

収録範囲	米国特許商標庁 (USPTO) が発行した公開特許・登録特許中の核酸およびアミノ酸配列		
ファイル種類	書誌情報, 配列データベース		
特徴	アラート (自動 SDI 検索) 毎週		
	CAS RN® (CAS登録番号) <input type="checkbox"/>	ページイメージ <input type="checkbox"/>	STN AnaVist <input type="checkbox"/>
	Keep & Share <input checked="" type="checkbox"/>	中間一致・ 後方一致検索 <input checked="" type="checkbox"/>	STN Easy <input type="checkbox"/>
	練習用ファイル <input type="checkbox"/>	構造図 <input type="checkbox"/>	
レコード内容	<ul style="list-style-type: none"> 米国特許商標庁 (USPTO) が発行した公開特許・登録特許中の核酸およびアミノ酸配列を収録しています。 発明の名称 (標題), 抄録, 特許出願人や完全な発明者名, 特許番号, PCT 出願番号や日付などを含む豊富な情報を収録しています。 		
レコード数	28,302,000 件以上 (2011 年 7 月現在) 20,937,000 件以上の核酸配列 (2011 年 7 月現在) 7,365,000 件以上のタンパク質配列 (2011 年 7 月現在)		
収録年代	1981 年 -		
更新頻度	毎週更新 (通常は特許が発行されてから 3 日以内に検索できます)		
言語	英語		
データベース	SequenceBase Corporation		
製作者	3 Dellview Drive Edison, NJ 08820-2545 USA E-mail : mgoffman@sequencebase.com Copyright Holder		
データベース	FIZ Karlsruhe		
代理店	STN Europe P.O. Box 2465 76012 Karlsruhe Germany Phone : +49 7247 808 555 Fax : +49 7247 808 259 E-mail : helpdesk@fiz-karlsruhe.de		
収録源	米国特許商標庁 (USPTO) が発行した公開特許・登録特許		
検索補助資料	<ul style="list-style-type: none"> STN 技術資料 http://www.jaici.or.jp/stn/stn_doc_01.html オンラインヘルプ => HELP DIRECTORY ですべての利用可能なヘルプメッセージが表示されます STNGUIDE ファイル オンラインヘルプの PDF : http://www.stn-international.de/pctgen_help.html USGENE Workshop Manual : http://www.stn-international.com/USGENE_workshop_manual.html GETSIM/BLAST のよくあるご質問 : http://www.stn-international.de/pctgen_help.html 		
利用可能なク	<ul style="list-style-type: none"> ALLBIB AUTHORS BIOSCIENCE CORPSOURCE 	<ul style="list-style-type: none"> MEDICINE PATENTS (HPATENTS) PHARMACOLOGY 	

ヨーロッパ
STN カールスルーエ

FIZ Karlsruhe
P.O. Box 2465
76012 Karlsruhe
Germany
Phone: +49-7247-808-555
Fax: +49-7247-808-259
E-mail: helpdesk@fiz-karlsruhe.de
Internet: www.stn-international.de

日本
STN 東京

一般社団法人 化学情報協会
〒113-0021 東京都文京区本駒込6-25-4 中居ビル
Phone: 0120-003-462 (Help Desk)
: 0120-151-462 (上記以外)
Fax: 03-5978-4090
E-mail: support@jaici.or.jp (Help Desk)
customer@jaici.or.jp (上記以外)
Internet: www.jaici.or.jp

北アメリカ
STN コロンバス

CAS
P.O. Box 3012
Columbus, Ohio 43210-0012 U.S.A
CAS Customer Care:
Phone: 800-753-4227 (North America)
614-447-3700 (worldwide)
Fax: 614-447-3751
E-mail: help@cas.org
Internet: www.cas.org

SEARCH および DISPLAY フィールド

中間一致および後方一致検索可能なフィールドはアスタリスク (*) で示してあります。

フィールド	SEARCH コード	SEARCH 例	DISPLAY コード
基本索引 * 標題 (/TI) 抄録 (/AB) 生物名 (/ORGN) 分子タイプ (/MTY) 配列の説明 (/DESC) (以上からの切り出し語)	なし または /BI	S PLANT ?GENE? AND RNA S DNA DEPENDENT	TI, AB, ORGN, MTY, DESC
抄録 出願国 (WIPO コードとテキスト) 出願日 ¹⁾ レコード番号 出願番号 ²⁾ 出願年 ¹⁾ 配列の説明 資料種類 (コードとテキスト) 入力日 ^{1) 4)} 代表クレーム 特徴表 * ファイルセグメント (コードとテキスト) 発明者 分子タイプ 注記 生物名 特許出願人 ³⁾ 特許発行国 (WIPO コードとテキスト) 特許発行日 ¹⁾ 特許種別 特許番号 ²⁾ 特許中の配列情報の記載位置 優先権主張国 優先権主張日 優先権出願番号 優先権主張年 特許期間調整の日数 特許発行年 ¹⁾ 関連出願国 (WIPO コードとテキスト) 関連出願日 ¹⁾ 関連出願番号 ²⁾ 関連出願年 ¹⁾	/AB /AC /AD /AN /AP /AY /DESC /DT または /TC /ED /ECLM または /MCLM /FEAT /FS /IN または /AU /MTY /NTE /ORGN /PA または /CS /PC /PD /PK /PN または /PATS /PSL /PRC /PRD /PRAI または /PRN /PRY /PTA /PY /RLC /RLD /RLN または /RLI /RLY	S ANAPHYLATOXIN/AB S US/AC AND L1 S JUN 2000/AD S 4305872.2/AN S US1980-162082/AP S 2000/AY S GHRH/DESC S PATENT/DT AND L1 S P/TC AND L1 S ED>=JAN 2007 AND L1 S NANOCRYSTAL/ECLM S ?COMBINAT?/FEAT S NUCLEIC/FS AND L1 S PROTEIN/FS AND L1 S MANDREKAR MICHELLE?/IN S RNA/MTY S DISCLAIMER/NTE S CRICETULUS AUREUS/ORGN S AMGEN/PA S US/PC AND L1 S PD=4 MAR 2003 S USB2/PK S US6686189/PN S CLAIM/PSL S JP/PRC S 20070202/PRD S US2002-500660/PRN S 2006/PRY S 100-150/PTA S 2002-2003/PY AND L1 S WO/RLC AND L1 S 20010313/RLD S W02000-EP4786/RLN S 2000-2001/RLY	AB AI (AP) AI (AP) AN AI (AP) AI (AP) AI (AP) DESC DT (TC) ED ECLM FEAT FS IN (AU) MTY NTE ORGN PA (CS) PI (PN) PI (PN) PI (PN) PI (PN) PSL PRAI PRAI PRAI PRAI PTA PI (PN) RLI (RLN) RLI (RLN) RLI (RLN) RLI (RLN)

(続く)

SEARCH および DISPLAY フィールド (続き)

フィールド	SEARCH コード	SEARCH 例	DISPLAY コード
配列の数 ¹⁾	/SEQC	S 3-4/SEQC	SEQC
配列同定番号 ¹⁾	/SEQN	S 337/SEQN	SEQN
配列長 ¹⁾	/SQL	S 150-175/SQL AND L1	SQL
配列の収録源	/SSO	S NCBI/SSO AND L1	SSO
標題	/TI	S HYBRIDIZATION ASSAY##/TI	TI
更新日 ^{1), 4)}	/UP	S UP=20100730	UP
失効日	/XPD	S 20120701-20120731/XPD	XPD
失効年	/XPY	S 2010-2012/XPY	XPY

1) 数値演算子あるいは範囲指定による検索が可能な数値検索フィールドです。

2) STN 形式またはダウエント形式のどちらでも利用できます。

3) このフィールドでは (S) 演算子はスペースで代用できます。

4) 2010 年 7 月 30 日以降利用可能です。

スーパー検索フィールド

必要な情報が含まれる一つまたは複数のフィールドを検索するときは、スーパー検索フィールドコードを利用します。スーパー検索フィールドを利用すると、クロスファイルおよびマルチファイル検索が簡単に実行できます。スーパー検索フィールドで EXPAND は利用できませんので、個々のフィールドコードで EXPAND してください。

フィールド	スーパー検索 フィールドコード	検索される フィールド	SEARCH 例	DISPLAY コード
出願番号グループ ¹⁾	/APPS	/AP, /RLN, /PRN	S US2001-809524/APPS	AI, RLI, PRAI

1) STN 形式またはダウエント形式のどちらでも利用できます。

配列ホモロジー検索 (BLAST/GETSIM)

BLAST (R) および GETSIM パッケージプログラムは、タンパク質および核酸配列データを相同性 (類似性) で検索することが可能です。BLAST は、National Library of Medicine (NLM) の National Center for Biotechnology Information (NCBI) より提供されています。GETSIM は、FIZ Karlsruhe GmbH より提供されており、FASTA アルゴリズムに基づいています。

BLAST および GETSIM で検索する際は、次の検索フィールドコードを使用します。

/SQP — タンパク質の配列検索 (デフォルト)

/SQN — 核酸の配列検索

/TSQN — 核酸配列から翻訳されたタンパク質配列のデータベースを検索

BLAST および GETSIM 検索は、バッチ検索やアラートの登録が可能です。検索フィールドに /SQN または /TSQN を使用する場合は、単独 (SIN)、相補鎖 (COM)、双方 (BOTH) を指定することができます。これらのオプションは /SQN COM のように TSQN あるいは SQN の検索フィールドコードと共に使用します。BLAST では BOTH、GETSIM では SIN がデフォルトです。

核酸およびタンパク質配列はさまざまな方法で相同性検索ができます。質問式には、QUERY コマンドを使用して作成・保存した L 番号を使うことができます。また RUN BLAST/GETSIM コマンドと同じ行に直接入力したり、UPLOAD コマンドでアップロードしたテキストファイルを利用することも可能です。検索を実行すると、質問式と得られた結果の配列の相同性を示すグラフが作成されます。Y 軸は相同性の数値を示し、X 軸は対応する回答数を表わします。すべての回答、もしくは任意の相同性以上の回答を選んで回答セットを作成することができます。最初に作成された回答セットでは、回答がレコード番号の降順に並んでいます。相同性の高い順に回答を並べ替える際は、SOR SCORE D と L 番号を入力してください。

ALIGN 表示形式を使用して、BLAST および GETSIM の検索結果を質問式の配列と対応づけて表示できます。上の列は質問式、下の列は検索でヒットした配列を示します。BLAST の ALIGN 表示形式は、NCBI の対応表示の標準規則に従います。GETSIM の ALIGN 表示形式では、同一のアミノ酸/核酸残基を 2 ドット、化学的に“等価”のマッチを 1 ドット、マッチがない場合はスペース (ドットなし)

で示されます。質問式または回答配列に挿入されたギャップは、下線で示されます。USGENE ファイルでは、SEQO 表示形式を利用するとオリジナルの配列情報を表示することができます。SEQO 表示形式では、特許出願人によって記述されたオリジナルの核酸配列に相当するタンパク質配列が表示されます。

BLAST/GETSIMの検索タイプ

RUN GETSIM/BLAST BATCH 検索では、BATCH は次のようにコマンドに続けて入力します：

```
RUN GETSIM L1 /TSQN BOTH BATCH
RUN BLAST L1 /TSQN BOTH BATCH
```

RUN GETSIM/BLAST アラート検索では、ALERT は次のようにコマンドに続けて入力します：

```
RUN GETSIM L1 /TSQN BOTH ALERT
RUN BLAST L1 /SQN COM ALERT
```

アラート検索はデータ更新毎、すなわち週一回実行されます。

記 述	SEARCH コード	SEARCH 例 ¹⁾
タンパク質ホモロジー	/SQP	RUN BLAST L1/SQP
核酸ホモロジー	/SQN	RUN BLAST L1/SQN
単独鎖 ²⁾		RUN GETSIM L1/SQN SIN
相補鎖		RUN GETSIM L1/SQN COM
双方 ³⁾		RUN BLAST L1/SQN BOTH
翻訳タンパク質ホモロジー	/TSQN	RUN BLAST L1/TSQN
単独鎖 ²⁾		RUN GETSIM L1/TSQN SIN
相補鎖		RUN BLAST L1/TSQN COM
双方 ³⁾		RUN BLAST L1/TSQN BOTH

1) L1は UPLOAD または QUERY を使用して生成された質問式配列です。

2) GETSIM ではデフォルト

3) BLAST ではデフォルト

BLAST の上級ユーザー向けオプション

BLAST の熟練ユーザーのために、多数のオプションが可能となっています。これらのパラメータの変更は検索回答出力に大きな効果を発揮します。FIZ Karlsruhe はこれらの機能パラメータのどれについても利用する前に、ユーザーが NCBI の文献を読み、理解されることを強く推奨します。詳細情報は次をご覧ください：

<http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/information3.html>

上級ユーザーオプションはハイフン + 1 文字コード + スペース + 値の書式で入力します：

```
RUN BLAST L1/SQN -E 0.1.
```

オプション	スイッチ	値
フィルタ	-f	T (初期値), F, C T がセットされた場合、タンパク質に対しては SEG, 核酸に対しては DUST フィルタが有効となる。 C 記号は“COILED COILED”フィルタ。
期待値	-e	浮動小数点数 (初期値10)
Word Size	-w	核酸：7-23, 初期値は 11 タンパク質：2-3, 初期値は 3
ストランド	-s	1 (SIN), 2 (COM), または 3 (BOTH) (初期値)
マトリックス	-m	BLOSUM62 (初期値), BLOSUM80, BLOSUM45, PAM30 または PAM70
ギャップ反則	-g	核酸：5 (初期値) タンパク質：11 (初期値)
ギャップ拡張	-x	核酸：2 (初期値) タンパク質：1 (初期値)
核酸のミスマッチに対する減点	-q	-3 (初期値)
核酸のマッチに対する加点	-r	1 (初期値)

BLAST マトリックスの設定

特定のマトリックスに対して、可能性のあるギャップとギャップ拡張値の組合せは制限されていません。各マトリックスに対して可能な設定値と初期値は以下の表のとおりです。これ以外の組合せはシステムによって拒絶され、警告メッセージが表示されます。

マトリックス	ギャップ	ギャップ拡張値
BLOSUM62	9	2
	8	2
	7	2
	12	1
	11	1 (デフォルト)
	10	1
BLOSUM80	8	2
	7	2
	6	2
	11	1
	10	1 (デフォルト)
	9	1
BLOSUM45	13	3
	11	3
	12	3
	9	3
	15	2 (デフォルト)
	14	2
	13	2
	12	2
	19	1
	18	1
	17	1
16	1	
PAM30	7	2
	6	2
	5	2
	10	1
	8	1
	9	1 (デフォルト)
PAM70	8	2
	7	2
	6	2
	11	1
	10	1 (デフォルト)
	9	1

検索例 (BLAST ホモロジー検索)

```
=> UPLOAD
IS THIS DATA A QUERY, OR FOR A RUN PACKAGE? Q/R/(END):R
ENTER NAME OF RUN PACKAGE, END OR (?):BLAST
START LOCAL KERMIT TRANSMIT PROCESS
```

```
UPLOAD SUCCESSFULLY COMPLETED
L1 GENERATED
```

```
=> D L1 LQUE
```

```
L1 ANSWER 1 USGENE COPYRIGHT 2010 SEQUENCEBASE CORP on STN
LQUE gtatatataa cgtgatgagc gtacgggtgc ggagacgcac
cggagcgcctcgcgccagccgcgctccaagc ccctgaggtt tccggggacc acaatgaacaagttgctgtg
ctgcgcgcctcgtgtttctgg acatctccat taagtggaccaccaggaaa cgtttctctcc
aaagtaccttcattatgacg aagaaaacctctcatcagctg ttgtgtgaca aatgtctctcc
tggtacctacctaanaaacaactgtacagc aaagtggaag accgtgtgcg
ccccttgccctgaccactactacacagaca gctggcacac cagtgacgag
tgtctatactgcagccccgtgtgcaaggag ctgcagtacg tcaagcagga
gtgcaatgcaccccacaaccgctgtgca atgcaaggaa gggcgcctacc
ttgagatagagttctgcttgaacatagga gctgccctcc tggatttggga
gtggtgcaagctggaaccccagagcgaat acagtttgca aaagatgtcc
agatgggttctctcaaatgagacgtcatc taaagcacc tgtagaaaac
acacaaattgcagtgcttttggctcctgc taactcagaa aggaaatgca
acacacgacaacatatgttccggaaacagt gaatcaactc aaaaatgtgg
aatagatgttaccctgtgtgaggaggcatt cticagggtt gctgttctta
caaagtttacgcctaactggcttagtgtct tggtagaaa tttgcctggc
accaaagtaaacgcagagagtagagagg ataaaaacgc aacacagctc
acaagaacagactttccagctgctgaagtt atggaacat caaaaacaag
cccaagatatagtcaagaagatcatccaag atattgacct ctgtgaaaac
agcgtgcagcggcacattggacatgctaac ctcaccttcg agcagcttcg
tagcttgatggaagcttaccgggaaagaa agtgggagca gaagacattg
aaaaaacaataaaggcatgcaaaccagtg accagatcct gaagctgctc
agttttgtggcgaataaaaaatggcgaccaa gacaccttga agggccta
gcacgcactaaagcactcaaagacgtacca ctttccaaa actgtcactc
agagtctaaagaagaccatcaggttccttc acagcttcac aatgtacaaa
ttgtatcagaagttattttagaaatgata ggtaaccagg tccaatcagt
aaaaataagctgcttataactggaaatggc cattgagctg tttctcaca attggcgaga
tcccatggatgataa
```

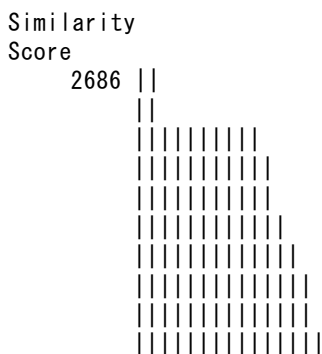
```
=> RUN BLAST L1/SQN
```

```
BLAST Version 2.2
```

The BLAST software is used herein with permission of the National Center for Biotechnology Information (NCBI) of the National Library of Medicine (NLM). See also, Altschul, :

1814 ANSWERS FOUND BELOW EXPECTATION VALUE OF 10.0

```
QUERY SELF SCORE VALUE IS 2686
BEST ANSWER SCORE VALUE IS 2686
```



検索例 (BLAST ホモロジー検索) (続き)

Strand = Plus / Plus

Query: 1 gtatatataacgtgatgagcgtacgggtgcggagacgcaccggagcgtcgcccagccgc

Sbjct: 1 gtatatataacgtgatgagcgtacgggtgcggagacgcaccggagcgtcgcccagccgc

Query: 61 cgctccaagcccctgaggtttccggggaccacaatgaacaagttgctgtgctgcgcgctc

Sbjct: 61 cgctccaagcccctgaggtttccggggaccacaatgaacaagttgctgtgctgcgcgctc

:

Query: 121 gtgtttctggacatctccattaagtggaccaccaggaaacgtttcctccaaagtacctt

Sbjct: 121 gtgtttctggacatctccattaagtggaccaccaggaaacgtttcctccaaagtacctt

Query: 1261 ggtaaccagggtccaatcagtaaaaataagctgcttataactggaaatggcattgagctg

Sbjct: 1261 ggtaaccagggtccaatcagtaaaaataagctgcttataactggaaatggcattgagctg

Query: 1321 tttcctcacaattggcgagatcccatggatgataa 1355

|||||

Sbjct: 1321 tttcctcacaattggcgagatcccatggatgataa 1355

L3 ANSWER 200 OF 396 USGENE COPYRIGHT 2010 SEQUENCEBASE CORP on STN

TI Albumin fusion proteins (Patent)

DESC Homo sapiens DNA; sequence 1215 of 2162

MTY DNA

SQL 1203

SCORE 2377 88% of query self score 2686

BLASTALIGN

Query = 1355 letters

Length = 1203

Score = 2377 bits (1199), Expect = 0.0

Identities = 1202/1203 (99%)

Strand = Plus / Plus

Query: 94 atgaacaagttgctgtgctgcgcgctcgtgtttctggacatctccattaagtggaccacc

Sbjct: 1 atgaacaagttgctgtgctgcgcgctcgtgtttctggacatctccattaagtggaccacc

Query: 154 caggaaacgtttcctccaaagtaccttcattatgacgaagaaacctctcatcagctgttg

Sbjct: 61 caggaaacgtttcctccaaagtaccttcattatgacgaagaaacctctcatcagctgttg

Query: 214 tgtgacaaatgtcctcctggtacctacctaatacaaacactgtacagcaaagtggaagacc

Sbjct: 121 tgtgacaaatgtcctcctggtacctacctaatacaaacactgtacagcaaagtggaagacc

:

Query: 1174 gtcactcagagtctaaagaagaccatcaggttccttcacagcttcacaatgtacaaattg

Sbjct: 1081 gtcactcagagtctaaagaagaccatcaggttccttcacagcttcacaatgtacaaattg

Query: 1234 tatcagaagttatTTTTAGAAATGATAGGTAACCAGGTCCAATCAGTAAAAATAAGCTGC

Sbjct: 1141 tatcagaagttatTTTTAGAAATGATAGGTAACCAGGTCCAATCAGTAAAAATAAGCTGC

Query: 1294 tta 1296

|||

Sbjct: 1201 tta 1203

検索例 (BLAST ホモロジー検索) (続き)

=> S L3 AND ((TNF## OR TUMOR NECROSIS FACTOR?) (2A) (RECEPTOR? OR BINDING PROTEIN?))/BI,ECLM AND AY>1990 AND GRANTED/SSO

L5 12 L4 AND ((TNF## OR TUMOR NECROSIS FACTOR?) (2A) (RECEPTOR? OR BINDING PROTEIN?))/BI,ECLM AND AY>1990 AND GRANTED/SSO

=> D BIB AB ECLM ALIGN 1

L5 ANSWER 1 OF 12 USGENE COPYRIGHT 2010 SEQUENCEBASE CORP on STN
 AN 7632922.124 DNA USGENE
 TI Osteoprotegerin (Patent)
 IN Boyle William J. (Moorpark, CA); Lacey David L. (Thousand Oaks, CA);
 Calzone Frank J. (Westlake Village, CA); Chang Ming-Shi (Newbury Park,
 CA)
 PA Amgen Inc (Thousand Oaks CA)
 PI US 7632922 B1 20091215
 AI US 2000-718725 20001122
 PRAI US 2000-718725 20001122
 US 1998-132985 19980812
 US 1996-771777 19961220
 US 1996-706945 19960903
 US 1995-577788 19951222
 XPD 20151222 (calculated)
 NTE Subject to any Disclaimer, the term of this patent is extended or
 adjusted under 35 USC 154(b) by 1275 days.
 PSL SEQ ID NO 124
 DESC Homo sapiens DNA; sequence 124 of 179
 DT Patent
 AB The present invention discloses a novel secreted polypeptide, termed
 osteoprotegerin, which is a member of the tumor necrosis factor
 receptor superfamily and is involved in the regulation of bone
 metabolism. Also disclosed are nucleic acids encoding osteoprotegerin,
 polypeptides, recombinant vectors and host cells for expression,
 antibodies which bind OPG, and pharmaceutical compositions. The
 polypeptides are used to treat bone diseases characterized by increased
 resorption such as osteoporosis.
 ECLM US7632922 B1: 1. An isolated antibody or fragment thereof which
 specifically binds to an OPG polypeptide consisting of residues 22-401
 inclusive of SEQ ID NO:121.

BLASTALIGN

Query = 1355 letters
 Length = 1355
 Score = 2686 bits (1355), Expect = 0.0
 Identities = 1355/1355 (100%)
 Strand = Plus / Plus

Query: 1 gtatatataacgtgatgagcgtacgggtgctggagacgcaccggagcgtcgcccagccgc
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 1 gtatatataacgtgatgagcgtacgggtgctggagacgcaccggagcgtcgcccagccgc

Query: 61 cgctccaagcccctgaggtttcggggaccacaatgaacaagttgctgtgctgctgcgcgctc
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 61 cgctccaagcccctgaggtttcggggaccacaatgaacaagttgctgtgctgctgcgcgctc

Query: 121 gtgtttctggacatctccattaagtggaccaccaggaacgtttcctccaaagtacctt
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 121 gtgtttctggacatctccattaagtggaccaccaggaacgtttcctccaaagtacctt

:

Query: 1321 ttctctcacaattggcgagatcccatggatgataa 1355
 ||||||||||||||||||||||||||||||||
 Sbjct: 1321 ttctctcacaattggcgagatcccatggatgataa 1355

配列検索 (GETSEQ)

GETSEQ で使う質問式には、QUERY コマンドで作成した L 番号を利用したり、配列検索可能な STN の他のファイル、例えば REGISTRY や DGENE ファイルで配列検索した L 番号を質問式として利用することも可能です。また、RUN GETSEQ の後に直接入力することもできます。

配列検索コード

質問式	SEARCH 例
一般アミノ酸の 1 文字コード ¹⁾	QUE LAGLL/SQSP
一般アミノ酸の 3 文字コード ¹⁾ コードまたはコードの列を一重引用符で囲む。 各コード間にハイフンを入力する。	QUE 'MET-GLY-LEU-TRP-TRP-ARG'/SQSFP
核酸の 1 文字コード ²⁾	QUE ACTACCTTCAAATACTAC/SQEN QUE ACTACCTTCAAATACTAC/SQSN

1) 矢印プロンプト (=>) の後で、HELP AAC と入力すると、一般アミノ酸の 1 および 3 文字コード表が表示されます。

2) 矢印プロンプト (=>) の後で、HELP NUC と入力すると、核酸のコード表が表示されます。

配列検索の種類

核酸およびタンパク質配列データは、SEQ フィールド中に1文字コードで、またタンパク質のみ SEQ3 フィールド中に 3 文字コードで表示されます。

RUN GETSEQ BATCH 検索では、BATCH は次のようにコマンドに続けて入力します：

RUN GETSEQ コード/SQSN BATCH

種類	定義	SEARCH コード	SEARCH 例
タンパク質, 完全配列	質問式に一致する配列を検索。 質問式は完全に定義されていなければならない。	/SQEP	QUE AKRSSKM/SQEP QUE 'ALA-LYS-SER-LYS'/SQEP
タンパク質, 完全配列ファミリー	質問式に一致する配列および質問式中のアミノ酸とファミリーに相当するアミノ酸の配列 ¹⁾	/SQEFP	QUE SETLR/SQEFP QUE 'SER-GLU-THR-LEU'/SQEFP
タンパク質, 部分配列	完全に一致する配列と質問式の配列を含む配列を検索。特殊記号が利用できる。	/SQSP	QUE SKGYF/SQSP QUE 'SER-LYS-GLN-PHE'/SQSP
タンパク質, 部分配列ファミリー	完全に一致する配列と質問式のアミノ酸とファミリーに相当するアミノ酸の配列が含まれる。特殊記号が利用できる。 ¹⁾	/SQSFP	QUE SYVVE/SQSFP QUE 'SER-TYR-VAL-GLU'/SQSFP
核酸, 完全配列	質問式に完全に一致する配列を検索。 非確定コードが利用できる。	/SQEN	QUE ACTACCTTCAAATACTAC/SQEN
核酸, 部分配列	完全に一致する配列と質問式の配列を含む配列を検索。 非確定コードおよび特殊記号が利用できる。	/SQSN	QUE ACTACCTTCAAATACTAC/SQSN

1) タンパク質の等価ファミリーは以下の通りです。

P, A, G, S, T	(弱疎水性, 中性)
Q, N, E, D, B, Z	(親水性, 酸アミン)
H, K, R	(親水性, 塩基性)
F, Y, W	(疎水性, 芳香族)
L, I, V, M	(疎水性)
C	(架橋)

部分配列検索の特殊記号 (/SQSP, /SQSFP, および /SQSN)

記号	機能	SEARCH 例
[]	代替残基の特定	QUE LGP[VL]/SQSP
[-]	特定の代替残基の除外	QUE LGP[VAL 'LEU' 'LYS']/SQSP QUE LGP[-H]/SQSP QUE LGP[-HIS']/SQSP QUE LGP[-HL]/SQSP
{m}	直前の配列, 配列質問式, または L 番号を m 回繰り返す	QUE (FL){2}/SQSP QUE L4{2}/SQSP QUE (CTG){2}/SQSN QUE TAA(TAAA){2}/SQSN
{m, u} または {m-u}	直前の配列, 配列質問式, または L 番号を m から n 回繰り返す	QUE GG(FL){1, 2}/SQSP QUE L3{1, 3}/SQSP QUE (CTG){1, 3}/SQSN
? または {0, 1} または {0-1}	直前の配列, 配列質問式, または L 番号をゼロまたは 1 回繰り返す	QUE FLRRI(RP)?K/SQSP QUE FLRRI(RP){0, 1}K/SQSP QUE L1{0, 1}NN/SQSP QUE CAT(CGA){0, 1}GGAC/SQSN
* または {0, } または {0-}	直前の配列, 配列質問式, または L 番号をゼロ回以上繰り返す	QUE KLK(WD){0, }N/SQSP QUE KLK(WD)*N/SQSP QUE L1{0-}NN/SQSP QUE L1{0-}NN/SQSP QUE CAT(CTG){0, }TATT/SQSN
+ または {1, } または {1-}	直前の配列, 配列質問式, または L 番号を 1 回以上繰り返す	QUE KLK(DLE){1, }/SQSP QUE KLK(DLE)+/SQSP QUE L2{1-}/SQSP QUE L2{1, }/SQSP QUE CAT(CTG){1, }TATT/SQSN
&	配列表記または質問式 (L 番号) を結合する	QUE L1&L3/SQSFP QUE L2&L5{1, 3}/SQSP

1) 部分配列質問式のさまざまな指定法に関する詳細は, USGENE ファイルで矢印プロンプト(=>)の後に HELP SQQ と入力すると表示できます。

上記の他に, キャレット (^), 垂直バー (|) も利用できます。

キャレット (^) は, 配列中の最初または最後の部分を指定して検索する場合に用います。

垂直バー (|) は, 代替記号です。例えば, 代替配列質問式に利用します。

部分配列検索のギャップ記号 (/SQSP, /SQSFP, および /SQSN)

記号	機能	SEARCH 例
.	1 残基のギャップ	QUE SY.RPG/SQSP QUE SY..RPG/SQSP QUE AAG...TGC/SQSN
.{m} または {m.}	m 残基のギャップ	QUE SY.{2}RPG/SQSP QUE SY[2.]RPG/SQSP
.{m, u} または .{m-u}	m から u 残基のギャップ	QUE GFF.{2, 10}LSS/SQSP QUE GFF.{2-10}LSS/SQSP QUE AAG.{2, 5}TGC/SQSN
: または .? または .{0, 1} または .{0-1}	ゼロまたは 1 残基のギャップ	QUE AGA:SRI/SQSFP QUE AGA.?SRI/SQSFP QUE AGA.{0, 1}SRI/SQSFP QUE AGA.{0-1}SRI/SQSFP
.* または .{0, } または .{0-}	ゼロ残基以上のギャップ	QUE HLC.*TYG/SQSP QUE HLC.{0, }TYG/SQSP QUE HLC.{0-}TYG/SQSP QUE AAGGCAGATG.*GCAA/SQSN
.+ または .{1, } または .{1-}	1 残基以上のギャップ	QUE SY.+TYG/SQSP QUE SY.{1, }TYG/SQSP QUE SY.{1-}TYG/SQSP QUE TCCTG.+GTGG/SQSN

DISPLAY および PRINT 形式

回答のディスプレイとプリントには下記の表示形式を自由に組み合わせることができます。複数のコードを指定するときは、"D L1 1-5 TI PA KW" のようにスペースやカンマで区切ってください。フィールドは指定された順序で表示されます。

DED, ED, FEAT, UP を除き、すべての検索フィールドでハイライト機能が利用できます。HIT, KWIC および OCC の各形式を使うためには、検索時にハイライト機能は ON であることが必要です。

形 式	英 語 名	内 容	入 力 例
AB	Abstract	抄録	D TI AB
AI (AP) ¹⁾	Application Information	出願情報	D AI
AIO ²⁾	Application Information, Original	オリジナルの出願情報	D AIO
AN	Accession Number	レコード番号	D AN
CLM	Claims	クレーム	D CLM
DESC	Description	配列の説明	D DESC
DT (TC)	Document Type	資料種類	D TI DT
ECLM (MCLM)	Ezemplary Claim Date	代表クレーム	D BIB AB ECLM
ED ²⁾	Entry Date	入力日	D ED
FEAT	Feature Table	特徴表	D FEAT
FS ²⁾	File Segment	ファイルセグメント	D FS
IDENT ^{2) 4)}	Percent Identity	同一性パーセント	D IDENT
IN (AU)	Inventor	発明者	D IN PA
MTY	Molecule Type	分子タイプ	D MTY
NTE	Note	注記	D NTE
ORGN	Organism Name	生物名	D ORGN
PA (CS)	Patent Assignee	特許出願人	D PA
PI ¹⁾	Patent Information	特許情報	D PI
(PN, PATS)			
PRAI (PRN) ¹⁾	Priority Information	優先権情報	D PRAI
PSL	Patent Sequence Location	特許中の記載位置	D PSL
PTA ²⁾	Patent Term Adjustment (number of days)	特許期間調整の日数	D PTA
RLI (RLN) ¹⁾	Related Application Number	関連出願情報	D RLI
RLIO ²⁾	Related Application Information, Original	オリジナルの関連出願情報	D RLIO
SCORE ^{2) 3)}	Similarity Score	相同性スコア	D SCORE
SEQ	Sequence (1-letter codes)	配列 (1文字コード)	D SEQ
SEQ3	Sequence (3-letter codes)	配列 (3文字コード)	D SEQ3
SEQC ²⁾	Sequence Number Count	配列の数	D SEQC
SEQN ²⁾	Sequence Identity Number	配列同定番号	D SEQN
SEQO	Original Sequence (alignment of nucleotide sequence and peptide sequence it express when given)	オリジナルの配列	D SEQO
SQL	Sequence Length	配列長	D SQL
SSO	Sequence Source	配列の収録源	D SSO
TI	Title	標題	D TI
UP ²⁾	Update Date	更新日	D UP
XPD	Expiration Date	失効日	D XPD

DISPLAY および PRINT 形式 (続き)

ALIGN ³⁾	質問式と相同性検索の結果の回答を対応させて表示 (RUN GETSIM または RUN BLAST)	D ALIGN
ALL ¹⁾	AN, MTY, TI, IN, PA, PI, AI, RLI, PRAI, XPD, NTE, PSL, DESC, AB, CLM, SSO, ORGN, SQL, SEQ, FEAT	D ALL
APPS ¹⁾	AI, PRAI, RLI	D APPS
BIB (STD) ¹⁾	AN, MTY, TI, IN, PA, PI, AI, RLI, PRAI, XPD, NTE, PSL, DESC, DT (デフォルト)	D BIB
BRIEF ¹⁾	AN, MTY, TI, IN, PA, PI, AI, RLI, PRAI, XPD, NTE, PSL, DESC, DT, AB, ECLM, SSO, ORGN, SQL, SEQ, FEAT	D BRIEF
CFAM ¹⁾	特許ファミリー中のすべての公報番号と発行日を表形式で表示 (INPADOC の情報)	D CFAM
FAM ¹⁾	すべての公報番号に対応する出願番号と出願日, 優先権出願番号と優 先日の情報を表形式で表示	D FAM
FASTA	AN, SEQN, PN, SEQ (FASTA 形式)	D FASTA
FASTA2	SEQ (FASTA 形式)	D FASTA2
IALL ¹⁾	フィールド名付きインデント型 ALL 形式	D IALL
IBIB ¹⁾	フィールド名付きインデント型 BIB 形式	D IBIB
IBRIEF ¹⁾	フィールド名付きインデント型 BRIEF 形式	D IBRIEF
LS ¹⁾	法的状況 (INPADOC の情報)	D LS
LS2 ¹⁾	法的状況 (INPADOC の情報), 各データをフィールドごとに分けて表示	D LS2
SCAN ⁵⁾	TI (回答番号なしのランダム表示)	D SCAN
SQIDE	AN, SQL, SEQ, FEAT	D SQIDE
SQ3IDE	AN, SQL, SEQ3, FEAT	D SQ3IDE
TRIAL (TRI, SAM)	AN, DESC, MTY, SQL	D TRI
HIT	ヒットタームを含むフィールド	D HIT
KWIC	ヒットタームの前後 50 語を表示 (KeyWord-In-Context)	D KWIC
OCC	ヒットタームの出現頻度をフィールドごとに表示	D OCC

- 1) デフォルトでは, 特許番号, 出願番号, および優先権出願番号は STN 形式で表示されます. ダウエント形式に変更するときには, 矢印プロンプト (=>) の後で SET PATENT DERWENT と入力します. STN 形式に戻す場合は, SET PATENT STN と入力します.
- 2) カスタム形式のみの表示です.
- 3) RUN GETSIM あるいは RUN BLAST の回答である必要があります. 詳しくは, 配列ホモロジー検索の項を参照してください.
- 4) RUN BLAST の回答である必要があります. 詳しくは, 配列ホモロジー検索の項を参照してください.
- 5) この表示形式のオンライン・ディスプレイ料金は無料です. SCAN は, コマンドに続けて入力します. 例:
D SCAN または DISPLAY SCAN

SELECT, ANALYZE および SORT フィールド

SELECT コマンドは、回答セットの指定したフィールドから抽出した語句に E 番号を付与します。
 ANALYZE コマンドは、回答セットの指定したフィールドから抽出した語句に L 番号を付与します。
 SORT コマンドは、検索結果を指定したフィールドのアルファベット順または数値順に並べ替えます。
 (該当項目は Y, 該当しないものは N で表示されています)

フィールド	フィールドコード	ANALYZE/SELECT ¹⁾	SORT
抄録	AB	Y	Y
出願国	AC	Y	N
出願日	AD	Y	N
出願情報	AI (AP)	Y	Y
オリジナルの出願情報	AIO	Y	N
レコード番号	AN	Y	Y
出願年	AY	Y	N
クレーム	CLM	Y	Y
配列の説明	DESC	Y	N
資料種類	DT (TC)	Y	Y
代表クレーム	ECLM (MCLM)	Y	Y
入力日	ED	Y	Y
特徴表	FEAT	Y	N
ファイルセグメント	FS	Y	Y
同一性パーセント	IDENT	Y ²⁾	Y ²⁾
発明者 (著者名)	IN (AU)	Y	Y
分子タイプ	MTY	Y	Y
注記	NTE	Y	N
生物名	ORGN	Y	Y
特許出願人	PA (CS)	Y	Y
特許発行国	PC	Y	Y
特許発行日	PD	Y	Y
特許種別	PK	Y	Y
特許番号	PN (PI)	Y	Y
優先権主張国	PRC	Y	N
優先権主張日	PRD	Y	N
優先権出願番号	PRN (PRAI)	Y	N
優先権主張年	PRY	Y	N
特許中の記載位置	PSL	Y	N
特許期間調整の日数	PTA	Y	Y
特許発行年	PY	Y	Y
関連出願国	RLC	Y	N
関連特許出願日	RLD	Y	N
オリジナルの関連出願情報	RLIO	Y	N
関連出願番号	RLN (RLI)	Y	Y
関連出願年	RLY	Y	N
相同性スコア	SCORE ⁴⁾	N	Y ²⁾
配列 (1 文字コード)	SEQ	Y ³⁾	N
配列 (3 文字コード)	SEQ3	Y ³⁾	N
配列の数	SEQC	Y	N
配列同定番号	SEQN	Y	Y
配列長	SQL	Y	Y
配列の収録源	SSO	Y	Y
標題	TI	Y (デフォルト)	Y
失効日	XPB	Y	N
失効年	XPY	Y	N
更新日	UP	Y	Y

1) ヒットタームだけを抽出させるには、HIT を使います。例：SEL HIT IN

2) BLAST により生成された L 番号を指定します。

3) SELECT したタームに /SQSP が付与されます。

4) GETSIM または BLAST により生成された L 番号を指定します。

サンプルレコード

ALL形式

AN 7553483.1 Protein USGENE
 TI Chemokine mutants acting as chemokine antagonists (Patent)
 IN Proudfoot Amanda (Chens sur Lemman, FR); Kosco-Vilbois Marie (Minzier, FR); Shaw Jeffrey (Geneva, CH)
 PA Merck Serono SA (Coinsins Vaud CH)
 PI US 7553483 B2 20090630
 US 20050220757 A1 20051006
 WO 2003051921 A 20030626
 AI US 2002-499100 20021216
 RLI WO 2002-EP14325 20021216
 PRAI EP 2001-761 20011217
 XPD 20221216 (calculated)
 NTE Subject to any Disclaimer, the term of this patent is extended or adjusted under 35 USC 154(b) by 433 days.
 PSL Claim 1; SEQ ID NO 1
 DESC Artificial protein; Mutant chemokine; sequence 1 of 22
 DT Patent
 AB Mutants of specific CC-chemokines containing a non-conservative substitution in a conserved consensus sequence act as CC-chemokine antagonists, and can be effectively used in the treatment of autoimmune and inflammatory diseases, cancers, and viral or bacterial infections. Particularly preferred are the RANTES/CCL5 mutants having the amino acid sequence of SEQ ID NO: 1, 2, 3, or 4.
 CLM US7553483 B2: 1. An isolated mutant of CCL5 (RANTES) comprising SEQ ID NO: 1, 4, 5, 6 or 7, wherein said mutant antagonizes in vivo cellular recruitment caused by CCL5.

 2. An isolated polypeptide comprising a heterologous polypeptide sequence and a polypeptide selected from the group consisting of SEQ ID NO: 1, 4, 5, 6 and 7.

 3. The polypeptide according to claim 2, wherein said heterologous polypeptide sequence comprises an amino acid sequence selected from one or more of the following protein sequences: extracellular domains of membrane-bound protein, immunoglobulin constant regions, multimerization domains, extracellular proteins, signal peptide-containing proteins, or export signal-containing proteins.
 :
 20. The pharmaceutical composition according to claim 4, wherein said mutant comprises SEQ ID NO: 7.

 SSO PROTEIN; USPTO; GRANTED
 ORGN Artificial sequence
 SQL 68
 SEQ
 1 spysdttpc cfayiarplp rahikeyfyt snkcsnpavv fvtrknrqvc
 51 anpekkwvre yinslems

FEATURE TABLE:

Key	Location
	mutant chemokine

FASTA 表示形式

FASTA:

```
>USGENE|20100017904.32958|Protein|sequence 32958 from US20100017904
mgevvatweateggagvkgpvvvtgasgflgswlvmkllqagytratrdrpanvvtkplldlpgater
lslwkadladesfddairgctgvfhvatpmdfeskdpenevikptvegmmssimrackeagtvrriivfts
sagtvnieerqrpyvdqdnwsdvdqcrvkmrgwmyfvskslaekaamayaahgldfisiiptlvvgpf
lsagmppsllitalalvtgneahysilkqvqfvhlddldahlfifehpaaagryvcsshdatihglaaml
rdrypeydi perfpgieddlqpvhfsskklldhgftfkytvedmfdaairmcrekgliplatagggralp
```

FASTA2 表示形式

FASTA2:

```
>USGENE|Protein
mgevvatweateggagvkgpvvvtgasgflgswlvmkllqagytratrdrpanvvtkplldlpgater
lslwkadladesfddairgctgvfhvatpmdfeskdpenevikptvegmmssimrackeagtvrriivfts
sagtvnieerqrpyvdqdnwsdvdqcrvkmrgwmyfvskslaekaamayaahgldfisiiptlvvgpf
lsagmppsllitalalvtgneahysilkqvqfvhlddldahlfifehpaaagryvcsshdatihglaaml
rdrypeydi perfpgieddlqpvhfsskklldhgftfkytvedmfdaairmcrekgliplatagggralp
```