul; Polson, Andrew; Vij, Rajesh; Wu, Yan; Liang, Wei-Ching  
World Intellectual Property Organization | English | WO2013149159

[Anti-Ly-6E antibodies and immunoconjugates for diagnosis and therapy](#)

By: Chang, Peter; Sakanaka, Chie  
World Intellectual Property Organization | English | WO2013177055

[Pyrrolobenzodiazepines and conjugates as Antitumor agents](#)

By: Flygare, John A.; Gunzner-Toste, Janet L.; Pillow, Thomas; Howard, Philip Wilson; Masterson, Luke  
World Intellectual Property Organization | English | WO2013055987

[Antibodies against endothelin B receptor \(ETBR\) and immunoconjugates in treating cancer, particularly melanoma](#)

By: Asundi, Jyoti; Clark, Suzanna; Polakis, Paul  
World Intellectual Property Organization | English | WO2014022680

[Immunoconjugates of anti-CD22 antibodies and nemorubicin for the treatment of leukemia and lymphoma](#)

By: Polakis, Paul; Polson, Andrew; Spencer, Susan Diane; Yu, Shang-Fan; Zheng, Bing  
World Intellectual Property Organization | English | WO2014011520

Showing the Top Documents

[View all results in](#) 

# Prior Art Analysis获取现有技术分析文献

Recent Search History View All Search History

March 17, 2022

|                       |   |  |
|-----------------------|---|--|
| References<br>8:03 AM | Prior Art Analysis (183)<br>Anti-HER2 antibodies and immunoconjugates and therapeutic and diagnostic uses thereof | <a href="#">View Results</a><br>Complete |
|-----------------------|---|--|

Filter Behavior

[Filter by](#) [Exclude](#)

Document Type

- Journal (88)
- Patent (95)
- Review (31)
- Clinical Trial (4)
- Commentary (1)
- Letter (1)

Language

- English (171)
- Japanese (6)
- Chinese (4)
- French (2)

Publication Year

References (183) Sort: Relevance View: Partial Abstract

Substances  Reactions  Citing

1

**Preparation of anti-human LgR5 antibodies and immunoconjugates for cancer diagnosis and therapy**  
By: Hongo, Jo-Anne; Mao, Weiguang; Polakis, Paul; Polson, Andrew; Vij, Rajesh; Wu, Yan; Liang, Wei-Ching  
World Intellectual Property Organization, WO2013149159 A1 2013-10-03 | Language: English, Database: CAPlus

The invention provides anti-protein LgR5 (leucine-rich repeat-containing G protein-coupled receptor 5) antibodies and immunoconjugates and methods of using the same. Positron emitter-labeled and cytotoxic agent-conjugated anti-LgR5 antibodies are used for cancer diagnosis and therapy.

[PatentPak](#) [Full Text](#)  Substances (95)  Reactions (0)  Citing (4)  Citation Map

2

**Anti-Ly-6E antibodies and immunoconjugates for diagnosis and therapy**  
By: Chang, Peter; Sakanaka, Chie  
World Intellectual Property Organization, WO2013177055 A2 2013-11-28 | Language: English, Database: CAPlus

The authors disclose the preparation and characterization of humanized antibodies that bind to lymphocyte antigen 6E (Ly-6E). In one example, auristatin E conjugated to an anti-Ly-6E antibody is shown to inhibit tumor growth in xenograft models.

可以去相关现有技术文献中获取相关序列

# 获取序列结果集的相关文献结果集

## CDR Segments

Select a segment below to view individual or intersecting CDR results.

CDR1 CDR2 CDR3

Apply Reset Segments

## Bioscape Analysis

Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)

Create Bioscape Analysis

## Filter by

E-Value

0 to 10<sup>6</sup>

Query Coverage %

0 to 100

## Biosequences (159)

Sort: Alignment Identity View: Expanded

References

Query Details View Less

- > CDR1  
RASQSVSGSRFTYMH
- > CDR2  
YASILES
- > CDR3  
QHSWEIPPWT

1 Alignment Identity: 100%

Subject 1 112

Matches: 32  
Mismatches: 0

View Less

Alignment Subject References

CAS Registry Numbers: 2411616-63-0, 2756302-24-4  
Length: 112 aa

Sequence

```
1 DIVLTQSPAS LVVSLGQRAT ISCRASQSVS GSRFTYMHWY QQKPGQPKL LIKYASILES GVPARFSGGG SGTDFTLNIH
81 PVEEDDTATY YCQHSWEIPP WTFGGGTKLE IK
```

# 获取序列结果集的相关文献结果集


Filter Behavior

**Filter by** Exclude

Document Type

Language

Publication Year



No Min to No Max Apply

View Larger

Available at My Institution

Author

Organization

Publication Name

**Concept**

CA Section

CAS Solutions

Formulation Purpose

Database

Search Within Results

## References (294)

Sort: Publication Date: Newest View: Partial Abstract

Substances Reactions Citing

Download Email Save

1

### Compositions and methods for treating autoimmune diseases and cancers by targeting IGSF8

By: Xiao, Tengfei; Hu, Xihao; Liu, Xiaole  
World Intellectual Property Organization, WO2022033419 A2 2022-02-17 | Language: English, Database: CAplus

The present invention provides methods and compositions for treating a cancer, and/or an autoimmune disease, by modulating the expression and/or activity of IGSF8 and its binding ligands. The pharmaceutical compositions may include, but are not limited to, antibodies that specifically bind human IGSF8, and have an activity of inhibiting IGSF8-mediated immunosuppression in a subject in need thereof.

PatentPak Full Text Substances (1,060) Reactions (0) Citing (0) Citation Map

2

### Dendritic cell activating therapy as an adjunct to radiation therapy

By: Guha, Chandan; Pandey, Sanjay  
World Intellectual Property Organization, WO2022032043 A1 2022-02-10 | Language: English, Database: CAplus

Provided herein are methods relating to administering a dendritic cell activating therapy as an adjunct to radiation therapy or an energy-based therapy for treating a tumor or cancer in an individual.

PatentPak Full Text Substances (41) Reactions (0) Citing (0) Citation Map

3

### Armed NK cells for universal cell therapy

By: Villalba, Martin; Jorgensen, Christian; Robert, Bruno; Martineau, Pierre; Hernandez, Francisco Javier  
France, FR3112939 A1 2022-02-04 | Language: French, Database: CAplus

The present invention relates to the field of therapeutic treatment, particularly of cell therapy based on CD 16+ cells and NK (Natural Killer) cells. In particular, the invention relates to a pharmaceutical composition comprising a CD 16+ cell, a NK cell or a NK cell



ACS  
International



A division of the  
American Chemical Society





# 通过Concept聚焦研究重点

Concept ×

Top Count | Alphanumeric | Search

1 Selected

|   |  |  |
|---|--|--|
| <input type="checkbox"/> Homo sapiens (256)                   | <input type="checkbox"/> Mucin 1 (59)                                | <input type="checkbox"/> Tumor antigens (48)                       |
| <input type="checkbox"/> Human (256)                          | <input type="checkbox"/> Single-chain antibodies (59)                | <input type="checkbox"/> Anti-inflammatory agents (47)             |
| <input type="checkbox"/> Antibodies and Immunoglobulins (149) | <input type="checkbox"/> Chimeric fusion proteins (58)               | <input type="checkbox"/> CD2 antigens (47)                         |
| <input type="checkbox"/> Antitumor agents (135)               | <input type="checkbox"/> Epidermal growth factor receptors (58)      | <input type="checkbox"/> FcεRII receptors (47)                     |
| <input type="checkbox"/> Immunoglobulin fragments (117)       | <input type="checkbox"/> Soluble tumor necrosis factors (58)         | <input type="checkbox"/> FcγRI receptors (47)                      |
| <input type="checkbox"/> Protein sequences (115)              | <input type="checkbox"/> Bispecific antibodies (57)                  | <input type="checkbox"/> Hormones, animal (47)                     |
| <input type="checkbox"/> Neoplasm (100)                       | <input type="checkbox"/> CD25 antigens (57)                          | <input type="checkbox"/> Stomach neoplasm (47)                     |
| <input type="checkbox"/> Monoclonal antibodies (93)           | <input type="checkbox"/> CD30 antigens (57)                          | <input type="checkbox"/> Uterine cervical neoplasm (47)            |
| <input type="checkbox"/> Proteins (92)                        | <input type="checkbox"/> Cytokines (57)                              | <input type="checkbox"/> CD52 antigens (46)                        |
| <input type="checkbox"/> Mammary gland neoplasm (85)          | <input checked="" type="checkbox"/> Immunoglobulin light chains (57) | <input type="checkbox"/> Inflammation (46)                         |
| <input type="checkbox"/> Humanized antibodies (80)            | <input type="checkbox"/> Multiple sclerosis (57)                     | <input type="checkbox"/> Passive immunotherapy (46)                |
| <input type="checkbox"/> Pancreatic neoplasm (80)             | <input type="checkbox"/> Ulcerative colitis (57)                     | <input type="checkbox"/> Protein motifs (46)                       |
| <input type="checkbox"/> Interleukin 2 (77)                   | <input type="checkbox"/> CD19 antigens (56)                          | <input type="checkbox"/> Radionuclides (46)                        |
| <input type="checkbox"/> Interleukin 6 (77)                   | <input type="checkbox"/> Interleukin 17 (56)                         | <input type="checkbox"/> Sialic acid-binding Ig-like lectin 3 (46) |
|   |  | <input type="checkbox"/> CD45 antigens (45)                        |

|  |   |  |
|--|---|--|
| <input type="checkbox"/> CD40 ligands (70)                           | <input type="checkbox"/> Kidney neoplasm (55)   | <input type="checkbox"/> CD4 antigens (44)                           |
| <input type="checkbox"/> Lung neoplasm (70)                          | <input checked="" type="checkbox"/> Epidermal growth factor receptor HER2 (54)          | <input type="checkbox"/> Insulin-like growth factor I receptors (44) |
| <input type="checkbox"/> CD40 antigens (68)                          | <input type="checkbox"/> Interleukin 15 (54)  | <input type="checkbox"/> Interleukin 6 receptors (44)                |
| <input type="checkbox"/> Interleukin 12 (66)                         | <input type="checkbox"/> Bladder neoplasm (53)  | <input type="checkbox"/> Lewis blood group antigen Ley (44)          |
| <input type="checkbox"/> Melanoma (65)                               | <input type="checkbox"/> Colorectal neoplasm (53)                                       | <input type="checkbox"/> CD38 antigens (43)                          |
| <input type="checkbox"/> CD20 antigens (64)                          | <input type="checkbox"/> Histocompatibility antigen HLA-DR (53)                         | <input type="checkbox"/> Drugs (43)                                  |
| <input type="checkbox"/> Rheumatoid arthritis (64)                   | <input type="checkbox"/> Immunomodulators (53)  | <input type="checkbox"/> Epithelial cell adhesion molecules (43)     |
| <input type="checkbox"/> Chimeric antibodies (62)                    | <input type="checkbox"/> Interferon α (53)  | <input type="checkbox"/> Epitopes (43)                               |
| <input type="checkbox"/> Peptides (62)                               | <input type="checkbox"/> Interferon β (53)  | <input type="checkbox"/> Folate receptors (43)                       |
| <input type="checkbox"/> CD3 antigens (61)                           | <input type="checkbox"/> Interleukin 18 (52)  | <input type="checkbox"/> Myasthenia gravis (43)                      |
| <input type="checkbox"/> CD80 antigens (61)                          | <input type="checkbox"/> Multiple myeloma (52)  | <input type="checkbox"/> Scleroderma (43)                            |
| <input type="checkbox"/> Immunoconjugates (61)                       | <input type="checkbox"/> Enzymes (51)   | <input type="checkbox"/> Interferons (42)                            |
| <input checked="" type="checkbox"/> Immunoglobulin heavy chains (61) | <input type="checkbox"/> Carcinoembryonic antigen-related cell adhesion molecule 6 (50) | <input type="checkbox"/> Interleukin 10 (42)                         |
| <input type="checkbox"/> Carcinoembryonic antigen (60)               | <input type="checkbox"/> Liver neoplasm (50)  |  |
| <input type="checkbox"/> Combination chemotherapy (60)               |   |  |

# 通过Concept,获取感兴趣主题的相关文献

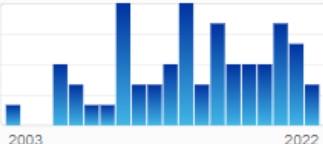
Filter Behavior

**Filter by** Exclude

Document Type

Language

Publication Year



No Min to No Max Apply View Larger

Available at My Institution

Author

Organization

Publication Name

Concept

CA Section

- Immunochemistry (25)
- Pharmaceuticals (16)
- Pharmacology (5)
- Biochemical Methods (4)
- Amino Acids, Peptides, and Proteins (3)

## References (54)

Sort: Publication Date: Newest View: Partial Abstract

Substances Reactions Citing

Filtering: Concept: Epidermal growth factor r... Clear All Filters

1

**Preparation of 6a,7,8,9,10,12-hexahydro-12-oxobenzo[e]pyrido[1,2-a][1,4]diazepine and related compounds as sequence-selective DNA mono-alkylating cytotoxic agents and their antibody conjugates containing peptidyl linkers**

By: Andriollo, Paolo; Jackson, Paul; Thurston, David  
World Intellectual Property Organization, WO2022023735 A1 2022-02-03 | Language: English, Database: CAplus

The invention is related to 6a,7,8,9,10,12-hexahydro-12-oxobenzo[e]pyrido[1,2-a][1,4]diazepine derivatives, analogs and their linker conjugates containing a sigma hole group, e.g., I, II and III [X = S, Se, Te, P, As, Sb, Bi, Si, Ge, Sn or Pb; Y = N, C-NH<sub>2</sub>, C-OH; Z = O, N-Me], their pharmaceutically acceptable salts, tautomers, stereoisomers or their mixtures as DNA-binding compounds, especially DNA-alkylating agents that are useful as medicaments, such as anti-proliferative agents. The pyridinobenzodiazepines derivatives and analogs of the invention linked, either directly or indirectly, to a...

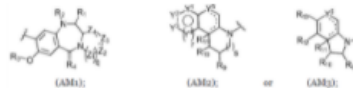
View More

PatentPak Full Text Substances (92) Reactions (61) Citing (0) Citation Map

2

**Preparation of 6a,7,8,9,10,12-hexahydro-12-oxobenzo[e]pyrido[1,2-a][1,4]diazepine and related compounds as sequence-selective DNA mono-alkylating cytotoxic agents and their antibody conjugates containing peptidyl linkers**

By: Andriollo, Paolo; Jackson, Paul Joseph; Thurston, David Edwin  
United Kingdom, GB2597532 A 2022-02-02 | Language: English, Database: CAplus



The invention is related to compounds (TX<sup>4</sup>)<sub>p</sub>B<sup>1</sup>X<sup>3</sup>AX<sup>2</sup>LX<sup>1</sup>-I [A, B<sup>1</sup> = independently (un)substituted Ph, C<sub>5-9</sub> heteroaryl, a sigma hole group; T = C<sub>1-8</sub> alkyl, A, B<sup>1</sup>; Z = (Z<sup>1</sup>)<sub>q</sub>; X<sup>1-4</sup> independently = O, S, CO, etc.; p = 0-1; R<sup>1</sup>, R<sup>2</sup> are selected so that R<sup>1</sup> and R<sup>2</sup> form a double bond, R<sup>1</sup> = H, OH, OC<sub>1-8</sub> alkyl and R<sup>2</sup> = H, N-protecting group, K<sup>1</sup>R<sup>A</sup>, or R<sup>1</sup> = SO<sub>3</sub>H, :O and R<sup>2</sup> =

# 序列检索小贴士

- BLAST 和 MOTIF: 可以提高E-Value 数值, 来增大找到结果的可能性, 特别是对于短序列。
- 如果没有获得检索结果, 可以试着提高检索序列数量的上限。
- BLAST支持同时检索多条序列 (最多100条) , 用FASTA格式的标题隔开即可。

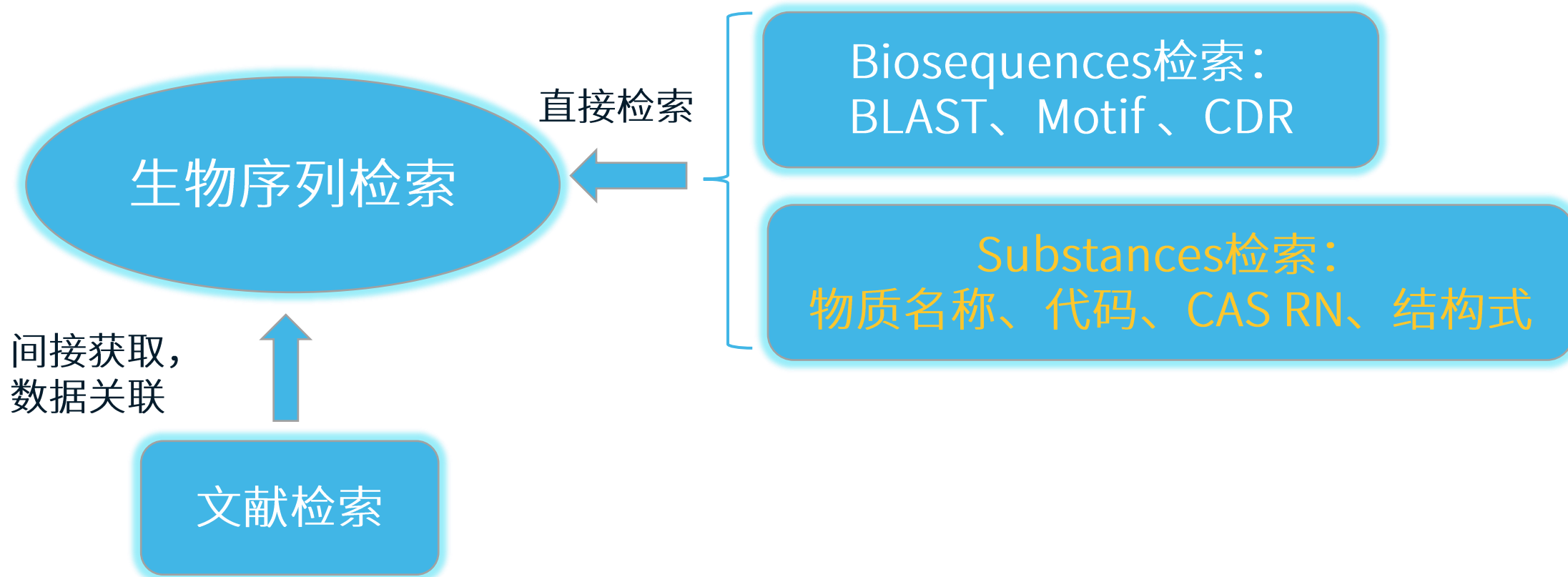
```
BLAST CDR Motif Upload Sequence Clear Search
> pdb|1E57|C Chain C, Physalis Mottle Virus
MDSSEVVKVKQASIPAPGSILSQPNTEQSPAIVLPPFQFEATTFGTAE TAAQVSLQTADPITKLTAPYRHAQIVECKAILT
PTDLAVSNPLTVYLAWVPANSPATPTQILRVYGGQSFVLGGAISSAAKTIEVPLNLDVNRMLKDSVTYTDTPKLLAYSRA
PTNPSKIPTASIQISGRIRLSKPMLIAN
> pdb|1E57|B Chain B, Physalis Mottle Virus
MDSSEVVKVKQASIPAPGSILSQPNTEQSPAIVLPPFQFEATTFGTAE TAAQVSLQTADPITKLTAPYRHAQIVECKAILT
PTDLAVSNPLTVYLAWVPANSPATPTQILRVYGGQSFVLGGAISSAAKTIEVPLNLDVNRMLKDSVTYTDTPKLLAYSRA
PTNPSKIPTASIQISGRIRLSKPMLIAN
```

Fasta 格式:

第一行由大于号">"开头的任意文字说明

第二行开始为序列本身

# 生物序列获取方法



# 通过Substances物质检索获取序列

通过物质名称、代码和CAS RN检索来获取序列

- Chemical name: e.g. [Trastuzumab](#)
- Brand name (drug name): e.g. [Argireline](#)
- Generic name: e.g. [Somavubove](#)
- CAS Registry Number: e.g. [1190939-33-3](#)
- GenBank Number: e.g. [GenBank KT001084](#)

# 通过标识符检索获取序列

支持单独或批量检索

物质标识符批量检索；最多可同时输入2000个字符

trastuzumab argireline somavubove 1190939-33-3 "GenBank KT001084"

Searching for...

- All
- Substances**
- Reactions
- References
- Suppliers
- Biosequences
- Retrosynthesis

### Substances

Search by Substance Name, CAS RN, Patent Number, PubMed ID, AN, CAN, and/or DOI. [Learn More](#)

trastuzumab argireline somavubove 1190939-33-3 "GenBank KT001084" ✕ Draw 🔍

+ Add Advanced Search Field Learn more about SciFinder<sup>®</sup> Advanced Search.

# 通过标识符检索获取序列

Filter Behavior

Commercial Availability

Reaction Role

Product (2)

Reactant (2)

Reagent (1)

Reference Role

Biological Study (5)

Biological Study, Unclassified (5)

Properties (5)

Therapeutic Use (4)

Uses (4)

[View All](#)

Stereochemistry

Number of Components

Substance Class

Isotopes

Metals

Molecular Weight

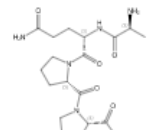
Regulatory Data by Country/Region

## Substances (5)

Sort: Relevance View: Partial

1

**1190939-33-3**



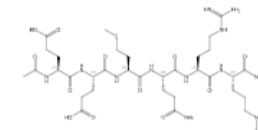
Absolute stereochemistry shown

$C_{18}H_{29}N_5O_6$   
L-Alanyl-L-glutaminyl-L-prolyl-L-proline  
Protein/Peptide Sequence  
Sequence Length: 4

References  Reactions  Suppliers

2

**616204-22-9**



Absolute stereochemistry shown

$C_{34}H_{60}N_{14}O_{12}S$   
**Argireline**  
Protein/Peptide Sequence  
Sequence Length: 6

References  Reactions  Suppliers

3

**180288-69-1**

Image Not Available

Unspecified  
**Trastuzumab**  
Protein/Peptide Sequence  
Sequence Length: 1328

References  Reactions  Suppliers

4

**126752-39-4**

Image Not Available

Unspecified  
**Somavubove**  
Protein/Peptide Sequence  
Sequence Length: 191

References  Reactions  Supplier

5



**2092642-35-6**

Image Not Available

Unspecified  
**GenBank KT001084**  
Nucleic Acid Sequence  
Sequence Length: 22089

Reference  Reactions  Suppliers

© 2022 American Chemical Society. All rights reserved.

 ACS International  
 CAS  
A division of the American Chemical Society

# 根据结构式进行序列检索

The image shows a screenshot of the CAS Draw software interface. The main window displays a complex peptide structure with the molecular formula  $C_{34}H_{60}N_{14}O_{12}S$  (889.00). A 'Templates' panel is open, showing a list of amino acids under the heading 'Amino Acid (25)'. The amino acid Arginine is highlighted in blue. An arrow points from the Arginine template to a search results panel on the right. This panel shows the search ID '616204-22-9', the chemical structure of Argireline, its molecular formula  $C_{34}H_{60}N_{14}O_{12}S$ , the name 'Argireline', and the protein/peptide sequence information: 'Protein/Peptide Sequence' and 'Sequence Length: 6'. At the bottom of the results panel, there are three buttons: '875 References', '120 Reactions', and '40 Suppliers'.

注：序列长度<50个单元时可使用结构进行序列检索



# 点击CAS RN查看序列详情

## Substance Detail (1 of 4)

References (875) Reactions (120) Suppliers (40)

CAS Registry Number  
616204-22-9

Absolute stereochemistry shown

$C_{34}H_{60}N_{14}O_{12}S$   
L-Argininamide, N-acetyl-L-α-glutamyl-L-α-glutamyl-L-methionyl-L-glutamyl-L-arginyl- (9CI, ACI)

| Key Physical Properties | Value                      | Condition                    |
|-------------------------|----------------------------|------------------------------|
| Molecular Weight        | 888.99                     | -                            |
| Density (Predicted)     | 1.54±0.1 g/cm <sup>3</sup> | Temp: 20 °C; Press: 760 Torr |
| pKa (Predicted)         | 4.43±0.10                  | Most Acidic Temp: 25 °C      |

Protein/Peptide Sequence  
Sequence Length: 6  
modified

Related Sequences (88)

Other Names and Identifiers

Sequence Details

Sequence: linear **氨基酸序列**

1 EEMQRR - - - -

Sequence Modifications **序列修饰信息**

| Type          | Location        | Description      |
|---------------|-----------------|------------------|
| terminal mod. | glutamic acid-1 | N-acetyl         |
| terminal mod. | arginine-6      | C-terminal amide |

Predicted Properties

Predicted Spectra

Target Indicators

Regulatory Information

82 **1435520-99-2**

Absolute stereochemistry shown

$C_{56}H_{96}N_{14}O_{20}S$   
Protein/Peptide Sequence  
Sequence Length: 6

1 Reference 2 Reactions 0 Suppliers

83 **1435520-98-1**

Absolute stereochemistry shown

$C_{36}H_{64}N_{14}O_{12}S$   
Protein/Peptide Sequence  
Sequence Length: 6

1 Reference 3 Reactions 0 Suppliers

84 **1336986-67-4**

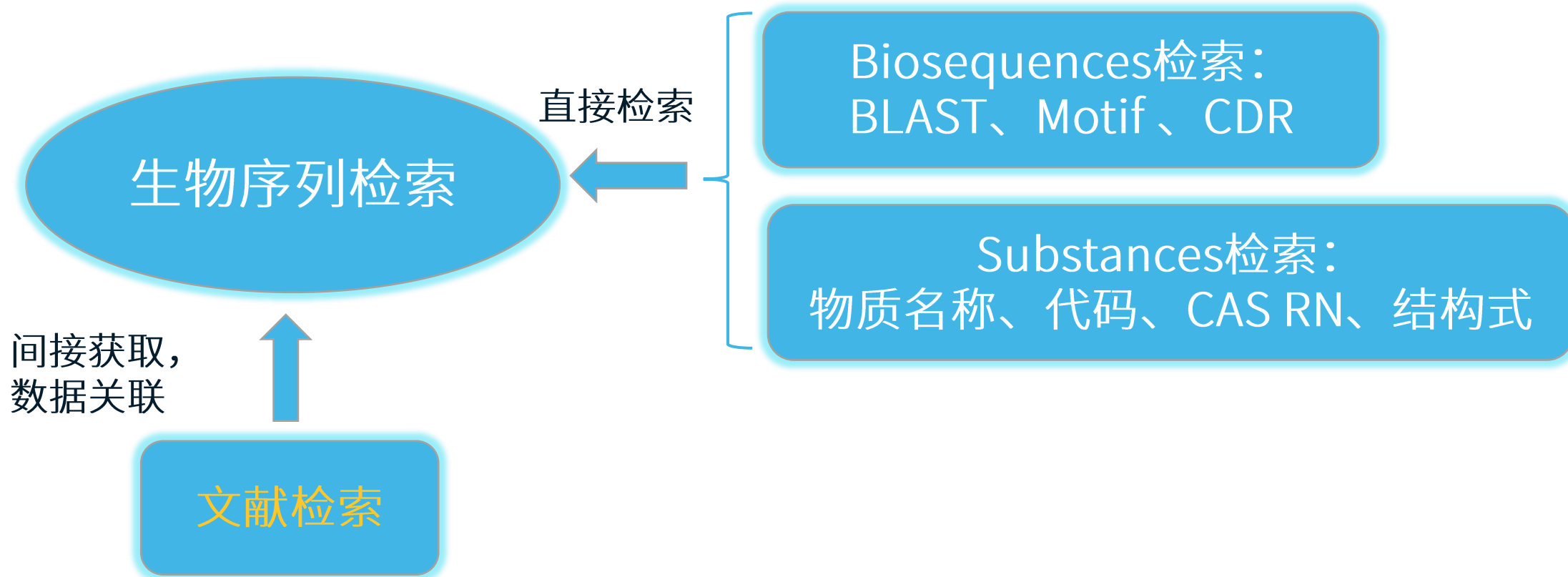
Absolute stereochemistry shown

$C_{48}H_{70}N_{16}O_{16}S$   
L-Ornithinamide, N-acetyl-L-α-glutamyl-L-α-glutamyl-L-methionyl-L-glutamyl-L-N<sup>5</sup>-...  
Protein/Peptide Sequence  
Sequence Length: 6

1 Reference 5 Reactions 0 Suppliers

获取到更多序列信息

# 生物序列获取方法



# 通过文献检索获取生物序列信息

示例 1: 通过主题检索获取菌种相关序列信息

主题词: *Escherichia coli*

Searching for...

- All
- Substances
- Reactions
- References**
- Suppliers
- Biosequences
- Retrosynthesis

## References

Search by Keyword, Substance Name, CAS RN, Patent Number, PubMed ID, AN, CAN, and/or DOI. [Learn More](#)

Escherichia coli ✕ Draw 🔍

+ Add Advanced Search Field Learn more about SciFinder<sup>®</sup> Advanced Search.

# 通过文献检索获取菌种相关序列信息

Filter Behavior

Document Type

Language

Publication Year

Available at My Institution

Author

Organization

Publication Name

Concept

- Escherichia coli (555K)
- Humans (115K)
- Animals (111K)
- Staphylococcus aureus (85K)
- Microbial gene (72K)

[View All](#)

CA Section

CAS Solutions

## References (555,907)

Sort: Relevance View: Partial Abstract

Substances  Reactions  Citing

Filtering: Concept: Escherichia coli  [Clear All Filters](#)

1

### Pathogenic Escherichia coli

By: Kaper, James; Nataro, James P.; Mobley, Harry L. T.  
Nature Reviews Microbiology (2004), 2(2), 123-140 | Language: English, Database: CAPLUS and MEDLINE

Few Microorganisms are as versatile as **Escherichia coli**. An important member of the normal intestinal microflora of humans and other mammals, **E. coli** has also been widely exploited as a cloning host in recombinant DNA technol. But **E. coli** is more than just a laboratory workhorse or harmless intestinal inhabitant; it can also be a highly versatile, and frequently deadly, pathogen. Several different **E. coli** strains cause diverse intestinal and extraintestinal diseases by means of virulence factors that affect a wide range of cellular processes.

2

### Diarrheagenic Escherichia coli

By: Nataro, James P.; Kaper, James B.  
Clinical Microbiology Reviews (1998), 11(1), 142-201 | Language: English, Database: CAPLUS and MEDLINE

A review with 719 references

# 通过文献检索获取菌种相关序列信息

Filter Behavior

Filter by Exclude

- Commercial Availability
- Reaction Role
- Reference Role
- Stereochemistry
- Number of Components
- Substance Class
  - Manual Registration (1.4M)
  - Nucleic Acid Sequence (854K)**
  - Protein/Peptide Sequence (727K)**
  - Organic/Inorganic Small Molecule (411K)
  - Salt and Compound With (41K)
- Isotopes
- Metals
- Molecular Weight
- Experimental Property
- Experimental Spectrum

## Substances (1,581,927)

Sort: Relevance View: Partial

References Reactions Suppliers

Filtering: Substance Class: 2 Selected X Clear All Filters

|   |   |   |
|---|---|---|
| <input type="checkbox"/> 1<br><b>63182-55-8</b><br>Image Not Available<br>Unspecified<br>Galactosidase, $\beta$ - (Escherichia coli reduced)<br>Protein/Peptide Sequence<br>Sequence Length: 1023<br>3 References 0 Reactions 0 Suppliers | <input type="checkbox"/> 2<br><b>86090-42-8</b><br>Image Not Available<br>Unspecified<br>DNA (Escherichia coli gene lacZ)<br>Nucleic Acid Sequence<br>Sequence Length: 3078<br>1 Reference 0 Reactions 0 Suppliers          | <input type="checkbox"/> 3<br><b>110736-92-0</b><br>Image Not Available<br>Unspecified<br>Protein (Escherichia coli gene tolQ)<br>Protein/Peptide Sequence<br>Sequence Length: 231<br>49 References 0 Reactions 2 Suppliers |
| <input type="checkbox"/> 4<br><b>111694-26-9</b><br>Image Not Available<br>$C_{436}H_{677}N_{111}O_{136}S_7$<br>Toxin SLT-II (bacteriophage 933W B-subunit precursor reduced)<br>Protein/Peptide Sequence<br>Sequence Length: 89          | <input type="checkbox"/> 5<br><b>127736-36-1</b><br>Image Not Available<br>$C_{331}H_{505}N_{85}O_{104}S_2$<br>Protein CS 7.4 (Escherichia coli clone pJJ G01 precursor)<br>Protein/Peptide Sequence<br>Sequence Length: 70 | <input type="checkbox"/> 6<br><b>115537-70-7</b><br>Image Not Available<br>$C_{487}H_{801}N_{141}O_{151}S_6$<br>Protein (Escherichia coli clone pRJ741 gene fis)<br>Protein/Peptide Sequence<br>Sequence Length: 98         |

# 通过文献检索获取菌种相关序列信息

文献主题检索 → 获取文献结果集，并通过Concept精炼 → 获取结果集的相关物质 → 物质分类 → 获取序列结果集

## Substance Detail (1 of 1)

Reference (1) Reactions (0) Suppliers (0)

CAS Registry Number  
**140104-20-7**

Image Not Available

**Unspecified**  
DNA (Escherichia coli clone λFP1 321-amino acid protein gene plus flanks) (9CI)

**Nucleic Acid Sequence**  
Sequence Length: 1402  
372 a, 326 c, 354 g, 350 t

Other Names and Identifiers

**Sequence Details**

Sequence: DNA; linear

|    |            |            |            |            |            |
|----|------------|------------|------------|------------|------------|
| 1  | aagcttatgc | cgatagcttc | gcactggacc | cggtcgtgga | aaaagaagag |
| 51 | tggtgccgta | ttaccggtcg | taagaaatta | accttcgcat | cgccgtaggt |

## GenBank Definitions & Feature Table

**Definition**  
E.coli orf1 (cotranscribed with fis promoter).

**Organism**  
Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia

**Accession Number**  
X62399

**Version Number**  
X62399.1

| Feature Key     | Location   | Qualifier  |
|-----------------|------------|--|
| source          | 1..1402    | /organism="Escherichia coli"<br>/mol_type="genomic DNA"<br>/strain="CSH50"<br>/db_xref="taxon:562" |
| regulatory      | 364..370   | /regulatory_class="minus_10_signal"  |
| gene            | 375..>1374 | /gene="ORF1"   |
| prim_transcript | 375..>1374 | /gene="ORF1"<br>/experiment="experimental evidence,<br>tails recorded"                             |
| CDS             | 409..1374  | /gene="ORF1"   |

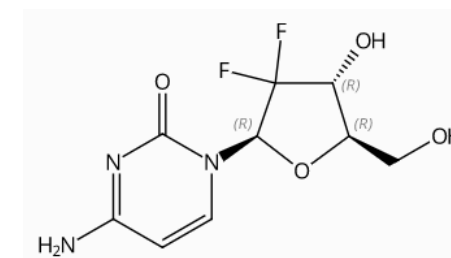
# 通过文献检索获取生物序列信息

示例 2: 主题和结构检索联用--获取吉西他滨相关PDC的序列信息

主题词: "peptide drug conjugate"

结构:

The screenshot shows the CAS SciFinder interface with search results for "peptide drug conjugate". The left sidebar contains filters for Structure Match (As Drawn, Substructure), Filter Behavior (Filter by, Exclude), Document Type (Journal, Patent, Review), Substance Role, Language, and Publication Year (2012-2022). The main area displays a list of references. The first reference is titled "Integrin-Targeting Knottin Peptide-Drug Conjugates Are Potent Inhibitors of Tumor Cell Proliferation" by Cox, Nick; Kintzing, James R.; Smith, Mark; Grant, Gerald A.; Cochran, Jennifer R. The second reference is titled "Peptide conjugates with small molecules designed to enhance efficacy and safety" by He, Rongjun; Finan, Brian; Mayer, John P.; DiMarchi, Richard D. A chemical structure of a peptide-drug conjugate is shown in a pop-up window above the first reference.



# 通过文献检索获取生物序列信息

Filter Behavior

Filter by Exclude

- Commercial Availability
- Reaction Role
- Reference Role
- Stereochemistry
- Number of Components
- Substance Class
  - Organic/Inorganic Small Molecule (857)
  - Protein/Peptide Sequence (98)
  - Manual Registration (81)
  - Polymer (45)
  - Salt and Compound With (21)
  - Element (20)
  - Ring Parent (7)
  - Coordination Compound (5)
  - Incompletely Defined Substance (5)
  - Mixture (2)

View Fewer

Substances (100) Sort: Relevance View: Partial

References Reactions Suppliers Save And Alerts

Filtering: Substance Class: 2 Selected X Clear All Filters

|  |   |   |
|--|---|---|
| <input type="checkbox"/> 1<br><b>1438846-77-5</b><br>Image Not Available<br><chem>C131H201N43O45S6</chem><br>Glycinamide, glycyl-L-cysteinyl-L-prolyl-L-arginyl-L-prolyl-L-arginylglycyl-L-α-...<br>Protein/Peptide Sequence<br>Sequence Length: 33<br>References: 5 Reactions: 9 Suppliers: 0 | <input type="checkbox"/> 2<br><b>1956423-40-7</b><br>Image Not Available<br><chem>C133H204N46O44S6</chem><br>Protein/Peptide Sequence<br>Sequence Length: 33<br>References: 2 Reactions: 5 Suppliers: 0 | <input type="checkbox"/> 3<br><b>1956294-73-7</b><br>Image Not Available<br><chem>C147H219F2N49O49S6</chem><br>Protein/Peptide Sequence<br>Sequence Length: 33<br>References: 2 Reactions: 5 Suppliers: 0 |
| <input type="checkbox"/> 4<br><b>1956294-72-6</b><br>Image Not Available<br><chem>C147H219F2N49O49S6</chem><br>Protein/Peptide Sequence<br>Sequence Length: 33   | <input type="checkbox"/> 5<br><b>33515-09-2</b><br>Absolute stereochemistry shown<br><chem>C55H75N17O13</chem><br>Gonadotropin-releasing hormone<br>Protein/Peptide Sequence                            | <input type="checkbox"/> 6<br><b>52671-12-2</b><br>Absolute stereochemistry shown<br><chem>C59H84N18O13</chem><br>Gonadoliberin A<br>Protein/Peptide Sequence   |



# 大纲

- CAS和CAS SciFinder<sup>n</sup>简介
- 生物序列及相关信息的获取策略
  - 序列获取方法
  - 序列检索结果的精炼和分析 & 序列相关信息的获取
  - 序列相关抗体偶联药物 (ADC) 信息的获取
- 在线演示以及Q&A

# 获取序列相关抗体偶联药物信息

示例：与emtansine偶联的含有特定CDR的抗体信息获取

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

## Biosequences

Enter a protein string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

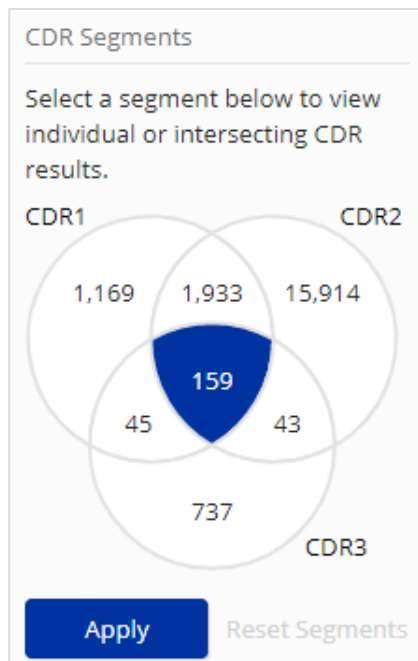
BLAST **CDR** Motif Upload Sequence Clear Search

|      |  |                                  |
|------|--|----------------------------------|
| CDR1 | <input type="text" value="RASQSVSGSRFTYMH"/> | <input type="button" value="x"/> |
| CDR2 | <input type="text" value="YASILES"/>         | <input type="button" value="x"/> |
| CDR3 | <input type="text" value="QHSWEIPPWT"/>      | <input type="button" value="x"/> |

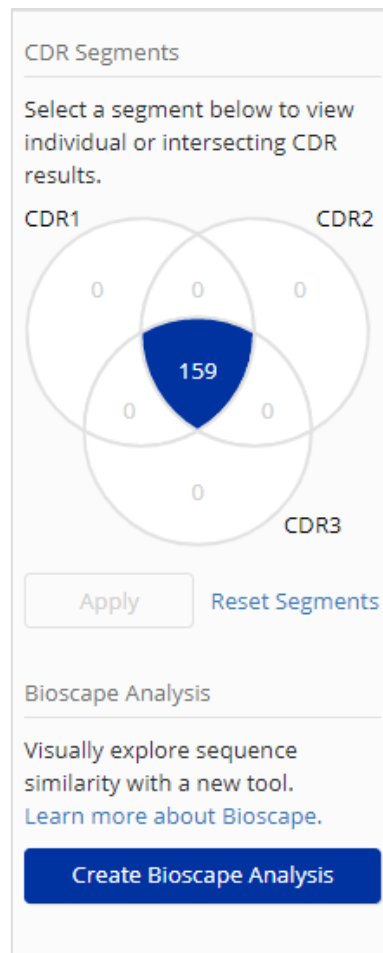
Include NCBI Sequences

Limit Total Sequence Results to:

# 通过CDR检索获取相关文献



查看3个CDR均被包含的序列结果



Biosequences (159) Sort: Alignment Identity View: Expanded

References 获取相关文献

Query Details View Less

> CDR1  
RASQSVSGSRFTYMH

> CDR2  
YASILES

> CDR3  
QHSWEIPPWT

1 Alignment Identity: 100%

Subject 1 112

Matches: 32  
Mismatches: 0


# 利用二次检索，定位与特定小分子相关文献

Filter Behavior

**Filter by** Exclude

- Document Type
- Language
- Publication Year
- Available at My Institution
- Author
- Organization
- Publication Name
- Concept
- CA Section
- CAS Solutions
- Formulation Purpose
- Database
- Search Within Results**

Filter Content Report

Download filter data from this result set. 

## References (294)

Sort: Publication Date: Newest View: Partial Abstract

Substances  Reactions  Citing

1

### Compositions and methods for treating autoimmune diseases and cancers by targeting IGSF8

By: Xiao, Tengfei; Hu, Xihao; Liu, Xiaole  
World Intellectual Property Organization, WO2022033419 A2 2022-02-17 | Language: English, Database: CAplus

The present invention provides methods and compositions for treating a cancer, and/or an autoimmune disease, by modulating the expression and/or activity of IGSF8 and its binding ligands. The pharmaceutical compositions may include, but are not limited to, antibodies that specifically bind human IGSF8, and have an activity of inhibiting IGSF8-mediated immunosuppression in a subject in need thereof.

2

### Dendritic cell activating therapy as an adjunct to radiation therapy

By: Guha, Chandan; Pandey, Sanjay  
World Intellectual Property Organization, WO2022032043 A1 2022-02-10 | Language: English, Database: CAplus

Provided herein are methods relating to administering a dendritic cell activating therapy as an adjunct to radiation therapy or an energy-based therapy for treating a tumor or cancer in an individual.

3

### Armed NK cells for universal cell therapy

By: Villalba, Martin; Jorgensen, Christian; Robert, Bruno; Martineau, Pierre; Hernandez, Francisco Javier

# 获取特定小分子相关文献

^ Search Within Results

Search for up to 3 text strings within the result set.

## References (15)

Sort: Publication Date: Newest ▾ View: Full Abstract ▾

Substances ▾  Reactions ▾  Citing ▾

Filtering: Search Within Results: emtansine ×

1

**Antibody-payload conjugates with enhanced delivery domain and uses thereof** [文献详情](#)

By: Lee, Keun Ho; Lin, Leo Yen-Cheng; Vaidyanathan, Ranjani; Lario, Paula  
World Intellectual Property Organization, WO2021092672 A1 2021-05-20 | Language: English, Database: CAPLUS

The present invention provides a covalent conjugate. The conjugate includes an antibody or antibody derivative, at least two LL37-derived polypeptides, and a payload. The antibody or antibody derivative targets a cell that has phosphatidylserine in its outer leaflet. The payload includes: a small mol. cytotoxic drug of less than 3 kDa, or a plurality thereof; or a peptide or protein of less than 100 kDa. Uses and methods of using these covalent conjugates are also provided, related to enhancing delivery of the antibody/derivative or the payload, e.g. to enhance therapeutic or diagnostic effectiveness.

2

**ANTIKOR-POLIMER KONJUGATLARI**

# 高亮显示目标小分子物质信息

CAS SciFinder<sup>®</sup> References

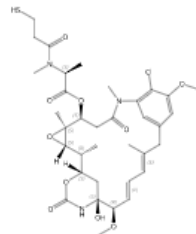
Reference Detail (1 of 15)

Substances (562) Reactions (0)

## Substances

Substances (562)

139504-50-0



Absolute stereochemistry shown  
Double bond geometry shown

$C_{35}H_{48}ClN_3O_{10}S$

**Emtansine**

**PatentPak**

Role: Therapeutic Use, Biological Study, Uses

Notes: conjugates with antibody and LL37-derived peptides

2646630-87-5

Image Not Available

**Unspecified**

HFE 7A

**PatentPak**

Role: Therapeutic Use, Biological Study, Uses

# 获取与抗体偶联的药物信息

Concept下获取更多与抗体偶联的小分子信息

| ^ Concepts  |   |
|---|---|
| <b>Anthracyclines</b><br>Modifier: <b>conjugates</b><br>Role: Therapeutic Use | <b>Histone deacetylase inhibitors</b><br><hr/> <b>Histone H1</b><br>Role: Biological Study, Unclassified    |
| <b>CD80 antigens</b><br>Role: Biological Study, Unclassified                  | <b>Pyrrrolobenzodiazepines</b><br>Modifier: D6.5, <b>conjugates</b> with antibody and LL37-derived peptides |
| <b>CD150 antigens</b><br>Role: Biological Study, Unclassified                 | <b>Maytansinoids</b><br>Modifier: <b>conjugates</b><br>Role: Therapeutic Use                                |

通过CDR检索  
获取相关序列结果集



获取序列结果集的相关文献



通过Search within results  
检索目标小分子



通过文献详情页面，  
Substances和Concepts下的标引，  
获取与抗体偶联的药物信息

# 在自动保存的检索历史中查看检索&检索结果

CAS SciFinder<sup>®</sup> Substances ▾ Enter a query... Draw 🔍 🔔 🕒 👤

Filter by

- Result Type
  - All (68)
  - Biosequences (180)
  - Patent Markush (16)
  - Prior Art Analysis (5)
  - Reactions (29)
  - References (348)
  - Retrosynthesis (17)
  - Substances (232)
  - Suppliers (4)
- Date
  - Start Date: mm/dd/yyyy to End Date: mm/dd/yyyy
  - March, 2022
  - SU MO TU WE TH FR SA
  - 27 28 1 2 3 4 5
  - 6 7 8 9 10 11 12
  - 13 14 15 16 17 18 19
  - 20 21 22 23 24 25 26
  - 27 28 29 30 31 + 2

## 🕒 Your Search History

180 Searches

March 16, 2022

|                          |                           |  |  |                              |
|--------------------------|---------------------------|--|--|------------------------------|
| <input type="checkbox"/> | 👁 Biosequences<br>9:29 AM | NCBI Included: No<br>Results will expire on<br>Apr 15, 2022. | CDR1: RASQSVSGSRFTYMH<br>CDR2: YASILES<br>CDR3: QHSWEIPPWT | <a href="#">View Results</a> |
|                          |                           |  |  | <a href="#">Edit Search</a>  |
|                          |                           |  |  | Complete                     |

March 15, 2022

|                          |                           |  |   |                                       |
|--------------------------|---------------------------|--|---|---------------------------------------|
| <input type="checkbox"/> | 👁 Biosequences<br>9:15 PM | Sequence Type: Protein<br>NCBI Included: No<br>Query Coverage: 90%<br>E-value: 10<br>Results will expire on<br>Apr 15, 2022. | [SG]X{4}GK[DT]<br>4 Biosequences Submitted<br><a href="#">View Individual Sequences</a> ▾ | <a href="#">View Results (4 of 4)</a> |
|                          |                           |  |   | <a href="#">Edit Search</a>           |

[View Individual Sequences](#) ▾



# 生物序列及相关信息的获取策略总结

## 一、序列获取方法：

- |   |    |
|---|----|
| ✓ Biosequences 检索：BLAST、Motif、CDR检索   | 直接 |
| ✓ Substances 检索：物质名称、代码、CAS RN、结构式检索  |    |
| ✓ References 检索：（关键词或文本和结构联用），获取文献，之后获取文献相关物质，再通过Substance Class筛选选项下选择序列相关选项，来获取序列。对目标专利文献，可利用PatentPak获取更多相关序列信息；利用Prior Art Analysis 获取相关现有技术文献中的相关序列信息。 | 间接 |
| ✓ 序列的物质详情页面：利用Related Sequences获取更多相关序列   |    |

## 二、序列检索结果的精炼和分析 & 序列相关信息的获取

- ✓ Filter by : 多个维度筛选精炼序列检索结果，为不同研究目的获取目标序列。
- ✓ BioScape 专利可视化地图分析：概览相似序列专利布局，了解专利热点，快捷获取披露感兴趣序列的专利信息。
- ✓ 序列信息（包括比对结果）下载，便于进一步整理、分析、保存。
- ✓ 序列相关文献结果集中利用Concepts：快速定位感兴趣的研究点，获取相关文献及序列。

三、获取序列相关ADC信息：CDR检索获取相关序列，再获取序列相关文献，利用Search Within Results筛选与目标小分子相关的文献，在文献详情页面Substances和Concepts下的标引获取目标ADC信息。

# 大纲

- CAS和CAS SciFinder<sup>n</sup>简介
- 生物序列及相关信息的获取策略
  - 序列获取方法
  - 序列检索结果的精炼和分析 & 序列相关信息的获取
  - 序列相关抗体偶联药物 (ADC) 信息的获取
- 在线演示以及Q&A

# 谢谢关注!



## 美国化学文摘社北京代表处

ACS International, Ltd. – Representing CAS  
Unit 1010, Tower B, Raycom InfoTech Park  
2 Kexueyuan Nanlu, Haidian District, Beijing, China  
+86-10-6250 8026/7

[cas.org](http://cas.org)

