

収録範囲	<ul style="list-style-type: none"> <li>・ 10 以上の塩基を持つ核酸配列</li> <li>・ 4 残基以上からなるすべてのアミノ酸配列</li> <li>・ あらゆる長さのプロープおよびプライマー</li> </ul>
ファイル種類	書誌情報, 配列データベース
特徴	<p>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。</p> <p>アラート (自動 SDI 検索) 隔週</p> <p>CAS RN<sup>®</sup> (CAS 登録番号) <input type="checkbox"/> ページイメージ <input type="checkbox"/> STN AnaVist <input type="checkbox"/></p> <p>Keep &amp; Share <input checked="" type="checkbox"/> 中間一致・<input checked="" type="checkbox"/> STN Easy <input type="checkbox"/></p> <p>後方一致検索</p> <p>練習用ファイル <input type="checkbox"/> 構造図 <input type="checkbox"/></p>
レコード内容	<ul style="list-style-type: none"> <li>・ 41 特許発行機関から発行された特許を収録する WPI ファイルのベーシック特許から抽出した核酸およびアミノ酸のデータベースです。</li> <li>・ 約 15 万件の世界中の配列特許から, 約 3,930 万件の配列レコードを収録しています。</li> <li>・ すべてのレコードには, Clarivate Analytics が作成した標題, 配列に関する簡潔な説明, 特許情報, 詳細な索引, 特徴表, および配列データが収録されています。</li> <li>・ WPI ファイルのアクセッション番号は OS フィールドに表示されます。</li> <li>・ WPI ファイルの特許ファミリー情報や INPADOC ファイルの法的状況を, DGENE ファイルで直接表示することができます。</li> </ul>
レコード数	<ul style="list-style-type: none"> <li>・ 3,930 万件以上 (2015 年 8 月現在)</li> <li>・ 2,970 万件以上の核酸配列 (2015 年 8 月現在)</li> <li>・ 950 万件以上のタンパク質配列 (2015 年 8 月現在)</li> </ul>
収録年代	1981 年 -
更新頻度	隔週更新
言語	英語
データベース 製作者	<p>Clarivate Analytics (UK) Limited</p> <p>Friars House, 160 Blackfriars Rd.</p> <p>London SE1 8EZ</p> <p>United Kingdom</p> <p>Copyright Holder: Clarivate Analytics</p>
データベース 代理店	<p>クラリベイト・アナリティクス・</p> <p>ジャパン株式会社 (日本法人)</p> <p>〒107-6119</p> <p>東京都港区赤坂 5-2-20 赤坂パークビル 19F</p>

<p><b>ヨーロッパ</b></p> <p><b>STN カールスルーエ</b></p> <p>FIZ Karlsruhe</p> <p>P.O. Box 2465</p> <p>76012 Karlsruhe</p> <p>Germany</p> <p>Phone: +49-7247-808-555</p> <p>Fax: +49-7247-808-259</p> <p>E-mail: helpdesk@fiz-karlsruhe.de</p> <p>Internet: www.stn-international.de</p>	<p><b>日本</b></p> <p><b>STN 東京</b></p> <p><b>化学情報協会</b></p> <p>〒113-0021 東京都文京区本駒込6-25-4 中居ビル</p> <p>Phone: 0120-003-462 (Help Desk)</p> <p>: 0120-151-462 (上記以外)</p> <p>Fax: 03-5978-4090</p> <p>E-mail: support@jaici.or.jp (Help Desk)</p> <p>customer@jaici.or.jp (上記以外)</p> <p>Internet: www.jaici.or.jp</p>	<p><b>北アメリカ</b></p> <p><b>STN コロンバス</b></p> <p>CAS</p> <p>P.O. Box 3012</p> <p>Columbus, Ohio 43210-0012 U.S.A</p> <p>CAS Customer Care:</p> <p>Phone: 800-753-4227 (North America)</p> <p>614-447-3700 (worldwide)</p> <p>Fax: 614-447-3751</p> <p>E-mail: help@cas.org</p> <p>Internet: www.cas.org</p>
--	--	---

データベースサマリーシート作成: 化学情報協会

---

収録源	• Derwent World Patents Index (WPIDS/WPIX/WPINDEX) に収録されている世界の主要特許発行機関から発行された特許
検索補助資料	• STN 技術資料 <a href="http://www.jaici.or.jp/stn/stn_doc_01.html">http://www.jaici.or.jp/stn/stn_doc_01.html</a> • 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 • DGENE Workshop Manual : <a href="http://www.stn-international.de/exploring_geneseq.html">http://www.stn-international.de/exploring_geneseq.html</a>
利用可能なク ラスター	• ALLBIB • AUTHORS • BIOSCIENCE • CORPSOURCE • MEDICINE • PATENTS (HPATENTS) • PHARMACOLOGY

---

## SEARCH および DISPLAY フィールド

特徴表 (/FEAT) フィールドでは, 後方一致検索が可能です.

フィールド	SEARCH コード	SEARCH 例	DISPLAY コード
基本索引 標題 (TI) キーワード (KW) 抄録 (AB) 説明 (DESC) 生物名 (ORGN) (以上からの切出し語)	なし または/BI	S F PROMOTER S HUMAN INSULIN(L)PREPARATION S JUNIPERUS VIRGINIANA	AB, DESC, KW, ORGN, TI
レコード番号	/AN	S 95P-R67826/AN	AN
アミノ酸 <sup>1)</sup>	/AA	S 5 S/AA AND 2 T/AA	AA
アミノ酸の数 <sup>2)</sup>	/AA. CNT	S L2 AND AA. CNT>9	AA
出願国 (WIPO コードとテキスト)	/AC	S AT/AC S AUSTRIA/AC	AI
出願日 <sup>2)</sup>	/AD	S 7 JUN 1999/AD	AI
出願番号 <sup>3)</sup>	/AP	S AU94-64520/AP S 94AU-0064520/AP	AI
出願年 <sup>2)</sup>	/AY	S 2000/AY	AI
クロスリファレンス (関連する DGENE 配列レコード)	/CR	S 90N-Q01810/CR	CR
データの入力日 <sup>2)</sup>	または/XR /DED	S 19901220/DED	DED
説明	/DESC	S DYNORPHIN-A./DESC	DESC
資料種類 (コードとテキスト)	/DT または/TC	S PATENT/DT S P/DT	DT
入力日 <sup>2)</sup>	/ED	S ED>20020100	ED
特徴表 <sup>4)</sup>	/FEAT	S CDS/FEAT S ?CYTE?/FEAT	FEAT
ファイルセグメント (コードとテキスト)	/FS	S L10 AND PROTEIN/FS S L10 AND NS/FS	FS
発明者	/IN または/AU	S GOSSEN M/IN	IN
キーワード	/KW	S F PROMOTER/KW	KW
言語 (ISO コードとテキスト)	/LA	S DE/LA S FRENCH/LA	LA
分子タイプ	/MTY	S RNA/MTY	MTY
核酸 <sup>1)</sup>	/NA	S 7 C/NA AND 0 H/NA	NA
核酸の数 <sup>2)</sup>	/NA. CNT	S L3 AND NA. CNT>10	NA
生物名	/ORGN	S ADENOVIRUS/ORGN S (SIMIAN AND ADENOVIRUS)/ORGN	ORGN
その他の収録源 (WPI ファイルのレコード番号)	/OS	S 94-151326/OS	OS
特許出願人 <sup>1)</sup>	/PA または/CS	S CHIRON SPA/PA	PA
特許出願人コード	/PACO	S BADI/PACO	PA
特許発行国 (WIPOコードとテキスト)	/PC または/PCS	S DE/PC S UNITED KINGDOM/PC	PI
特許種別コード	/PK	S C/PK	PI
特許番号 <sup>3)</sup>	/PN または/PATS	S DE4244565/PN S EP-348819/PATS	PI
特許中の配列情報の記載位置	/PSL	S (CLAIM(S)1)/PSL	PSL

(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
優先権主張国 (WIPO コードとテキスト)	/PRC	S L1 AND US/PRC	PRAI
優先権主張日 <sup>2)</sup>	/PRD	S 19930907/PRD	PRAI
優先権主張日, 最先のものに限定 <sup>2)</sup>	/PRDF	S 17 JUNE 1993/PRDF	PRAI
優先権出願番号 <sup>3)</sup>	/PRN	S 93US-0078471/PRN	PRAI
		S US93-78471/PRN	
優先権主張年 <sup>2)</sup>	/PRY	S 1993/PRY	PRAI
優先権主張年, 最先のものに限定 <sup>2)</sup>	/PRYF	S 1993/PRYF	PRAI
特許発行日 <sup>2)</sup>	/PD	S 1 NOV 2001/PD	PI
特許発行年 <sup>2)</sup>	/PY	S 1997-1999/PY	PI
配列長 <sup>2)</sup>	/SQL	S 70-90/SQL	SQL
標題	/TI	S HUMAN INSULIN/TI	TI
更新日 <sup>2)</sup>	/UP	S L1 AND UP>=20020101	UP

- 1) このフィールドでは, (S) 演算子をスペースで代用できます。
- 2) 数値演算子また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。
- 3) 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。
- 4) このフィールドでは, 前方一致, 後方一致, および中間一致が利用できます。語幹として少なくとも 4文字が必要で。

スーパー検索フィールド <sup>1)</sup>

フィールド	スーパー検索 フィールドコード	検索された フィールド	SEARCH 例	DISPLAY コード
出願番号グループ <sup>2)</sup>	/APPS	/AP, /PRN	S 94AU-0064520/APPS	AI, PRAI
特許発行国グループ (WIPO コードとテキスト)	/PCS	/PC	S UNITED KINGDOM/PCS	PI
特許番号グループ <sup>2)</sup>	/PATS	/PN	S EP-348819/PATS	PI

- 1) 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は, クロスファイルおよびマルチファイル検索を容易にします。スーパー検索フィールドで EXPAND は利用できません。代わりに個々のフィールドで EXPAND してください。
- 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。

## 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

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。

### BLAST/GETSIM の検索タイプ

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

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

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

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

```
RUN BLAST L1/SQN COM ALERT
```

```
RUN GETSIM L1/TSQN BOTH ALERT
```

アラート検索はデータ更新毎、すなわち隔週で実行されます。

検索タイプ	SEARCH コード	SEARCH 例 <sup>1)</sup>
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 <sup>2)</sup>		RUN GETSIM L1/SQN SIN
相補鎖		RUN GETSIM L1/SQN COM
双方 <sup>3)</sup>		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 <sup>2)</sup>		RUN GETSIM L1/TSQN SIN
相補鎖		RUN BLAST L1/TSQN COM
双方 <sup>3)</sup>		RUN BLAST L1/TSQN BOTH

1) L1は UPLOAD または QUERY を使用して生成された配列質問式です。

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://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE\\_TYPE=BlastDocs](http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs)

上級ユーザーオプションはハイフンが先行する 1 文字コードの後にスペースを入れ値を入力します：

```
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 <sup>1)</sup>	-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す。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)
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
# N3
```

```
UPLOAD SUCCESSFULLY COMPLETED
L1 GENERATED
```

```
=> D LQUE L1
```

```
L1 ANSWER 1 DGENE COPYRIGHT 2010 CLARIVATE ANALYTICS on STN
LQUE LDHILQKTERGVRLHPLARTAKVKNEVNSFKAALSSLAKHGEYAPFARLLNLSGVNNLEHGLFPQLSAIA
```

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

```
RUN GETSIM AT 05:09:58 ON 06 SEP 2010
COPYRIGHT (C) 2010 FIZ KARLSRUHE GMBH
```

```
180000 SEQUENCES PROCESSED
```

```
:
```

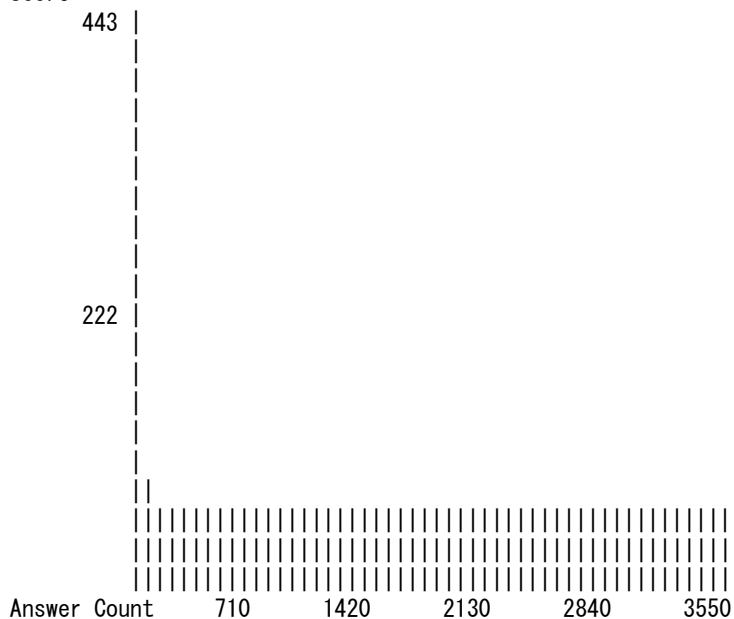
```
6260000 SEQUENCES PROCESSED
```

```
3544 ANSWERS FOUND ABOVE A THRESHOLD OF 57
```

```
QUERY SELF SCORE VALUE IS 443
```

```
BEST ANSWER SCORE VALUE IS 443
```

```
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 100%)
```

```
ENTER (ALL) OR ? :80%
```

```
L2 RUN STATEMENT CREATED
```

```
L2 29 LDHILQKTERGVRLHPLARTAKVKNEVNSFKAALSSLAKHGEYAPFARLL
NLSGVNNLEHGLFPQLSAIA/SQP
```

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

```
=> SOR SCORE D
```

```
PROCESSING COMPLETED FOR L2
```

```
L3 29 SOR L2 SCORE D
```

検索例 (GETSIM ホモロジー検索) (続き)

=> D SCORE ALIGN SEQ 1 10

L3 ANSWER 1 OF 29 DGENE COPYRIGHT 2010 CLARIVATE ANALYTICS on STN

SCORE 443 100% of query self score 443

ALIGN Smith-Waterman score: 443

70 aa overlap starting at 251

```
ldhilqktergvr|lhplartakvknevnsfkaalsslakhgeyapfar|lnlsgvnnleh
.....
ldhilqktergvr|lhplartakvknevnsfkaalsslakhgeyapfar|lnlsgvnnleh
glfpqlsaia
.....
glfpqlsaia
```

SEQ

```
1 mdsrpqkiwm apsltesdmd yhkiltagls vqqgivrqr ipvyqvnnle
51 eicqliiqaf eagvdfqesa dsflmlclh hayqgdyklf lesgavkyle
101 ghgfrfevkk rdgvrleel lpavssgkni krtlaampee etteanagqf
151 lsfaslflpk lvvgekacle kvqrqiqvha eqgliqypta wqsvghmmvi
201 frlmrtnfli kfllihqgmh mvaghdanda visnsvaqar fsgllivktv
251 ldhilqkter gvr|lhplart akvknevnsf kaalsslakh geypapfar|l
301 nlsqvnnleh glfpqlsaia lgvatahgst lagvvnveqy qqlreaatea
351 ekqlqyaes reldhlgidd qekkilnmfh qkkneisfq tnamvtlrke
401 rlaklteait aaslpktsgh ydddddipfp gpinddnpq hqdddptdsq
451 dttipdvvd pddgsygeyq sysengmap ddlvlfldl dedtkpvpn
501 rstkggqqkn sqkgqhiegr qtqfrpiqnv pghrtihha sapltdndrr
551 nepsgstspr mltpineead plddaddets slpplesdde eqdrdgtsnr
601 tptvappapy yrdhsekkel pqdeqqdqdh tqearnqdsd ntqsehslee
651 myrhilrsqg pfdavlyyhm mkdepvfst sdgkeytypd sleeyppwl
701 tekeamneen rfvtldgqqf ywpvmnhknk fmailqhqq
```

L3 ANSWER 10 OF 29 DGENE COPYRIGHT 2010 CLARIVATE ANALYTICS on STN

SCORE 443 100% of query self score 443

ALIGN Smith-Waterman score: 443

70 aa overlap starting at 150

```
ldhilqktergvr|lhplartakvknevnsfkaalsslakhgeyapfar|lnlsgvnnleh
.....
ldhilqktergvr|lhplartakvknevnsfkaalsslakhgeyapfar|lnlsgvnnleh
glfpqlsaia
.....
glfpqlsaia
```

SEQ

```
1 hgfrfevkk rdgvrleell pavssgknik rtaampeee tteanagqfl
51 sfasflfplk vvgekacle vqrqiqvhae qgliqyptaw qsvghmmvif
101 rlmrtnflik fllihqgmhx vaghdandav isnsvaqarf sglivktvl
151 dhilqkterg vr|lhplarta kvknevnsfk aalsslakhg eyapfar|ln
201 lsgvnnleh lfpqlsaial gvatahgstl agvvnveqy qqlreaatea
251 kqlqyaesr eldhlgidd ekkilnmfhq kkneisfqqt namvtlrker
301 laklteaita aslpktsghy ddddipfpg pinddnpgh qdddptdsq
351 ttipdvvd pddgsygeyqs ysengmap dlvlfdl dedtkpvpnr
401 stkggqqkns qkgqhtegrq tqsrtqnv pghrtihhas apltdndrrn
451 epsgstspr mltpineead plddaddets slpplesdee qdrdgtsnr
501 ptvappavy rdhsekkel pqdeqqdqdh tqearnqsdn tqpehsfeem
551 yrhilrsqg pfdavlyyhm kdepvfsts dgkeytyp
```



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
```

```
L1 ANSWER 1 DGENE COPYRIGHT 2010 CLARIVATE ANALYTICS on STN
LQUE ATGGCCCTGAAGAATGATGAGATAATAGATGCCACTCAAAAAGGAAATTGCTCTCGTTTCATGAATCACAGCT
GTGAACCAAATTTGTGAAACCCAAAAATGGACTGTGAACGGACAACCTGAGGGTTGGGTTTTTACCACCAAACCT
GGTT
```

```
=> RUN BLAST L1/SQN -E 0.1
```

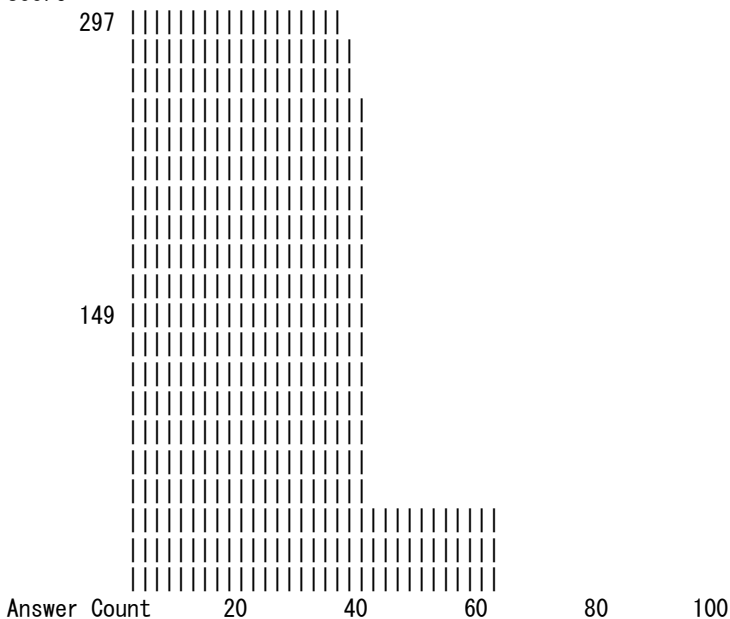
```
BLAST Version 2.2
```

```
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,
:
```

```
61 ANSWERS FOUND BELOW EXPECTATION VALUE OF 1.0e-01
```

```
QUERY SELF SCORE VALUE IS 297
BEST ANSWER SCORE VALUE IS 297
```

```
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 100%)
```

```
ENTER (ALL) OR ? :80%
```

```
L2 RUN STATEMENT CREATED
```

```
L2 39 ATGGCCCTGAAGAATGATGAGATAATAGATGCCACTCAAAAAGGAAATTG
CTCTCGTTTCATGAATCACAGCTGTGAACCAAATTTGTGAAACCCAAAAAT
GGACTGTGAACGGACAACCTGAGGGTTGGGTTTTTACCACCAAACCTGGTT
/SQN. -E 0.1
```

```
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          39 SORT L2 SCORE D IDENT D
```

```
=> D 1 5 SCORE ALIGN
```

```
L3 ANSWER 1 OF 39 DGENE COPYRIGHT 2010 CLARIVATE ANALYTICS on STN
SCORE 297      100% of query self score 297
```

```
BLASTALIGN
```

```
Query = 150 letters
Length = 7987
Score = 297 bits (150), Expect = 2e-84
Identities = 150/150 (100%)
Strand = Plus / Plus
```

```
Query: 1   atggccctgaagaatgatgagataatagatgccactcaaaaaggaaattgctctcgttc
          |||
Sbjct: 4434 atggccctgaagaatgatgagataatagatgccactcaaaaaggaaattgctctcgttc
```

```
Query: 61  atgaatcacagctgtgaaccaaattgtgaaacccaaaaatggactgtgaacggacaactg
          |||
Sbjct: 4494 atgaatcacagctgtgaaccaaattgtgaaacccaaaaatggactgtgaacggacaactg
```

```
Query: 121 agggttgggtttttaccaccaaactggtt 150
          |||
Sbjct: 4554 agggttgggtttttaccaccaaactggtt 4583
```

```
L3 ANSWER 5 OF 39 DGENE COPYRIGHT 2010 CLARIVATE ANALYTICS on STN
SCORE 297      100% of query self score 297
```

```
BLASTALIGN
```

```
Query = 150 letters
Length = 7987
Score = 297 bits (150), Expect = 2e-84
Identities = 150/150 (100%)
Strand = Plus / Plus
```

```
Query: 1   atggccctgaagaatgatgagataatagatgccactcaaaaaggaaattgctctcgttc
          |||
Sbjct: 4434 atggccctgaagaatgatgagataatagatgccactcaaaaaggaaattgctctcgttc
```

```
Query: 61  atgaatcacagctgtgaaccaaattgtgaaacccaaaaatggactgtgaacggacaactg
          |||
Sbjct: 4494 atgaatcacagctgtgaaccaaattgtgaaacccaaaaatggactgtgaacggacaactg
```

```
Query: 121 agggttgggtttttaccaccaaactggtt 150
          |||
Sbjct: 4554 agggttgggtttttaccaccaaactggtt 4583
```

## 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 や PCTGEN 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. 詳細については, DGENE ファイルで, 矢印プロンプト(=>)の後で, HELP GETSEQ と入力してください.

## 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 <sup>1), 2)</sup>	QUE LAGLL/SQSP
一般アミノ酸の3文字コード <sup>1), 2)</sup> 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 'THR-SER-GLY-MET-THR'/SQSFP QUE 'GLP' GY/SQSP
核酸の1文字コード <sup>3)</sup>	RUN GETSEQ 'CYS-ASN-THR-ALA'/SQSP QUE ATGAAN/SQEN RUN GETSEQ ATGAAN/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) 配列中で特殊アミノ酸は, 該当するものがあれば対応する一般アミノ酸で, もしくは X (またはXXX) で表されます. 特殊アミノ酸についての詳細は, 特徴表 (FEAT) 中にあります.
- 3) 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 S EQ3 フィールド中に3 文字コードで表示されます.

RUN GETSEQ バッチ検索では, 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/SQEN 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 <sup>1)</sup>	/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. <sup>1)</sup>	/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)

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

下記の他に、脱字符号(ˆ), 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。

部分配列質問式のさまざまな指定法に関する詳細は、DGENE 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。

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 RUN GETSEQ 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 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 RUN GETSEQ L2{1-}/SQSP
&	配列表記または質問式 (L 番号) を結合する	RUN GETSEQ CAT(CTG){1, }TATT/SQSN RUN GETSEQ L1&L3/SQSFP RUN GETSEQ L2&L5{1, 3}/SQSP

## 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 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 QUE AAGGCAGATG.*GCAA/SQSN
. + または . {1, } または .{1-}	1 またはそれ以上の残基のギャップ	RUN GETSEQ SY.+TH/SQSFP QUE SY.{1-}TH/SQSFP QUE SY.{1, }TH/SQSP RUN GETSEQ TCCTG.+GTGG/SQSN

## 特許出願人コード (/PACO) シソーラス

/PACO フィールドのシソーラスを用いて、特許出願人コードと特許出願人名の対応を表示できます。この機能を用いると、簡便に会社名とコードを関連付けることができ、またコードに対応する会社名を調べることもできます。コードのアルファベット順リストを表示させるには、会社名の単語あるいはフルネームを /PA フィールドで調べたのち、/PACO フィールドで EXPAND してください。各コードは、出現頻度および関連語数とともに表示されます。会社名に +ALL を付けて EXPAND すると、対応するコードの一覧表が表示されます。また、コードに +DEF をつけて EXPAND すると、会社名コードに対応する会社名が表示されます。すべての関係コードは SEARCH および DISPLAY コマンドとともに利用できます。

関係コード	内 容	入 力 例
ALL	会社名に関連するすべての特許出願人コード (SELF, CODE)	E PFIZER+ALL/PACO
DEF	コードに対応するすべての会社名 (SELF, DEF)	E PFIZ+DEF/PACO

## 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は ON であることが必要です。

形 式	英 語 名	内 容	入 力 例
AA	Amino Acid	アミノ酸	D AA TI KW
AB	Abstract	抄録	D TI AB
AI (AP) <sup>1)</sup>	Application Information	出願情報	D AI
AN <sup>2)</sup>	Accession Number	レコード番号	D AN
CR (XR)	Cross Reference (to related DGENE records)	クロスリファレンス (関連する DGENEのレコードへの)	D CR
DED <sup>2)</sup>	Data Entry Date	データの入力日	D DED
DESC <sup>2)</sup>	Description	説明	D DESC
DT (TC) <sup>2)</sup>	Document Type	資料種類	D TI DT
ED <sup>2), 3)</sup>	Entry Date	入力日	D ED
FEAT	Feature Table	特徴表	D AA FEAT
FS <sup>2), 3)</sup>	File Segment	ファイルセグメント	D FS
IDENT <sup>3), 6)</sup>	Percent Identity	同一性パーセント	D IDENT
IN (AU)	Inventor	発明者	D IN PA
KW (ST) <sup>2)</sup>	Keyword	キーワード	D KW
LA <sup>2)</sup>	Language	言語	D LA
LS	Legal Status	法的状況	D LS
LS2	Legal Status, Detailed Version	法的状況の詳細	D LS2
MTY <sup>2)</sup>	Molecule Type	分子タイプ	D MTY KW
NA	Nucleic Acid	核酸	D NA
ORGN <sup>2)</sup>	Organism Name	生物名	D ORGN
OS <sup>2)</sup>	Other Source (DWPI accession number)	その他の収録源 (WPIレコード番号)	D OS
PA (CS)	Patent Assignee (includes PACO)	特許出願人	D IN PA
PACO	Patent Assignee Code	特許出願人コード	D PACO
PI <sup>1)</sup>	Patent Information	特許情報	D PI
(PN, PATS)			
PRAI (PRN) <sup>1)</sup>	Priority Information	優先権出願情報	DIS PRAI L5
PSL	Patent Sequence Location	特許中の配列情報の記載位置	D PSL
SCORE <sup>2), 4)</sup>	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 1-3
SQL <sup>2)</sup>	Sequence Length	配列長	D SQL
TI <sup>2)</sup>	Title	標題	D TI IN PA
UP <sup>2), 3)</sup>	Update Date	更新日	D TI UP

## DISPLAY および PRINT 形式 (続き)

形式	内容	入力例
ABS ALIGN <sup>2), 4)</sup>	AN, MTY, AB 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 TI ABS 1-5 D ALIGN 1, 10, 100
ALL <sup>1)</sup>	AN, MTY, TI, IN, PA, PI, AI, PRAI, PSL, DED, DT, LA, OS, CR, DESC, KW, ORGN, AB, AA, NA, SQL, SEQ, FEAT	D ALL
APPS <sup>1)</sup>	AI, PRAI	D APPS
BIB <sup>1)</sup>	AN, MTY, TI, IN, PA, PI, AI, PRAI, PSL, DT, LA, OS, CR, DESC (デフォルト)	D BIB ABS
FAM	WPIの特許ファミリー情報 (PI, ADT, FDT, PRAI)	D FAM
IALL <sup>1)</sup>	フィールド名付きインデント型ALL形式	D IALL 1-3 L4
IBIB <sup>1)</sup>	フィールド名付きインデント型BIB形式	D IBIB KW
SCAN <sup>2), 5)</sup>	AN, TI, DESC (回答番号なしのランダム表示)	D SCAN
SQIDE	AN, MTY, AA, NA, SQL, SEQ, FEAT	D SQIDE ABS
SQ3IDE	AN, MTY, AA, NA, SQL, SEQ3, FEAT	D SQ3IDE
TRIAL <sup>2)</sup> (TRI, SAM)	AN, MTY, TI, DESC, KW, SQL	D TRI TOTAL
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 <sup>2)</sup>	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) 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。

3) カスタム形式のみの表示です。

4) RUN GETSIM もしくは RUN BRAST の回答である必要があります。詳しくは、相同性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

6) RUN GETSIM の回答である必要があります。詳しくは、相同性検索の項を参照してください。

## 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 <sup>1)</sup>	SORT
抄録	AB	Y <sup>2)</sup>	N
レコード番号	AN	Y	N
アミノ酸	AA	Y	N
アミノ酸, 数	AA. CNT	Y	N
出願国	AC	Y	Y
出願日	AD	Y	N
出願情報	AI	Y <sup>3)</sup>	Y
出願番号	AP	Y	Y
出願年	AY	Y	N
所属機関 (特許出願人)	CS	Y <sup>5)</sup>	Y
クロスリファレンス	CR	Y	N
(関連するDGENEファイルのレコードへの)	XR	Y <sup>6)</sup>	N
データの入力日	DED	Y	Y
資料種類	DT	Y	Y
入力日	ED	Y <sup>7)</sup>	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	N	Y <sup>16)</sup>
発明者	IN (AU)	Y	Y
キーワード	KW	Y	N
言語	LA	Y	Y

## SELECT, ANALYZE および SORT フィールド (続き)

フィールド	フィールドコード	ANALYZE/SELECT <sup>1)</sup>	SORT
法的状況	LS	Y	N
法的状況の詳細	LS2	Y	N
分子タイプ	MTY	Y	Y
核酸	NA	Y	N
生物名	ORGN	Y	Y
その他の収録源 (WPIのレコード番号)	OS	Y	Y
出願番号グループ	APPS	Y <sup>8), 9)</sup>	N
特許出願人	PA (CS)	Y	Y
特許出願人コード	PACO	Y	Y
特許発行国	PCS	Y <sup>10)</sup>	Y
特許発行国グループ	PC	Y	Y
特許情報	PI	Y <sup>11)</sup>	Y
特許種別コード	PK	Y	Y
特許番号	PN	Y	Y
	PATS	Y	N
特許中の配列情報の記載位置	PSL	Y	Y
優先権主張国	PRC	Y	Y
優先権主張日	PRD	Y	Y
優先権主張日, 最先のものに限定	PRDF	Y <sup>7)</sup>	Y
優先権出願情報	PRAI	Y <sup>12)</sup>	Y
優先権出願番号	PRN	Y	Y
優先権主張年	PRY	Y	Y
優先権主張年, 最先のものに限定	PRYF	Y	Y
特許発行日	PD	Y	Y
特許発行年	PY	Y	Y
配列 (1文字コード)	SEQ	Y <sup>7), 13)</sup>	N
配列 (3文字コード)	SEQ3	Y <sup>7), 13)</sup>	N
配列長	SQL	Y	Y
相同性のスコア	SCORE	N	Y <sup>14)</sup>
処理コード	TC	Y <sup>15)</sup>	Y
標題	TI	Y (デフォルト)	Y
更新日	UP	Y <sup>7)</sup>	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 で抽出されたタームに /BI が付与されます。
- 3) 特許出願番号が SELECT または ANALYZE され、SELECT で抽出されたタームには /AP が付与されます。
- 4) SELECT で抽出されたタームに /IN が付与されます。
- 5) SELECT で抽出されたタームに /PA が付与されます。
- 6) SELECT で抽出されたタームに /CR が付与されます。
- 7) SELECT HIT および ANALYZE HIT はこのフィールドでは使えません。
- 8) ダウエント形式で特許番号、出願番号、および優先権出願番号を SELECT するときには、矢印プロンプト (=>) で SET PATENT DERWENT と入力します。
- 9) AP と PRN が SELECT または ANALYZE され、SELECT で抽出されたタームには /APPS が付与されます。
- 10) PI フィールドから国名コードが SELECT または ANALYZE され、SELECT で抽出されたタームには /PC が付与されます。
- 11) 特許番号が SELECT または ANALYZE され、SELECT で抽出されたタームには /PN が付与されます。
- 12) 優先権出願番号が SELECT または ANALYZE され、SELECT で抽出されたタームには /PRN が付与されます。
- 13) SELECT で抽出されたタームに /SQSP が付与されます。
- 14) GETSIM で作成した L 番号を指定します。
- 15) SELECT で抽出されたタームに /DT が付与されます。
- 16) BLAST/GETSIM ホモロジー検索で作成した L 番号に対して使用します。

## サンプルレコード

## TRIAL 形式

AN AAD05559 cDNA DGENE  
 TI Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition -  
 DESC Human secreted protein-encoding gene 19 cDNA clone HYASC80, SEQ ID NO:78.  
 KW Human; secreted protein; proliferative disorder; cancer; tumour; foetal abnormality; developmental abnormality; haematopoietic disorder; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; inflammation; allergy; neurological disorder; Alzheimer's disease; Parkinson's disease; cognitive disorder; schizophrenia; asthma; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; angiogenic disorder; kidney disorder; gastrointestinal disorder; pregnancy-related disorder; gene therapy; endocrine disorder; infection; wound healing; vulnerary; cell culture; chemotaxis; food additive; binding partner identification; ss.

## SQIDE 形式

AN AAW79650 peptide DGENE  
 AA 0 A; 2 R; 0 N; 2 D; 0 B; 2 C; 0 Q; 0 E; 0 Z; 2 G; 0 H; 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 2 S; 0 T; 0 W; 0 Y; 0 V; 0 Others  
 SQL 10  
 SEQ  
 1 crgdsrgdsc

## FEATURE TABLE:

Key	Location	Qualifier
Modified-site	1	note  "N-alpha-acetyl-Cys"
Modified-site	2..6	note  "optionally one of Arg(2) and Arg(6) is in the form of MeArg"
Modified-site	10	note  "Cys-NH2"
Disulfide-bond	1..10	

## SQ3IDE 形式

AN AAW79650 peptide DGENE  
 AA 0 A; 2 R; 0 N; 2 D; 0 B; 2 C; 0 Q; 0 E; 0 Z; 2 G; 0 H; 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 2 S; 0 T; 0 W; 0 Y; 0 V; 0 Others  
 SQL 10  
 SEQ3  
 1 Cys-Arg-Gly-Asp-Ser-Arg-Gly-Asp-Ser-Cys

## FEATURE TABLE:

Key	Location	Qualifier
Modified-site	1	note  "N-alpha-acetyl-Cys"
Modified-site	2..6	note  "optionally one of Arg(2) and Arg(6) is in the form of MeArg"
Modified-site	10	note  "Cys-NH2"
Disulfide-bond	1..10	



## IALL 表示形式

**ACCESSION NUMBER:** AAD05559 cDNA DGENE  
**TITLE:** Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition -  
**INVENTOR:** Soppet D R; Komatsoulis G; Shi Y; Olsen H S; Ruben S M  
**PATENT ASSIGNEE:** (HUMA-N)HUMAN GENOME SCI INC.  
**PATENT INFO:** WO 2001034767 A2 20010517 540  
**APPLICATION INFO:** WO 2000-US30036 20001101  
**PRIORITY INFO:** US 1999-163576 19991105  
 US 2000-221366 20000727  
**PAT. SEQ. LOC:** Claim 1; Page 459  
**DATA ENTRY DATE:** 18 JUL 2001 (first entry)  
**DOCUMENT TYPE:** Patent  
**LANGUAGE:** English  
**OTHER SOURCE:** 2001-316492 [33]  
**CROSS REFERENCES:** P-PSDB: AAE01738  
**DESCRIPTION:** Human secreted protein-encoding gene 19 cDNA clone HYASC80, SEQ ID NO:78.  
**KEYWORD:** Human; secreted protein; proliferative disorder; cancer; tumour; foetal abnormality; developmental abnormality; haematopoietic disorder; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; inflammation; allergy; neurological disorder; Alzheimer's disease; Parkinson's disease; cognitive disorder; schizophrenia; asthma; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; angiogenic disorder; kidney disorder; gastrointestinal disorder; pregnancy-related disorder; gene therapy; endocrine disorder; infection; wound healing; vulnerary; cell culture; chemotaxis; food additive; binding partner identification; ss.  
**ORGANISM:** Homo sapiens.  
**ABSTRACT:** AAD05492-AAD05564 represent cDNAs corresponding to 22 human secreted protein genes, and AAE01672-AAE01743 represent the proteins they encode. AAE01744-AAE01763 represent human secreted protein fragments or variants. The secreted proteins and their genes are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 22 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein-encoding cDNA of the invention.  
**NUC. ACID COUNTS:** 8 A; 16 C; 20 G; 14 T; 2 other  
**SEQUENCE LENGTH:** 60  
**SEQUENCE**  
 1 ggggatggga acccaaggct gtccacatcc cagctggctg mtmcttctgg

## IALL 形式 (続き)

51 ggctgtcttg

## FEATURE TABLE:

Key	Location	Qualifier	
CDS	5..58	*tag= a	
		product	"Human secreted protein precursor"
		transl_except	(pos:41..43, aa:Xaa)
		note	"Xaa = any of the naturally occurring L-amino acids. CDS does not include stop codon"
			!/partial