

収録範囲	出願人により世界的な所有権機関 (WIPO) に電子的に提出された核酸およびアミノ酸配列情報
ファイル種類	書誌情報, 配列データベース
特徴	完全配列検索および部分配列検索用に GETSEQ, ホモロジー検索用に BLAST および GETSIM の RUN パッケージを提供しています. アラート (自動 SDI 検索) 毎週 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"/> STN Viewer <input type="checkbox"/>
レコード内容	<ul style="list-style-type: none"> レコードには, 出願人により提供された配列および特許情報が収録されています. 各レコードは配列データと配列に関する追加情報, たとえば分子タイプ, 生物名, および特許情報を含みます. 被引用情報 米国特許の法的情報 (1980 年以降)
レコード数	<ul style="list-style-type: none"> 9,120,000 件以上 (2011 年 7 月現在) 7,504,000 件以上の核酸配列 (2011 年 7 月現在) 1,616,000 件以上のタンパク質配列 (2011 年 7 月現在)
収録年代	2001 年 8 月-
更新頻度	毎週更新
言語	英語
データベース	World Intellectual Property Organisation
製作者	34, Chemin des Colomnettes 1211 Geneva Switzerland Phone: +41 22 338 91 11 Fax: +41 22 338 98 20
データベース提供者	FIZ Karlsruhe 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

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

収録源	特許出願人により電子的に提出されたデータ
検索補助資料	<ul style="list-style-type: none">・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)・オンラインヘルプの PDF : http://www.stn-international.de/pctgen_help.html・STNGUIDE・GETSIM/BLAST のよくあるご質問 : http://www.stn-international.de/pctgen_help.html
利用可能なクラスター	<ul style="list-style-type: none">・ ALLBIB・ BIOSCIENCE・ CORPSOURCE・ MEDICINE・ PATENTS (HPATENTS)・ PHARMACOLOGY
価格	<ul style="list-style-type: none">・ STN 料金表・ オンライン上での確認 (=> HELP COST)

SEARCH および DISPLAY フィールド

後方一致検索可能なフィールド (/FEAT) 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), 生物名 (ORGN), および分子タイプ (MTY) フィー ルドからの切り出し語)	なし または/BI	S ANAPHYLATOXIN S PLANT GENE# AND RNA	TI, ORGN, MTY
レコード番号	/AN	S 2002060924.37/AN	AN
出願国 (コードおよび国名)	/AC	S US/AC	AI
出願日 ¹⁾	/AD	S 20011129/AD	AI
出願番号 ²⁾	/AP	S US2001-809003/AP	AI
出願年 ¹⁾	/AY	S 2002/AY	AI
	/DT	S PATENT/DT	DT
入力日 ¹⁾	または/TC /ED	S 20021004/ED	ED
特徴表 *	または/UP /FEAT	S (RNA AND BINDING)/FEAT S ?COMBINAT?/FEAT	FEAT
ファイルセグメント (コードおよびテキスト)	/FS	S PROTEIN/FS	FS
分子タイプ	/MTY	S RNA/MTY	MTY
生物名	/ORGN	S CRASSOSTREA GIGAS/ORGN	ORGN
特許出願人 ³⁾	/PA	S MOLECULAR DYNAMICS/PA	PA
	または/CS		
特許発行国 (コードおよび国名)	/PC	S WO/PC	PI
特許番号 ²⁾	/PN	S WO 2002074961/PN	PI
	または/PATS		
特許発行日 ¹⁾	/PD	S 20030130/PD	PI
特許発行年 ¹⁾	/PY	S 2003/PY	PI
関連出願国 (コードのみ)	/RLC	S FR/RLC	RLI
関連出願日 ¹⁾	/RLD	S 20020208/RLD	RLI, RLIO
関連出願番号 ²⁾	/RLN	S EP2001-1102050/RLN	RLI, RLIO
	または/RLI		
関連出願年 ¹⁾	/RLY	S L1 AND 2000-2001/RLY	RLI, RLIO
配列識別番号 ¹⁾	/SEQN	S 337/SEQN	SEQN
配列長 ¹⁾	/SQL	S 150-175/SQL	SQL
標題	/TI	S HYBRIDIZATION ASSAY#/TI	TI
更新日 ¹⁾	/UP	S L1 AND UP>=20030200	UP

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。

スーパー検索フィールド

必要な情報を含むと思われる1または複数の項目の検索を実行するためにはスーパー検索フィールドコードを入力します。スーパー検索フィールドはクロスファイル検索およびマルチファイル検索を容易にします。EXPAND ではスーパー検索フィールドは使用できません。そのかわりに EXPAND には個々のフィールドを使用します。

SEARCH フィールド	SEARCH 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	S US2001-809003/APPS	AI, RLI

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 — DGENE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。BLAST では BOTH、GETSIM では SIN がデフォルトです。

BLAST および GETSIM 検索では、QUERY コマンドで作成した L 番号を質問式として使用できます。その他、RUN BLAST/GETSIM コマンドと同じ行に直接入力する、あるいは UPLOAD コマンドでテキストファイルをアップロードして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、SORTSCORE 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することができます。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。

GETSIM/BLAST 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が先行する1文字コードに、スペースをつづけ値を入力します：

```
RUN BLAST LI /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 ¹⁾	-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)
Penalty for nucleotide mismatch	-q	-3 (初期値)
Reward for nucleotide match	-r	1 (初期値)

BLAST マトリックスの設定

特定のマトリックスのためには、可能性のあるギャップとギャップ拡張値の指定された組合せのみが可能です。各マトリックスに可能な設定は以下の表にまとめられており、初期設定はそこに示されています。これ以外のどんな組合せも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)
BLOSUM45	9	1
	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	
PAM30	17	1
	16	1
	7	2
	6	2
	5	2
	10	1
PAM70	8	1
	9	1 (デフォルト)
	8	2
	7	2
	6	2
	11	1
	10	1 (デフォルト)
	9	1

検索例 (GETSIM 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 (?):GETSIM
START LOCAL KERMIT TRANSMIT PROCESS
```

```
UPLOAD SUCCESSFULLY COMPLETED
L1 GENERATED
```

```
=> D LQUE
```

```
L1 ANSWER 1 PCTGEN COPYRIGHT 2010 WIPO on STN
LQUE MAVMAPRTLL LVLSGVLALT QTWAGSHSMR YFYTSMSRPG
RGEPRFFAVGYVDDTQFVRFSDAASQRME PRAPWVEQEG PEYWDRETQN MKAQTQNAPV
NLRNLRGYYNQSEAGSHTLQTMHGCDLGPD GRLLRGYYQS AYDGKDYFAL NEDLRSWTAA
DLAAQNTQRKWEAADVAEQIRAYLEGRCVE WLRRYLENGK ETLQRADPPK THVTHHPVSD
HEATLRCWAVGFYPAEITLTWQRDGEDQTQ DTELMETRPA GDGTFQKWA VVVP SGKEQR
YTCHVQHEGLPKPLTLRWEPSSTIPIVG ILAGLVLLGA MVIGAVVA V MWRKSSDRK
GGSYSQAASSDSAQGSVSLTACKV
```

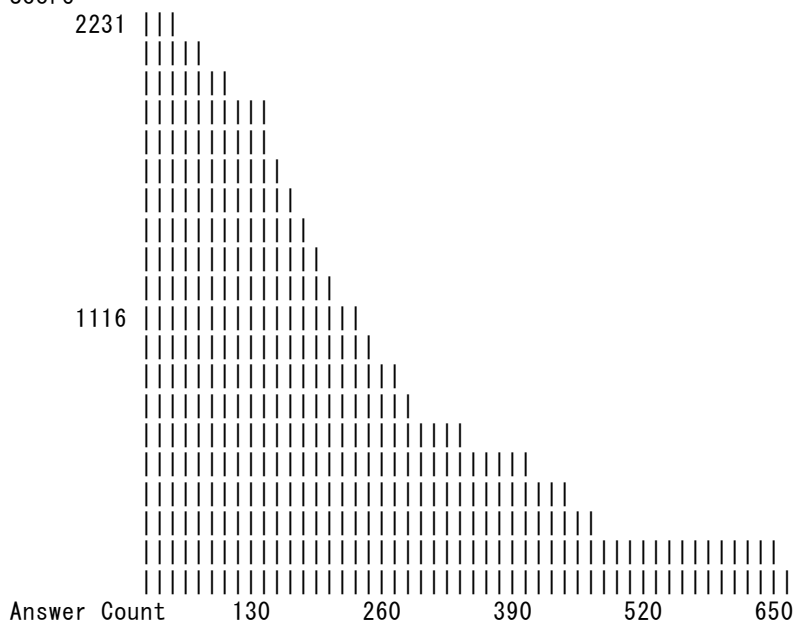
```
=> RUN GETSIM L1/SQP
```

```
RUN GETSIM AT 09:44:13 ON 10 SEP 2010
COPYRIGHT (C) 2010 FIZ KARLSRUHE GMBH
```

```
30000 SEQUENCES PROCESSED
60000 SEQUENCES PROCESSED
:
1360000 SEQUENCES PROCESSED
```

```
642 ANSWERS FOUND ABOVE A THRESHOLD OF 155
QUERY SELF SCORE VALUE IS 2497
BEST ANSWER SCORE VALUE IS 2231
```

```
Similarity
Score
```



```
ENTER EITHER THE NUMBER OF ANSWERS YOU WISH TO KEEP
OR ENTER MINIMUM PERCENT OF SELF SCORE FOLLOWED BY %
(BEST ANSWER PERCENTAGE OF SELF SCORE IS 89%)
ENTER (ALL) OR ? :80%
```

```
L2 RUN STATEMENT CREATED
L2 73 MAVMAPRTLLLVLGVLALTQTWAGSHSMRYFYTSMSRPGRGEPRFFAVG
YVDDTQFVRFSDAASQRMEPRAPWVEQEGPEYWDRETQN MKAQTQNAPV
NLRNLRGYYNQSEAGSHTLQTMHGCDLGPDGRLLRGYYQSAYDGKDYFAL
```

検索例 (GETSIM ホモロジー検索) (続き)

NEDLRSWTAADLAAQNTQRKWEAADVAEQIRAYLEGRVWLRRLRYLNGK
ETLQRADPPKTHVTHHPVSDHEATLRWAVGFYPAEITLWQRDGEDQTQ
DTELMETRPAGDGTfQKWAaVVPsGKEQRYTCHVQHEGLPKPLTLRWEp
SSQSTIpVgIIAGLVLlGAMVIGAVVAaVMWRrkSSDRKGGsYSQAASS
DsaQGSdVSLTACKV/SQP

Answer set arranged by accession number; to sort by descending
similarity score, enter at an arrow prompt (=>) "sor score d".

=> SORT SCORE D
L3 73 SORT L2 SCORE D

=> D 1 BIB SCORE ALIGN

L3 ANSWER 1 OF 73 PCTGEN COPYRIGHT 2010 WIPO on STN
AN 2008016356.13215 PRT PCTGEN
TI GeneMap of the human genes associated with psoriasis
PA Genizon BioSciences
PI WO 2008016356 20080207
ED 20100802
DT Patent
SCORE 2231 89% of query self score 2497
ALIGN Smith-Waterman score: 2231
365 aa overlap starting at 1
mavmaprtlllvsqvaaltqtwaqshsmryfytsmsrppgrgeprffavgvddtqfvrf
::
mavmaprtllllsgalaltqtwaqshsmryfftsvsrppgrgeprfiavgvddtqfvrf
dsdaasqrmeprapwveqegpeywdretqnmkagtqnapvlnrlrgyynqseagshtlq
::
dsdaasqrmeprapwieqegpeywdqetrvkaqsqtdrvdlgtlrgyynqseagshti
tmhgcdlqpdgrllrgyyqsaydgkdyfalnedlrswwtaadlaaqntqrkweaadva
::::: : : : : : : :
imygcdvgsdgrflrgyrqdaydgkdyialnedlrswwtaadmaaqtkrkweahea
raylegrcvewlrrylengketlqradppkthvthhpsdheatlrcwavgfypaeitlt
::::: : : : : : : :
rayldgtcvewlrrylengketlqrdppkthmthhpsdheatlrcwavgfypaeitlt
wqrdgedqtqdtelmetrpagdggtfQKWAaVVPsGKEQRYTCHVQHEGLPKPLTLRWEp
::
wqrdgedqtqdtelvetrpagdggtfQKWAaVVPsGEEQRYTCHVQHEGLPKPLTLRWEp
ssqstipivgiiaglvllgamvigavvaavmwrkssdrkGGsYSQAASSdsaqgsdvs
: : : : : : : : : : : : : : : : : :
ssqstipivgiiaglvllgavitgavvaavmwrkssdrkGGsYtQAASSdsaqgsdvs
tackv
:::::
tackv

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 LQUE

L1  ANSWER 1  PCTGEN COPYRIGHT 2010 WIPO on STN
LQUE CATGGTGGTTAAACTTACCTCATTAGCAGCATCCCTCTACAAGGTGCATTAACTATAAGTATACT

=> RUN BLAST
PLEASE ENTER SEQUENCE QUERY OR ?:END

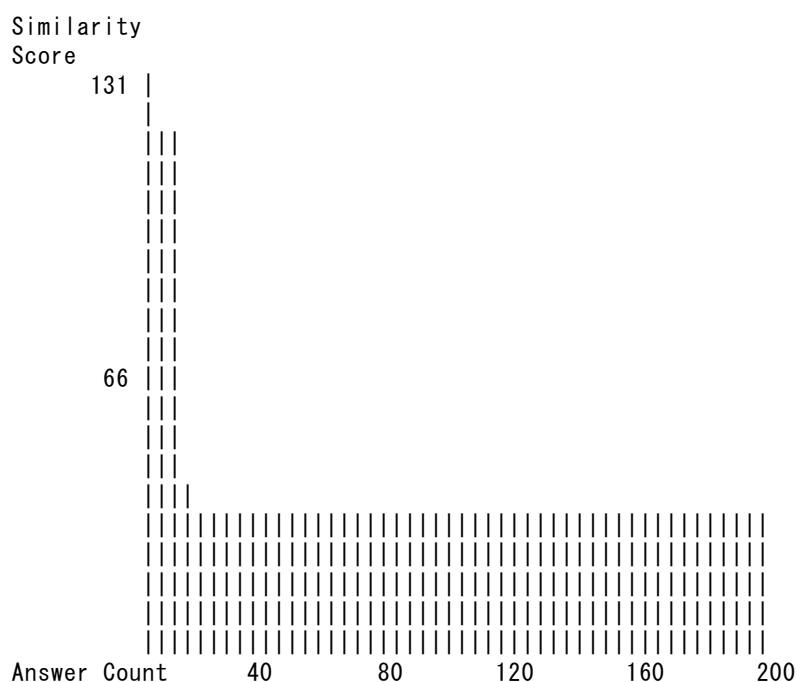
=> RUN BLAST L1/SQN -E 100
```

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, :

189 ANSWERS FOUND BELOW EXPECTATION VALUE OF 100.0

```
QUERY SELF SCORE VALUE IS 131
BEST ANSWER SCORE VALUE IS 131
```



```
ENTER EITHER THE NUMBER OF ANSWERS YOU WISH TO KEEP
OR ENTER MINIMUM PERCENT OF SELF SCORE FOLLOWED BY %
(BEST ANSWER PERCENTAGE OF SELF SCORE IS 100%)
ENTER (ALL) OR ? :85%
L2  RUN STATEMENT CREATED
L2  11 CATGGTGGTTAAACTTACCTCATTAGCAGCATCCCTCTACAAGGTGCATT
    TAACTATAAGTATACT/SQN. -E 100
```

Answer set arranged by accession number; to sort by descending similarity score, enter at an arrow prompt (=>) "sor score d".

検索例 (BLAST ホモロジー検索) (続き)

```
=> SORT SCORE D IDENT D
L3          11 SORT L2 SCORE D IDENT D
```

```
=> D 1 7 TI PA PI SCORE ALIGN
```

```
L3  ANSWER 1 OF 11  PCTGEN COPYRIGHT 2010 WIPO on STN
TI   RNS-sekretierende Bakterien
PA   Bachmann, Till; Villatte, Francois
PI   WO 2002024904    20020328
SCORE 131          100% of query self score 131
BLASTALIGN
```

```
Query = 66 letters
Length = 452
Score = 131 bits (66), Expect = 7e-36
Identities = 66/66 (100%)
Strand = Plus / Plus
```

```
Query: 1  catggtggttaaacttacctcattagcagcatccctctacaaggtgcatttaactataag
          |||
Sbjct: 201 catggtggttaaacttacctcattagcagcatccctctacaaggtgcatttaactataag
```

```
Query: 61  tatact 66
          |||||
Sbjct: 261 tatact 266
```

```
L3  ANSWER 7 OF 11  PCTGEN COPYRIGHT 2010 WIPO on STN
TI   Methods of Diagnosis of Soft Tissue Sarcoma, Compositions and Methods for
      Screening for Soft Tissue Sarcoma Modulators
PA   Aziz, Natasha
      Ginsburg, Wendy M.
      Zlotnik, Albert
PI   WO 2004048938    20040610
SCORE 117          89% of query self score 131
BLASTALIGN
```

```
Query = 66 letters
Length = 911
Score = 117 bits (59), Expect = 2e-31
Identities = 66/67 (98%), Gaps = 1/67 (1%)
Strand = Plus / Plus
```

```
Query: 1  catggtggttaaa-cttacctcattagcagcatccctctacaaggtgcatttaactataa
          |||
Sbjct: 239 catggtggttaaaacttacctcattagcagcatccctctacaaggtgcatttaactataa
```

```
Query: 60  gtatact 66
          |||||
Sbjct: 299 gtatact 305
```

配列検索 (GETSEQ)

配列情報 (タンパク質および核酸配列) は GETSEQ パッケージプログラムで利用できる多種の検索フィールドから得られます。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、直接 RUN GETSEQ の後に入力することもできます。また、配列検索可能な STN の他のファイル、たとえば REGISTRY ファイルや 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。

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⁾	QUE 'THR-SER-GLY-MET-THR'/SQSFP
コードまたはコード列を一重引用符で囲む。	QUE 'GLP'GY/SQSP
各コード間にハイフンを入力する。	RUN GETSEQ 'CYS-ASN-THR-ALA'/SQSP
核酸の1文字コード ²⁾	QUE ATGAAN/SQEN
	RUN GETSEQ ATGAAN/SQSN

1) 一般アミノ酸の 1 文字および 3 文字コードの表を表示するためには矢印プロンプトの位置に HELP AAC を入力して下さい。

2) 核酸のコードの表を表示するためには矢印プロンプトの位置に HELP NUC を入力して下さい。

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配列検索。質問式は完全に定義されていなければならない。	/SQEP	QUE AFFFF/SQEP QUE 'ALA-PHE-PHE-PHE-PHE'/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	RUN GETSEQ YGGFL/SQEFP QUE 'TYR-GLY-GLY-PHE-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	RUN GETSEQ LAGLL/SQSP QUE 'GLP'GY/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	RUN GETSEQ ATCXAW/SQSFP QUE 'LEU-ALA-GLY-LEU-LEU'/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 ATGAAN/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	RUN GETSEQ ATGAAN/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)

GETSEQ による部分配列検索の特殊記号

詳細情報は矢印プロンプト(=>)で HELP SQQ と入力してください。

記号	機能	SEARCH 例
[]	代替残基を特定	RUN GETSEQ LGP[VL]/SQSP QUE LGP['VAL'LEU'LYS']/SQSP
[-]	特定または代替残基を除く	QUE LGP[-H]/SQSP QUE LGP[-'HIS']/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	RUN GETSEQ LGP[-HL]/SQSP QUE LGP[-HL]/SQSP RUN GETSEQ L4{2}/SQSP RUN GETSEQ TAA(TAAA){2}/SQSN RUN GETSEQ(CTG){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	RUN GETSEQ GG(FL){1, 2}/SQSP RUN GETSEQ L3{1, 3}/SQSP RUN GETSEQ(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	RUN GETSEQ FLRRI(RP)?K/SQSP RUN GETSEQ L1{0-1}NN/SQSP RUN GETSEQ L1{0, 1}NN/SQSP
* または {0, } または {0-}	直前の配列, 配列質問式, またはL番号をゼロまたはそれ以上繰り返す	RUN GETSEQ CAT(CGA){0, 1}GGAC/SQSN RUN GETSEQ KLK(WD)*N/SQSP RUN GETSEQ L1{0-}NN/SQSP
+ 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またはそれ以上繰り返す	RUN GETSEQ CAT(CTG){0, }TATT/SQSN RUN GETSEQ KLK(DLE){1, }/SQSP RUN GETSEQ KLK(DLE)+/SQSP
&	配列表記または質問式(L番号)を結合する	RUN GETSEQ L2{1-}/SQSP RUN GETSEQ CAT(CTG){1, }TATT/SQSN RUN GETSEQ L1&L3/SQSFP RUN GETSEQ L2&L5{1, 3}/SQSP

上記の他に, 脱文字符(-), 垂直バー(|)も利用できます。

-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。

|は, 代替記号です。例えば, 代替配列質問式に利用します。

GETSEQ による部分配列検索のギャップ記号

記号	機能	SEARCH 例
.	1残基のギャップ	QUE SY.RPG/SQSP RUN GETSEQ SY..RPG/SQSP
.{m} または {m.}	m残基のギャップ	QUE AAG...TGC/SQSN 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	RUN GETSEQ GFF.{2, 10}LSS/SQSP RUN GETSEQ GFF.{2-10}LSS/SQSP
: 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 AAG.{2, 5}TGC/SQSN RUN GETSEQ AGA:SRI/SQSFP RUN GETSEQ AGA.?SRI/SQSFP QUE AGA.{0, 1}SRI/SQSFP QUE AGA.{0-1}SRI/SQSFP
* または .{0, } または .{0-}	ゼロまたはそれ以上の残基のギャップ	RUN GETSEQ HLC.*TYG/SQSP QUE HLC.{0, }TYG/SQSP RUN GETSEQ HLC.{0-}TYG/SQSP
.+ または .{1, } または .{1-}	1またはそれ以上の残基のギャップ	QUE AAGGCAGATG.*GCAA/SQSN RUN GETSEQ SY.+TH/SQSFP QUE SY.{1-}TH/SQSFP QUE SY.{1, }TH/SQSP RUN GETSEQ TCCTG.+GTGG/SQSN

DISPLAY および PRINT 形式

回答のディスプレイとオフライン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。複数のコードは、“D L1 1-5 TI PA SEQ”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。

AN, DT, MTY, ORGN, PA, RLI, RLIO, SEQN, SQL, TI の検索フィールドでハイライト機能が使えます。HIT, KWIC, OCC 形式を使うためには、検索時にヒットターム・ハイライト機能が ON になっていることが必要です。

形式	英語名	内容	入力例
AI (AP) ¹⁾	Application Information	出願情報	D AI
AIO ^{1) 2)}	Application Information Original	出願情報 (オリジナル)	D AIO
AN	Accession Number	受入れ番号	D AN
DT (TC)	Document Type	資料種類	D DT
ED	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) 3)}	Percent Identity	同一性パーセント	
MTY	Molecule Type	分子タイプ	D MTY
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
(PATS, PN)			
RLI ¹⁾	Related Application Information	関連出願情報	D RLI
RLIO ²⁾	Related Application Information Original	関連出願情報 (オリジナル)	D RLIO
SEQ ⁵⁾	Sequence (1-letter codes)	配列 (1文字コード)	D SEQ
SEQ3 ⁵⁾	Sequence (3-letter codes)	配列 (3文字コード)	D SEQ3
SEQN	Sequence Identity Number	配列識別番号	D SEQN
SEQO ^{2) 5)}	Original Sequence (alignment of nucleotide sequence and peptide sequence it expresses when given)	オリジナル配列 (核酸配列と翻訳されたタンパク質配列の対応表示)	D SEQO
SQL	Sequence Length	配列長	D SQL
TI	Title	標題	D TI
UP ²⁾	Update Date	更新日	D UP
ALIGN ⁴⁾	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) 5)}	AN, MTY, TI, PA, PI, AI, RLI, DT, ORGN, SQL, SEQ, FEAT		D ALL
APPS ¹⁾	AN, RLI		D APPS
BIB ¹⁾	AN, MTY, TI, PA, PI, AI, RLI, DT (デフォルト)		D BIB
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) 5)}	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
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コード毎に分けて表示		D LS2
SCORE ⁴⁾	類似性スコア		D SCORE
SQIDE ⁵⁾	TI, SQL, SEQ, FEAT		D SQIDE
SQ3IDE ⁵⁾	TI, SQL, SEQ3, FEAT		D SQ3IDE
SCAN ⁶⁾	TI (回答番号なしのランダム表示)		D SCAN
TRIAL (TRI, SAM, FREE)	TI, MTY, SQL		D TRIAL
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 20語を表示 (KeyWord-In-Context)		D KWIC NOH
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。ダウンロード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 BLAST の回答に対して使用します。
- 4) RUN GETSIM または RUN BLAST の回答に対して使用します。
- 5) 配列は WIPO の WST.25 に従って提供されています。
- 6) 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
レコード番号	AN	N	Y
出願国	AC	Y ²⁾	Y
出願日	AD	Y ²⁾	Y
出願情報	AI	Y ^{2), 3)}	Y
出願番号	AP	Y ²⁾	Y
出願番号および関連出願番号	APPS	Y ²⁾	N
出願年	AY	Y ²⁾	Y
資料種類	DT	Y	Y
入力日	ED	Y ²⁾	Y
ファイルセグメント	FS	Y ²⁾	Y
同一性パーセント	IDENT	Y ⁷⁾	Y ⁷⁾
分子タイプ	MTY	Y	Y
生物名	ORGN	Y	Y
特許出願人	PA	Y	Y
特許発行国	PC	Y	Y
特許情報	PI	N	Y
特許番号	PN	Y ²⁾	Y
	PATS	Y ²⁾	Y
特許発行日	PD	Y ²⁾	Y
特許発行年	PY	Y ²⁾	Y
関連出願国	RLC	Y	Y
関連出願日	RLD	Y	Y
関連出願情報	RLI	Y ⁴⁾	Y
関連出願番号	RLN	Y	Y
関連出願年	RLY	Y	Y
配列 (1文字コード)	SEQ	Y ^{2), 5)}	N
配列 (3文字コード)	SEQ3	Y ^{2), 5)}	N
配列識別番号	SEQN	Y	Y
配列長	SQL	Y	Y
類似性スコア	SCORE	N	Y ⁶⁾
標題	TI	Y (デフォルト)	Y
更新日	UP	Y ²⁾	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) SELECT HIT および ANALYZE HIT はこのフィールドでは使えません。
- 3) 出願番号が SELECT または ANALYZE され、SELECT したタームに /AP が付与されます。
- 4) 関連出願番号が SELECT または ANALYZE され、SELECT したタームに /RLN が付与されます。
- 5) SELECT したタームに /SQSP が付与されます。
- 6) GETSIM または BLAST により生成された L 番号を指定します。
- 7) 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

TRIAL 形式

TI Gene Expression Profiles in Liver Cancer
 MTY DNA
 SQL 2093

SQIDE 形式

TI Compositions and Methods Relating to Osteoarthritis
 SQL 100
 SEQ
 1 agttgngtgc cgttggaccg naggaaaact catagactca tgggagcgtg
 51 aggcttcgag ccgcctaatt ttttaaccct aaatgtcgaa aggcttctgg

FEATURE TABLE:

Key	Location
misc_feature 6, 21	n = A, T, C or G

IALL 形式

ACCESSION NUMBER: 2001057272.15599 DNA PCTGEN
 TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
 FORANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
 PATENT ASSIGNEE: Molecular Dynamics, Inc.
 Penn, Sharron G.
 Rank, David R.
 Hanzel, David K.
 Chen, Wensheng
 PATENT INFO: WO 2001057272 20010809
 REL APPL INFO: US 2000-180312P 20000204; US 2000-207456P 20000526; US
 2000-632366 20000803; GB 2000-24263 20001003; US 2000-236359P
 20000927; US 2000-234687P 20000921; US 2000-608408 20000630
 FILE ENTRY DATE: 20020923
 DOCUMENT TYPE: Patent
 ORGANISM: Homo sapiens
 SEQUENCE LENGTH: 100
 SEQUENCE
 1 cccagagatt ctgattctgc aaatcttgag cagcctgaga ttctgcagtt
 51 ctatgaagct tccaggtagt gtcaatgctg gtgctaggct gaccatagta

FEATURE TABLE:

Key	Location
	MAP TO ALO35448.28
	EXPRESSED IN PLACENTA, SIGNAL
	= 1.5
	NT HIT: U29185.1, EVALUE
	7.00e-04
	EST_HUMAN HIT: AA047634.1,
	EVALUE 2.20e-01

FASTA 形式

FASTA:

```
>PCTGEN|2009009830.168|PRT|sequence 168 from W02009009830
mappqpk|prpptrrractpsssarrssqrrtpsslswrpswppagsrpprwpqssraqrrtrkhqt
wwtacsgcwrrttcralwrrvrmaaspqgtaprrcasspptrtvspwrrsrstrtrsswragttrtrsl
taasrstrrtacrrssttaqtrastacstkrtrtirsssprssssttasaraspsstsaasappwppsp
httppsrgststsptrrrrsasptsaatcsrrcpwatpssgssttgvtshivarcrtatrcqctv
rwcwwsascrtrrpsarrgcststscsrttqaagrtrgstrrspgelaspasphtstptpgpssps
```

FASTA2 形式

FASTA2:

```
>PCTGEN|PRT
mappqpk|prpptrrractpsssarrssqrrtpsslswrpswppagsrpprwpqssraqrrtrkhqt
wwtacsgcwrrttcralwrrvrmaaspqgtaprrcasspptrtvspwrrsrstrtrsswragttrtrsl
taasrstrrtacrrssttaqtrastacstkrtrtirsssprssssttasaraspsstsaasappwppsp
httppsrgststsptrrrrsasptsaatcsrrcpwatpssgssttgvtshivarcrtatrcqctv
rwcwwsascrtrrpsarrgcststscsrttqaagrtrgstrrspgelaspasphtstptpgpssps
```