



Presenting

STN[®] Patents & Pizza Sessions

*Speedy Delivery of STN
Search Techniques!*

USGENE[®] on STN[®]

The USGENE Genetic Sequence Database

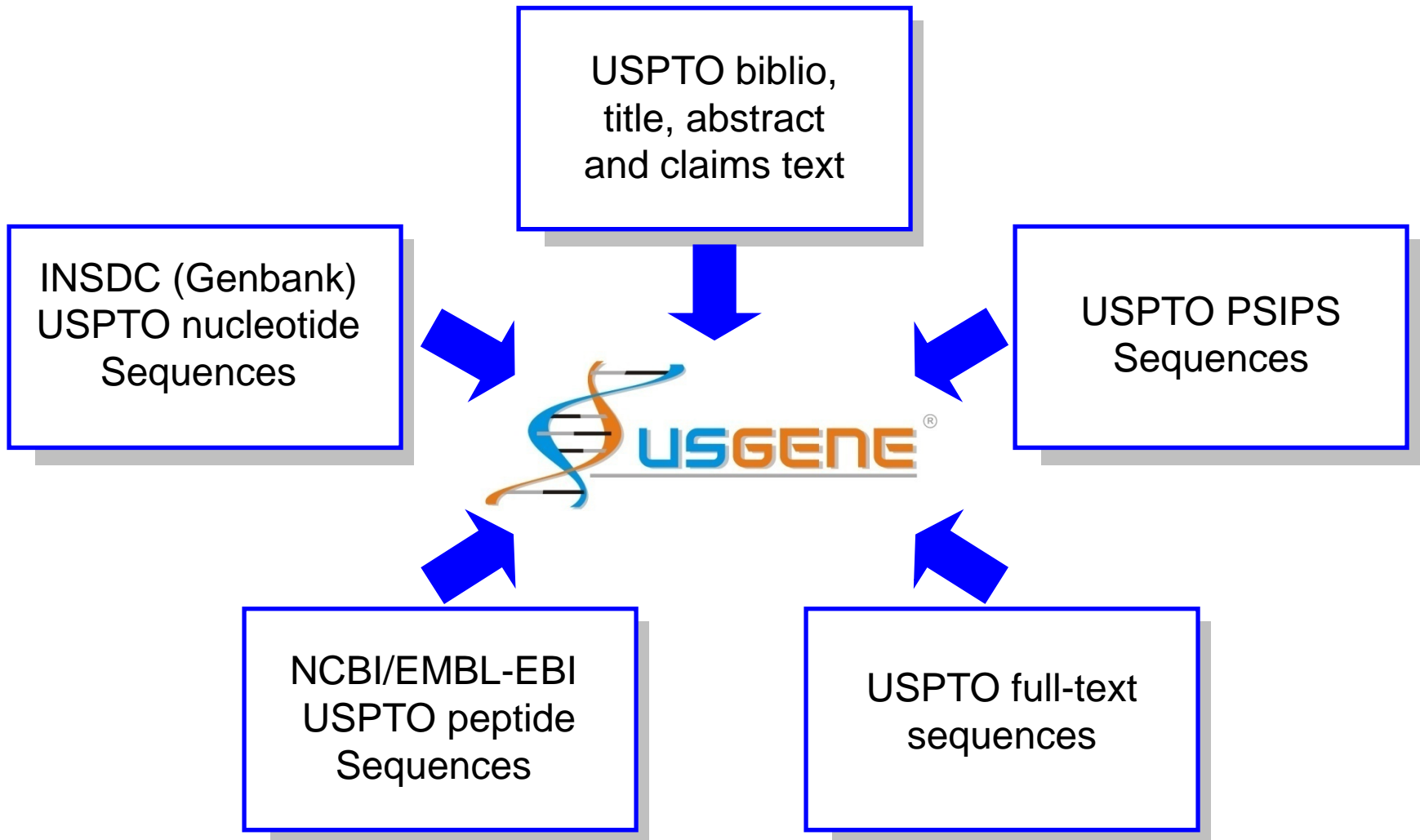
- Sequences from all relevant USPTO published patent applications and granted (issued) patents
- Assignee and full inventor names; publication, application and parent case PCT numbers and dates; original publication **title**, **abstract**, and **claims**
- Organism name, sequence length, Molecule Type, SEQ ID, and feature tables for features/annotations
- Produced by the SequenceBase Corporation
- Updated weekly – within **3 days** of publication
- 1982 – present

USGENE consolidates unique USPTO sequence data from different sources

- USPTO Publication Site for Issued and Published Sequences (PSIPS)
 - The official mega-publication download site, 2001-date
- International Nucleotide Sequence Database Collaboration (INSDC) (NCBI/EMBL/DDBJ, Genbank)
 - U.S. granted patent nucleotide sequences, 1982-date
- USPTO Protein Database (NCBI/EMBL)
 - U.S. granted patent protein/peptide sequences, 1982-date
- USPTO Published Applications and Patents Full-Text
 - Filling in omissions, coverage gaps and to enhance timeliness

The USGENE Sequence Source (/SSO) field indicates which source any given USGENE sequence record was derived from.

USGENE combines sequences with bibliographic data and claims text



USGENE includes all available sequences from U.S. published applications

```

AN      20110002942.45 ← Protein          USGENE
TI      Engineered Anti-IL-23 Antibodies (PublishedApp
IN      Presta Leonard G. (San Francisco, CA); Beyer B
        Ingram Richard N. (Kenilworth, NJ); Orth Peter
        Yan-Hui (Kenilworth, NJ)
PA      Schering Corporation (Kenilworth NJ)
PI      US 20110002942          A1      20110106
AI      US 2010-882035          20100914
RLI     US 2006-511635          20060829
PRAI    US 2005-713585P        20050831
PSL     SEQ ID NO 45
DESC    Mus musculus protein; sequence 45 of 93
DT      Patent
AB      Engineered antibodies to human IL-23p19 are provided, as well as uses
        thereof, e.g. in treatment of inflammatory, autoimmune, and
        proliferative disorders.
ECLM    US20110002942 A1: 1. A binding compound that binds to human IL-23,
        comprising: at least one antibody light chain variable region, or
        binding fragment thereof, comprising at least one CDR sequences . . .
SSO     PROTEIN; USPTO; APPLICATION
ORGN    Mus musculus
SQL     115
SEQ     1  evqlqqsgpe  lvkpgasvkm  sckasg
        51  inpnnggtny  nqkfkkgkatl  tvnkss
        101 dldywgqgtt  ltvss
    
```

AN 20110002942.45 is
SEQ ID NO: 45 from
US20110002942.

Each USGENE sequence
record includes searchable
bibliographic details.

AN 20110002942.45 is displayed
here in **BRIEF** format, which includes
the Exemplary Claim (**ECLM**).

USGENE includes all available sequences from U.S. issued (granted) patents

```
L1 ANSWER 1 OF 1 USGENE COPYRIGHT 2011 SEQUENCEBASE CORP on STN
AN 7863412.3 ← Protein USGENE
TI Method for mixing object into gelled assembly (
IN Uesugi Koji (Aichi, JP); Nagai Yusuke (Aichi, J
PA Menicon Co Ltd (Aichi JP)
PI US 7863412 B2 20110104
US 20100227398 A1 20100909
AI US 2010-705914 20100215
PRAI JP 2009-54982 20090309
XPD 20300215 (calculated)
PSL SEQ ID NO 3
DESC Artificial protein; Self-assembling peptide; sequence 3 of 3
DT Patent
AB This invention provides a mixing method by which a mixing object can
be uniformly mixed into a gelled assembly within a short . . .
ECLM US7863412 B2: 1. A method for mixing a mixing object into a gelled
assembly, the method comprising: freezing the gelled assembly;melting
the frozen assembly to obtain a sol;mixing the resultant sol . . .
SSO PROTEIN; USPTO; GRANTED
ORGN Artificial
SQL 14
SEQ
1 rldlrlalrl dlrl
```

AN 7863412.3 is
SEQ ID NO: 3 from
US7863412.

Each USGENE sequence
record includes searchable
bibliographic details.

AN 7863412.3 is displayed here
in **BRIEF** format, which includes
the Exemplary Claim (**ECLM**).

USGENE also has an extensive backfile

United States Patent: 5210028 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=

Most Visited STN International STN/CAS Home Page STN on the Web STN Viewer STN Viewer

Escherichia coli LC13 / transformed with pCL.sub.85 / and pPLMu/IGFII

SEQUENCE LISTING (1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 4 (2) INFORMATION FOR SEQ ID NO:1: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 67 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (iii) HYPOTHETICAL: N (iv) ANTI-SENSE: N (vi) ORIGINAL SOURCE: (A) ORGANISM: Human IGF-II (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
AlaTyrArgProSerGluThrLeuCysGlyGlyGluLeuValAspThr 151015 LeuGlnPheValCysGlyAsp ArgGlyPheTyrPheSerArgProAla 202530
SerArgValSerArgArgSerArgGlyIleValGluGluCysCysPhe 35 4045 ArgSerCysAspLeuAlaLeuLeuGluThrTyrCysAlaThrProAla 505560 LysSerGlu 65 (2)

INFORMATION FOR SEQ ID NO:2: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 216 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: N (iv) ANTI-SENSE: N (vi) ORIGINAL SOURCE: (A) ORGANISM: Synthetic Human IGF-II DNA Sequence; E.coli pref (B) STRAIN: pBB8/IGF-II (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2: CATATGGCATAACG CCCGAGCGAGACCCCTGTGCGGTGGCGAGCTCGTAGACACTCTGCAG60
TTCGTTTGTGGTGACCGTGGCTTCTACTTCTCTCGTCCTGCTAGCCGTGTATCTCGCCGT120
TCTAGAGGCATCGTTGAAGAGTGCTGTTTCCGCAGCTGTG
TACTGCGCAACTCCAGCAAATCCGAATAAGGATCC216 (2) INFORMATION FOR SEQ ID NO:3: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 233 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: N (iv) ANTI-SENSE: N (vi) ORIGINAL SOURCE: (A) ORGANISM: Human IGF-II DNA Sequence; E.coli pref (B) STRAIN: pIGF-II/3 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3: GAATTCGACGCTTATGGCTTACAGACCATCCGAAACCTTGT
CACCTTGCAATTCGTTTGTGGTGACAGAGGTTT CTA
TTCTAGAAGATCCAGAGGTATCGTTGAAGAATGTTGTTTCA
GTTGAAACCTACTGTGCTACCCAGCTAAGTCTGAATGAATGCGTCGAATTC233 (2) INFORMATION FOR SEQ ID NO:4: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 150 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: N (iv) ANTI-SENSE: N (vi) ORIGINAL SOURCE: (A) ORGANISM: ner gene ribosomal binding site of phage Mu (B) STRAIN: pPLMu (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4: GAATT
CTTACACTTAGTTAAATTGCTAACTTTATAGATTACAAA
ACCTTAGGAGGGTTTTT60
ACCATGGTTACGAATCCCGGGGATCCGTCGACCTGCAGCCAAGCTTGGCTGCCTCGCGC120
GTTTCGGTGATGACGGTGAAAACCTCTGAC 150

Done

M 2

This USPTO example is US5210028, which was issued on May 11th, 1993.

Published sequence data like this are identified, extracted, standardized and loaded into USGENE on STN (compare this to the STN record on the next slide).

To facilitate precise searching all USGENE sequences are in STN standardized format

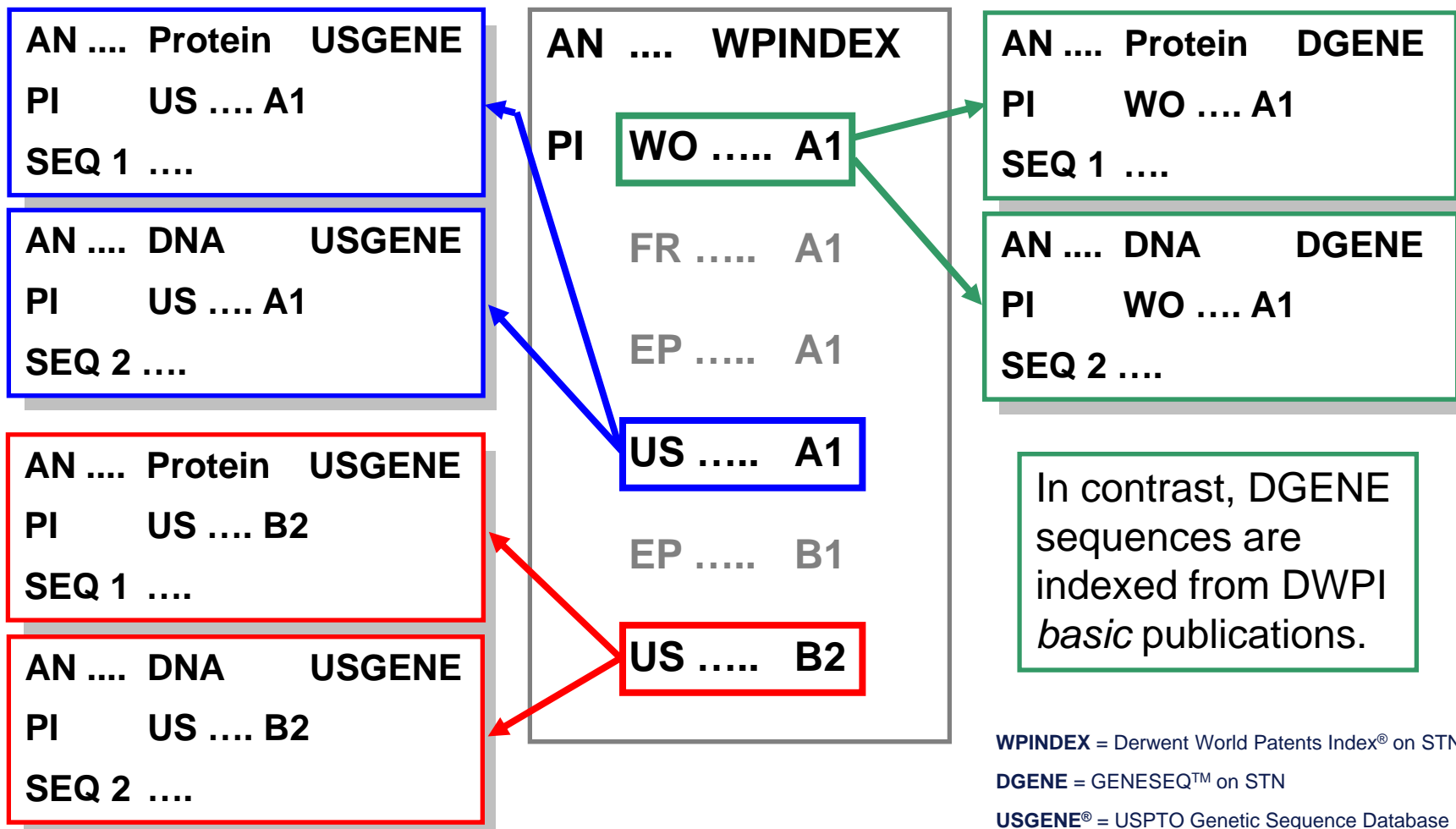
```
L1 ANSWER 1 OF 1 USGENE COPYRIGHT 2008 SEQUENCEBASE CORP on STN
AN 5210028.1 protein USGENE
TI Process for the production of unfused IGF-II (Patent)
IN Schmitz Albert (Basel, CH); Marki Walter (M
PA Ciba Geigy Corporation (Ardsley NY)
PI US 5210028 A 19930511
AI US 1990-616470 19901121
AB A process for the preparation of a recombinant IGF-II (rIGF-II)
without a covalently attached foreign protein moiety and without N-
terminal attached methionine or a derivative of methionine or of a
salt of said IGF-II, rIGF-II produc
ECLM US5210028 A: We claim:1. A process
recombinant IGF-II without a covale
moiety and without N-terminal attac
methionine, said process comprising
of E. coli, said strain being a lon.sup.- and htpR.sup.- double
mutant, with a hybrid vector comprising an expression cassette
consisting of the following elements in the 5' to 3' direction, said
elements which are operably linked: an inducible promoter, a
ribosomal binding site, and the cod
linked in proper reading frame to a
IGF-II having the amino acid sequenc
SSO PROTEIN; USPTO; GRANTED
ORGN Human IGF II
SQL 67
SEQ 1 ayrpsetlcg gelvdtlqfv cgdrgyfysr pasrvsrrsr giveeccfrs
51 cdlalletyc atpakse
```

AN 5210028.1 is SEQ ID NO: 1 from US5210028.

AN 5210028.1 is displayed here in BRIEF format, which includes the Exemplary Claim (ECLM).

Compare the STN standardized USGENE record to the original data source on the previous slide.

USGENE provides sequences from both USPTO published applications and granted patents



Example: sequence listing variation often occurs between publication stages



```
L1 ANSWER 1 OF 1 WPINDEX COPYRIGHT 2011 THOMSON REUTERS on STN
AN 1994-358278 [44] WPINDEX
TI New polynucleotide(s) specific for hepatitis C virus types 4, 5 and 6 -
and related antigenic peptide(s) and
diagnosis, HCV typing and treatment
DC B04; D16; S03
IN PIKE I H; SIMMONDS P; YAP P L
PA (COMM-N) COMMON SERVICES AGENCY; (MURE-N) MUREX DIAGNOSTICS INT INC; . . .
PI WO 9425602 A1 19941110 (199444)* EN 70[5]
AU 9465797 A 1994
FI 9505224 A 1995
EP 698101 A1 1996
JP 09500009 W 1997
AU 695259 B 1998
EP 698101 B1 2004
DE 69434116 E 2004
US 20050032047 A1 20050210 (200512) EN
US 6881821 B2 20050419 (200527) EN
. . . . .
ADT WO 9425602 A1 WO 1994-GB957 19940505 . . . .
PRAI GB 1994-263 19940107
GB 1993-9237 19930505
```

A patent family display from Derwent World Patents Index® on STN.











In this example the DWPI patent family has:

- 9 sequences from [WO9425602](#) in DGENE
- 50 sequences from [US20050032047](#) in USGENE
- 58 sequences from [US6881821](#) in USGENE

How does USGENE compare to other USPTO sequence data sources?

	Update Frequency	Typical Timeliness	Backfile coverage	Value added
USGENE	Weekly	3 days	1982 -	
DGENE (DWPI basics)	Biweekly	65 days	1981 -	
REGISTRY (CAplus basics)	Daily	27 days	1957 -	
NCBI/EMBL	Daily	1-3 months	1982 -	

How does USGENE compare to other USPTO sequence data sources? (cont.)

	USPTO Pub Apps	USPTO Patents	USPTO claims text	Value added
USGENE				
DGENE (DWPI basics)				
REGISTRY (CAplus basics)				
NCBI/EMBL				

There are three main sequence search options for searching USGENE on STN

- BLAST similarity
 - RUN BLAST
- FASTA similarity
 - RUN GETSIM
- Sequence Code Match (SCM)
 - RUN GETSEQ

Note: in this session the focus is on just Sequence Code Match (SCM) searching.

Sequence code match (SCM) searching in USGENE using RUN GETSEQ

- GETSEQ is designed to retrieve either exact matches to a sequence query or answers with conservative variation using special symbols
- It can also be used to retrieve exact length matches or subsequence hits, i.e. where the query is a small part of a larger hit sequence
- GETSEQ can prove to be a fast, precise and effective alternative to BLAST for very short sequence queries, e.g. DNA probes and primers

RUN GETSEQ command syntax

Sequence Code Match (SCM) searching with GETSEQ

=> **RUN GETSEQ L1** (sequence or query L-number)

/SQEP (exact protein) (default)

/SQEFP (exact family protein)

/SQSP (subsequence protein)

/SQSFP (subsequence family protein)

/SQEN (exact nucleotide)

/SQSN (subsequence nucleotide)

Sequence Code Match

- SCM options
 - Exact
 - Exact Family
 - Subsequence
 - Subsequence Family

Sequence Code Match

- SCM options
 - Exact
 - Exact Family
 - Subsequence
 - Subsequence Family

Sequence Code Match: Exact search

- Exact search
 - Matches the sequence query as entered
 - Identical sequences and exact length
- File codes identify the type of sequence to search, e.g.
 - /SQEP = SeQ uence Exa ct P eptide**
 - /SQEN = SeQ uence Exa ct N ucleotide**

EXACT amino acid search (/SQEP) in USGENE

```
=> RUN GETSEQ SMAEP/SQEP
```

```
L1      RUN STATEMENT CREATED
```

```
L1      13 SMAEP/SQEP
```

```
=> D L1 1 SQL SEQ
```

```
L1 ANSWER 1 OF 13 USGENE COPYRIGHT 2011 SEQUENCEBASE CORP on STN
```

```
SQL      5  
SEQ      1 smaep  
=====  
HITS AT: 1-5
```

In all sequence databases,
the typed order of the display
fields will be the order that the
fields are displayed.

```
=> D SEQ3
```

```
L1 ANSWER 1 OF 13 USGENE COPYRIGHT 2011
```

```
SEQ3     1 Ser-Met-Ala-Glu-Pro
```

```
HITS AT: 1-5
```

In USGENE, amino acids
can be displayed as single
letter codes or three letter
codes.

EXACT nucleic acid search (/SQEN) in USGENE

```
=> RUN GETSEQ GCCGCCGT/SQEN
```

```
L1      RUN STATEMENT CREATED
```

```
L1      3 GCCGCCGT/SQEN
```

```
=> D L1 1 SEQ SQL
```

```
L1 ANSWER 1 OF 3 USGENE COPYRIGHT 2011 SEQUENCEBASE CORP on STN  
SEQ
```

```
      1 gccgccgt
```

```
      =====
```

```
HITS AT: 1-8
```

```
SQL     8
```

Sequence Code Match

- **SCM**
 - Exact
 - **Exact Family**
 - Subsequence
 - Subsequence Family

Sequence Code Match: Exact Family search

- Exact Family search
 - Matches the sequence query as entered and allows family substitution to occur
 - Retrieves identical sequences and family sequences with exact length
 - Family substitutions only occur for proteins and not nucleic acids

/SQEFP = Sequence Exact Family Peptide

Amino acid families for RUN GETSEQ SQEFP and QSFP search options

GROUP	AMINO ACIDS
Neutral-Weak Hydrophobics	P, A, G, S, T
Acid Amines-Hydrophilic	Q, N, E, D, B, Z
Basic-Hydrophilic	H, K, R
Hydrophobics	I, M, L, V
Aromatic	F, W, Y
Cross-Linking	C

EXACT FAMILY amino acid search (/SQEFP) in USGENE

```
=> RUN GETSEQ SMAEP/SQEFP
```

```
L1      RUN STATEMENT CREATED
```

```
L1      75 SMAEP/SQEFP
```

```
=> D L1 3 SEQ SQL
```

```
L1 ANSWER 3 OF 75  USGENE COPYRIGHT 2011 SEQUENCEBASE CORP on STN
```

```
SEQ      1 gites
```

```
=====
```

```
HITS AT:  1-5
```

```
SQL      5
```

In USGENE:

SMAEP/SQEP retrieved 13 records.

SMAEP/SQEFP retrieved 75 records.

Possible family
substitutions for SMAEP:

S	M	A	E	P
P	I	G	Q	A
A	L	T	N	G
G	V	P	D	S
T		S	B	T

Sequence Code Match

- **SCM**
 - Exact
 - Exact Family
 - **Subsequence**
 - Subsequence Family

Sequence Code Match: Subsequence search

- Subsequence search
 - Retrieves exact answers plus sequences that are embedded in longer sequence

/SQSP = SeQ uence S ubsequence P eptides

/SQSN = SeQ uence S ubsequence N ucleotides

SUBSEQUENCE amino acid search (/SQSP) in USGENE

```
=> RUN GETSEQ KGPSYSLR/SQSP
```

```
L1 RUN STATEMENT CREATED
```

```
L1 284 KGPSYSLR/SQSP
```

```
=> D L1 19 SEQ SQL
```

```
L1 ANSWER 19 OF 284 USGENE COPYRIGHT 2011 SEQUENCEBASE CORP on STN  
SEQ
```

```
1 kgpsyslrst tmmirpldf
```

```
=====
```

```
HITS AT: 1-8
```

```
SQL 19
```

SUBSEQUENCE nucleic acid search (/SQSN) in USGENE

```
=> RUN GETSEQ ACCCTGCAAA TAGCA/SQSN
```

```
L1 RUN STATEMENT CREATED
```

```
L1 164 ACCCTGCAAATAGCA/SQSN
```

```
=> D L1 10 SEQ SQL
```

```
L1 ANSWER 10 OF 164 USGENE COPYRIGHT 2011 SEQUENCEBASE CORP on STN  
SEQ
```

```
1 accctgcaaa tagcagaa  
===== =====
```

```
HITS AT: 1-15
```

```
SQL 18
```

STN will ignore spaces
within the query sequence.

Sequence Code Match

- **SCM**
 - Exact
 - Exact Family
 - Subsequence
 - **Subsequence Family**

Sequence Code Match: Subsequence Family search

- Subsequence Family search
 - Exact sequence match, subsequence match, and sequences that contain family substitution of amino acid

/SQSFP = SeQ uence S ubsequence F amily P eptides

SUBSEQUENCE FAMILY amino acid search (/SQSFP) in USGENE

=> RUN GETSEQ KGPSYSLR/SQSFP

L1 RUN STATEMENT CREATED

L1 5460 KGPSYSLR/SQSFP

=> D L1 4 SEQ SQL

L1 ANSWER 4 OF 5460 USGENE COPYRIGHT 2011 SEQUENCEBASE CORP on STN SEQ

1 raapygvrlc grefiravif tcggsrwrrs dilaheamgd tfpdadaded
 =====

51 slageldeam gssewlaltk spqa

101 sscckwgcsk seisslc

HITS AT: 1-8

SQL 117

In USGENE:

KGPSYSLR/SQSP

retrieved 284 records.

KGPSYSLR/SQSFP

retrieved 5460 records.

Possible family substitutions for KGPSYSLR:

<u>K</u>	<u>G</u>	<u>P</u>	<u>S</u>	<u>Y</u>	<u>S</u>	<u>L</u>	<u>R</u>
H	P	A	T	F	P	I	H
R	A	G	P	W	A	M	K
	S	S	A		G	V	
	T	T	G		T		

Summary of the Sequence Code Match options

Search Type	Polypeptides	Nucleic Acids
EXACT	/SQEP	/SQEN
EXACT FAMILY	/SQEFP	—
SUBSEQUENCE	/SQSP	/SQSN
SUBSEQUENCE FAMILY	/SQSFP	—

Special variability symbols allow flexibility in sequence motif searching

- Allow users to specify motif patterns of different amino acid(s) at one location of a sequence
- Provide the ability to specify gaps of an unknown number of amino acids
- Provide the ability to search for motif patterns at the beginning or end of a sequence
- Allow users to specify the number or range of repeats for amino acid(s) or gaps

Note: a complete table of all variability symbols, with search examples, is given in the USGENE database summary sheet:

http://www.stn-international.com/stndatabases/databases/onlin_db.html

Variability symbols for RUN GETSEQ sequence code match searches

<u>Symbol</u>	<u>Function</u>
[]	Specify alternate residues
[-]	Exclude a specific residue or alternate residues
{ }	Repeat the preceding symbol(s) (number or range)
?	Repeat the preceding symbol(s) zero or one time
*	Repeat the preceding symbol(s) zero or more times
+	Repeat the preceding symbol(s) one or more times
^	Query appears at the beginning or the end of a sequence
	Alternate sequence expressions
.	A gap of one residue
:	A gap of zero or one residues
&	Concatenate (join together) sequence queries

Case study: using SCM variability symbols to search USGENE

Search Question:

Find U.S. patent references disclosing one or more of the sequences represented by this Markush peptide sequence formula:

LGPX₁QLCX₂LVX₃CAP

X₁ = V or L

X₂ = any amino acid except, G or H

X₃ = any amino acid

RUN GETSEQ SCM in USGENE search strategy

⇒ **RUN GETSEQ LGP[VL]QLC[-GH]LV.CAP/SQSP**

– Possible sequence retrieval

- *LGPVQLCALVHCAP*
- *LGPVQLCSLVVCAP*
- *LGPLQLCVLVACAP*
- *LGPLQLCPLVTCAP*

Run the GETSEQ SCM search in USGENE

=> FILE USGENE

=> RUN GETSEQ LGP[VL]QLC[-GH]LV.CAP/SQSP

RUN GETSEQ AT 21:34:43 ON 13 JAN 2011
COPYRIGHT (C) 2011 FIZ KARLSRUHE GMBH

L1 RUN STATEMENT CREATED
L1 63 LGP[VL]QLC[-GH]LV.CAP/SQSP

=> D TRI SEQ

L14 ANSWER 1 OF 63 USGENE COPYRIGHT 2011 SEQUENCEBASE CORP on STN
TI Compositions and Methods for Treating Tumors, Fibrosis, and Pulmonary
Alveolar Proteinosis (PublishedApplication)
DESC Homo sapiens protein; sequence 8 of 30
MTY Protein
SQL 417
SEQ 1 mrfawtvlll gplqlcalvh cappaagqqq :
= =====
51 ngqvfsllsl gsqyqpqrrr dpgaavpgaa nasaqqprtp illirdnrta
101 aartrtagss gvtagrprpt arhwfqagys tsrareagas raenqtapge . . .
HITS AT: 10-23

63 sequence hits (L1) have been found in USGENE containing the sequence fragment(s) of interest.

The hit portion of the answer sequence is highlighted with double underlining.

Sequence code match (SCM) searching in USGENE in BATCH mode

- Allows searcher to have system run searches while user is offline
- Up to 16 BATCH searches can be run on one STN ID
- Searches can be run concurrently
- Email can be sent by the system to let searcher know when search is complete
- Simply add the word BATCH to the end of the search string

SUBSEQUENCE FAMILY amino acid BATCH search example in USGENE

```
=> RUN GETSEQ KGPSYSLR/SQSFP BATCH
```

```
PLEASE ENTER BATCH IDENTIFIER (MAX. 8 CHARS):USGENE3
```

```
TO BE NOTIFIED WHEN THIS BATCH SEARCH IS COMPLETE,  
PLEASE ENTER YOUR EMAIL ADDRESS (MAX. 50 CHARS) OR "NONE"  
INPUT: OR (END):JIM.BROWN@FIZ-K.COM
```

```
RUN GETSEQ AT 21:01:54 ON 13 JAN 2011  
COPYRIGHT (C) 2011 FIZ KARLSRUHE GMBH
```

```
PREVIOUS BATCH REQUEST STILL RUNNING  
BATCH PROCESSING QUEUED FOR USGENE3
```

```
=> LOG Y
```

SUBSEQUENCE FAMILY amino acid search in USGENE (cont.)

USGENE USGENE3 BATCH GETSEQ SEARCH RESULTS ARE READY FOR COLLECTION. - Windows Internet Ex...

https://remote.fiz-karlsruhe.de/owa/DanaInfo=exchange.fiz-karlsruhe.de,SSL+?ae=Item&a=Open&t=IPM.Note&id=RgAf

Reply Reply All Forward

USGENE USGENE3 BATCH GETSEQ SEARCH RESULTS ARE READY FOR COLLECTION.

HelpDesk@FIZ-Karlsruhe.DE

To: Brown, Jim

Attachments: ATT00001.txt (424 B)

Thursday, January 13, 2011 3:11 PM

USGENE USGENE3 BATCH GETSEQ SEARCH RESULTS ARE READY FOR COLLECTION. PLEASE LOG BACK IN TO STN, AND USE THE RUN GETBATCH COMMAND IN FILE USGENE TO COLLECT YOUR SEARCH RESULTS.

THIS IS AN AUTOMATICALLY GENERATED EMAIL.
FOR MORE INFORMATION PLEASE CONTACT HelpDesk@Fiz-karlsruhe.de

Done Internet | Protected Mode: On 100%

Email is sent when search has run to completion.

SUBSEQUENCE FAMILY amino acid BATCH search example in USGENE

=> **FILE USGENE**

FILE 'USGENE' ENTERED AT 21:16:54 ON 13 JAN 2011
COPYRIGHT (C) 2011 SEQUENCEBASE CORP

=> **RUN GETBATCH**

Please enter your batch identifier

or enter # for batch id list

or enter * for batch id at top of list

or enter - before batch id to delete

or enter . for (end)

BATCH REQUEST: #

Batch result files remaining:

USGENE1 Retrieved (getseq)

USGENE2 Retrieved (getseq)

USGENE3 Completed (getseq)

USGENE Retrieved (getseq)

Use # to see list of BATCH
searches and their status.

SUBSEQUENCE FAMILY amino acid BATCH search example in USGENE

```
Please enter your batch identifier  
or enter # for batch id list  
or enter * for batch id at top of list  
or enter - before batch id to delete  
or enter . for (end)
```

```
BATCH REQUEST:USGENE3
```

```
RUN GETSEQ AT 16:44:17 ON 14 JAN 2011  
COPYRIGHT (C) 2011 FIZ KARLSRUHE GMBH
```

```
L1      RUN STATEMENT CREATED  
L1      5460 KGPSYSLR/SQSFP
```

Resources for sequence searching on STN

- *Sequence Searching on STN* modular workshop
<http://www.fiz-k.com/bostonsequenceworkshop>
 - DGENE, USGENE, PCTGEN content and searching
 - CAS REGISTRY and REGISTRY BLAST
 - Sequence Code Match (SCM) searching
 - Multifile search techniques
- USGENE resources, reference materials and FAQ
<http://www.sequencebase.com>

Upcoming 2011 STN Patents and Pizza Sessions

DATE	TOPIC
March 8	Reaction Searching on STN®
April 12	Finding Elusive Substances with CAS Registry Numbers®
May 10	Increase Confidence in Your Search Results
June 14	The Secret Life of CAS REGISTRY SM

learningsolutions.cas.org

Don't forget to register!