

STN[®]

Biosequence Searching on STN

e-Seminar, December 2007

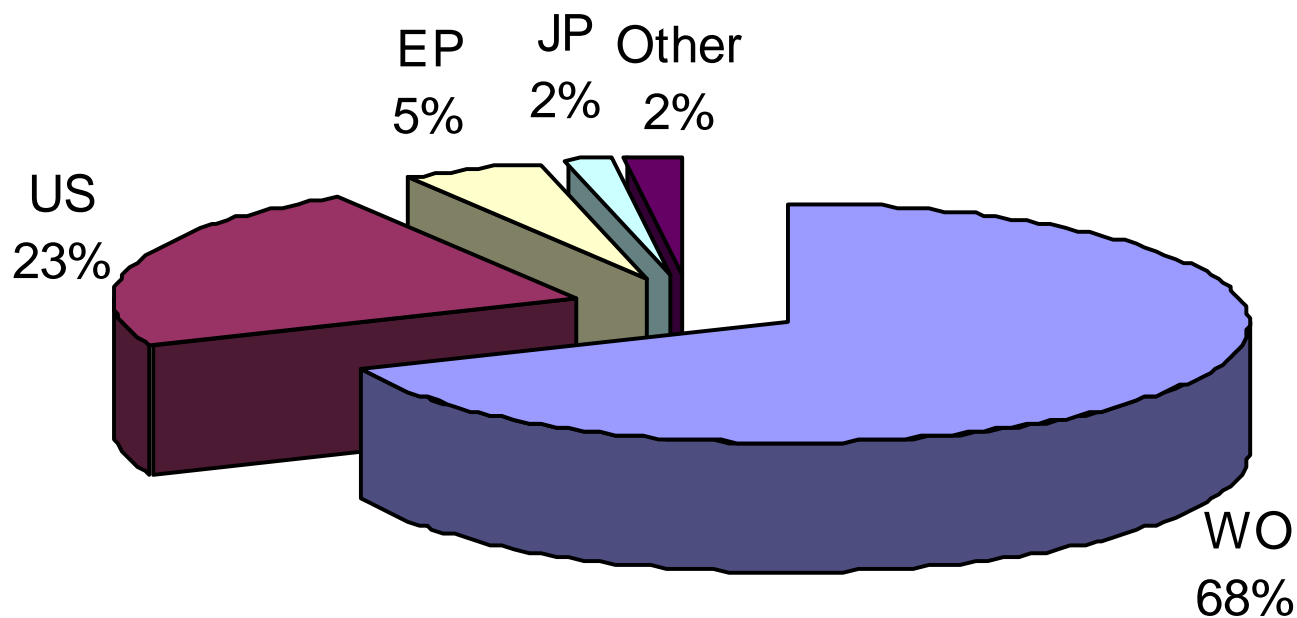
Agenda

- STN Biosequence databases – content
 - DGENE
 - USGENE
 - PCTGEN
- The STN on the Web Sequence Search Assistant
- Sequence search results and processing
- Basic search strategies

DGENE

- Produced by Thomson Scientific
- Sequences from the **basic** patents of the 41 authorities of the Derwent World Patents Index®
- Bibliography, enhanced title, abstract, indexing and patent location provided for each sequence
- Patent Family and Legal Status display
- 10 million patent sequence records (Dec 2007)
- Updated every two weeks
- 1981 - present

DGENE biosequences are indexed from DWPI basic patent publications



Note: the majority of DGENE patent sequences come from WIPO/PCT published patent applications (WO).

Relationship between DWPI patent family & DGENE sequence database

AN WPINDEX
TI
PA
PI FR A1
WO A1
US A1
US B2
AB

AN Protein DGENE
PI FR A1
SEQ 1
AB

AN DNA DGENE
PI FR A1
SEQ 2
AB

AN Peptide DGENE
PI FR A1
SEQ 3
AB

WPINDEX = Derwent World Patents Index® on STN®

DGENE = GENESEQ™ on STN

DGENE records provide

- Enhanced patent titles from DWPISM
- Enhanced English abstracts *per sequence* written by Thomson Scientific experts
 - Including foreign translations (~20% of the database)
- **Basic** patent bibliographic details from DWPI
- Cross-reference Accession Numbers from corresponding DWPI records
- Displayable DWPI patent family and INPADOCDB Legal Status information

How to read a DGENE record

ACCESSION NUMBER: AFS70510 DNA DGENE

TITLE: New isolated nucleic acid molecules comprising a Single Nucleotide Polymorphism, are useful in identifying an individual who has an altered risk for developing stenosis.

INVENTOR: Cargill M; Devlin J J; Luke M

PATENT ASSIGNEE: (APPL-N)APPLERA CORP.

PATENT INFO: WO 2004058990 A2 20040715 129

APPLICATION INFO: WO 2003-US40977 20031222

PRIORITY INFO: US 2002-434741P 20021220
US 2003-453050P 20030310
US 2003-466437P 20030430

PAT. SEQ. LOC: Claim 13; SEQ ID NO 23166

DATA ENTRY DATE: 20 SEP 2007 (first entry)

DOCUMENT TYPE: Patent

LANGUAGE: English

OTHER SOURCE: 2004-543463 [52]

DESCRIPTION: Human SNP-containing genomic DNA fragment SEQ ID No 23166.

How to read a DGENE record

KEYWORD: ds; SNP; single nucleotide polymorphism; SNP detection; diagnosis; prophylaxis; therapeutic; stenosis; vasotropic.

ORGANISM: Homo sapiens.

ABSTRACT:

The present invention relates to an isolated nucleic acid molecule comprising at least 8 contiguous nucleotides where one of the nucleotides comprises a single nucleotide polymorphic site. The invention further relates to an isolated polypeptide encoded by the nucleic acid; an antibody (or antigen-binding fragment) that specifically binds to a polypeptide; an isolated polynucleotide which specifically hybridizes to a nucleic acid molecule of the invention; a kit for detecting an SNP comprising a polynucleotide which specifically hybridizes to a nucleic acid molecule of the invention; a method for identifying an individual who has an altered risk for developing stenosis; a method for detecting an SNP in a nucleic acid molecule; a method for detecting a variant polypeptide; and a method for identifying an agent useful in the treatment or prophylaxis of stenosis. The present sequence is that of a genomic DNA-based nucleic acid sequence showing a stenosis-associated SNP of the invention.

NUC. ACID COUNTS: 61 A; 52 C; 44 G; 43 T; 0 U; 1 Other

SEQUENCE LENGTH: 201

How to read a DGENE record

SEQUENCE

```
1 tcactacttc tctgaccata actgatcact aaagccgctt tgtcattacc
51 atgtaagcca ccacaaaaca acaatttata gctcatagaa aaatctatag
101 magcaggctg ggctcggggg ctcacacctg taatcccagg actttgagag
151 gccaaaggcgg gcggatcaca aggtcaggag atcgagacca tcctggctaa
201 t
```

FEATURE TABLE:

Key	Location	Qualifier	
====+	====+	====+	====+
variation	101	*tag= a	
		standard_name	"Single nucleotide
			polymorphism"

DGENE patent family display (FAM)

```
L1 ANSWER 1 OF 1 DGENE COPYRIGHT 2007 THE THOMSON CORP on STN
PI WO 2004058990 A2 20040715 (200452)* EN 129[1]
  US 20040166519 A1 20040826 (200457) EN
  AU 2003299817 A1 20040722 (200476) EN
  EP 1583770 A2 20051012 (200567) EN
  AU 2003299817 A8 20051110 (200634) EN
  US 20070031847 A1 20070208 (200713) EN
ADT WO 2004058990 A2 WO 2003-US40977 20031222; US 20040166519 A1 Provisional
  US 2002-434741P 20021220; US 20040166519 A1 Provisional US 2003-453050P
  20030310; US 20040166519 A1 Provisional US 2003-466437P 20030430; AU
  2003299817 A1 AU 2003-299817 20031222; AU 2003299817 A8 AU 2003-299817
  20031222; EP 1583770 A2 EP 2003-800091 20031222; US 20040166519 A1 US
  2003-741601 20031222; EP 1583770 A2 WO 2003-US40977 20031222; US
  20070031847 A1 Provisional US 2003-453050P 20030310; US 20070031847 A1
  Provisional US 2003-466437P 20030430; US 20070031847 A1 US 2004-796280
  20040310
FDT AU 2003299817 A1 Based on WO 2004058990 A; EP 1583770 A2 Based
  on WO 2004058990 A; AU 2003299817 A8 Based on WO 2004058990 A
PRAI US 2003-466437P 20030430
  US 2002-434741P 20021220
  US 2003-453050P 20030310
  US 2003-741601 20031222
  US 2004-796280 20040310
```

DGENE legal status display (LS)

```
LEGAL STATUS  INPADOCDB  COPYRIGHT 2007 EPO / FIZ KARLSRUHE on STN
AN      AFS70510 DGENE
20040715 WOAK      + DESIGNATED STATES
                        WO                      A2
                        AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR
                        CU CZ DE DK DM DZ EC EE EG ES FI GB GD GE GH GM HR HU ID
                        IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG
                        MK MN MW MX MZ NI NO NZ OM PG PH PL PT RO RU SC SD SE SG
                        SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
20040715 WOAL      + DESIGNATED COUNTRIES FOR REGIONAL PATENTS
                        WO                      A2
                        BW GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW AM AZ BY KG KZ
                        MD RU TJ TM AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU
                        IE IT LU MC NL PT RO SE SI SK TR BF BJ CF CG CI CM GA GN
                        GQ GW ML MR NE SN TD TG
20040908 WO121     EP: THE EPO HAS BEEN INFORMED BY WIPO THAT EP WAS
                        DESIGNATED IN THIS APPLICATION
20050617 WOENP     ENTRY INTO THE NATIONAL PHASE IN:
                        JP 2005510048          A
20050617 WOWWE     + WIPO INFORMATION: ENTRY INTO NATIONAL PHASE
                        JP 2005510048
                        200744.....20071102
```

PCTGEN

- Produced by WIPO / FIZ Karlsruhe
- Sequences submitted electronically as a formal part of WIPO/PCT patent applications
- Information **as given by patent applicant**
- Bibliography and original publication title provided for each sequence
- Publication, application and related application numbers and dates
- 5.7 million sequence records (Dec 2007)
- Updated weekly
- August 2001 - present

Relationship between PCTFULL and PCTGEN databases

AN ... PCTFULL

TI

PA

PI WO A1

AB

DETD

CLM

AN Protein PCTGEN

PI WO A1

SEQ 1

AN DNA PCTGEN

PI WO A1

SEQ 2

AN Peptide PCTGEN

PI WO A1

SEQ 3

PCTFULL = WIPO/PCT patent applications full-text

PCTGEN = WIPO/PCT patent application biosequences

How to read a PCTGEN record

ACCESSION NUMBER: 2007127522.84 DNA PCTGEN
TITLE: MODULATING ALKALOID BIOSYNTHESIS
PATENT ASSIGNEE: Ceres, Inc.
PATENT INFO: WO 2007127522 20071108
REL APPL INFO: US 2006-360039 20060222; US 2005-654927P 20050222
FILE UPDATE DATE: 20071108
DOCUMENT TYPE: Patent
ORGANISM: Arabidopsis thaliana
SEQUENCE LENGTH: 996
SEQUENCE

Related application information

```
1 gtcgattgga tgatgaacat tctacatata taattattat gtttaagcac
51 ttagacagca taaattcttt ctaattatat aaatctaacc ttggttacatt
101 gtacatctat aaattacttg aagaaataac gagttctatt tctttttaaa
151 aattaaaaat actataccat atctcagtga ttaagttgaa ccaaaaggta
201 cggaggagaa acaagcattt gattcttcct tattttattt tattcatctc
251 tactaatga tgggtggagaa aaaaagaaaa tacctaacaa acaaatatat
301 attgtcatac aaaaatattt ctatatTTTT agttaattag tttatattcc
351 tcacttttca gggcttatat aagaaagtga gcaaacacaa atcaaaatgc
```

.....

How to read a PCTGEN record

```
401 agcagcaaat actatcatca cccatctcct tagttctatt ttataattcc
451 tcttcttttt gttcatagct ttgtaattat agtcttattt ctctttaagg
501 ctcaataaga ggaggtacta ttactacact tctctctact tttacttgta
551 ttttagcatt aaaatcctaa aatccgtttt aaattcaaaa ataaacttag
601 agatgtttta tctcgattcg gtttttcggc tttaggagaa taattatatg
651 aaattagtat ggatatcttt actagtttcc attcaaatga ttctgatttc
701 aatctaatac tctcactctt taattaaact atatgtagtg taatttcaca
751 ctgttaaatt tctaccatgt catgtatatt agagttgcat agaaaattgt
801 aaaacatcca tttgaattcg aatgaaacaa aatgttttaa aataaaattt
851 tggtttttaa aagaaaaatc taaaactgaa ttatatcgtt taaccaagtt
901 gtaaaagtca taaaacgtag tatcttgtaa atcgctcttc cacggtccaa
951 atagacttct agtaataaac aagtaaaact aattttggtt tcttac
```

FEATURE TABLE:

```
Key          |Location  |
=====+=====+=====
misc_feature|(1)..(996)|Ceres Promoter YP0115
```

PCTGEN original sequence (SEQO)

SEQO

```

acatcacatc tttcactctg acccctggac gtgcgcccac ctgcctgctc tgcggggcgcc      60
caaccttttc tcaacttctg ggcaagatgc ccggttgaag ggacagtgcc tggcgaactg      120
ggagggagag gggacgaggg cggaacctca gccacttctc ccctccctta gagtcacaat      180
caaaagcttg ggtgaggaca aacagacctc agtgggtcga gcccttacia aactcgcccc      240
agtgggccag caaggcgagc gtcactgcgg ggcctctcca atctgtgtcc cgccggggac      300
atggacgcct gagaacctgt gggtcggtta ggtcacagaa cacagaaaac accgggactc      360
gggaccgtc atg cac tcc aag act gct cca agg ttc ctg gtg ttc ctg ctg      411

```

Met His Ser Lys Thr Ala Pro Arg Phe Leu Val Phe Leu Leu

1

5

10

```

ctt acc ctg ctg cta ctc ctg gct gcc tct cct gtg gct agc aag ggc      459
Leu Thr Leu Leu Leu Leu Leu Ala Ala Ser Pro Val Ala Ser Lys Gly

```

15

20

25

30

```

tgt gtc tgc aaa ggc aaa gga caa tgc ctc tgc gcg ggg acc aaa ggg      507
Cys Val Cys Lys Gly Lys Gly Gln Cys Leu Cys Ala Gly Thr Lys Gly

```

35

40

45

```

gag aag ggg gag aaa ggg gtt cct ggt tcc cct gga ttt cct ggc cag      555
Glu Lys Gly Glu Lys Gly Val Pro Gly Ser Pro Gly Phe Pro Gly Gln

```

50

55

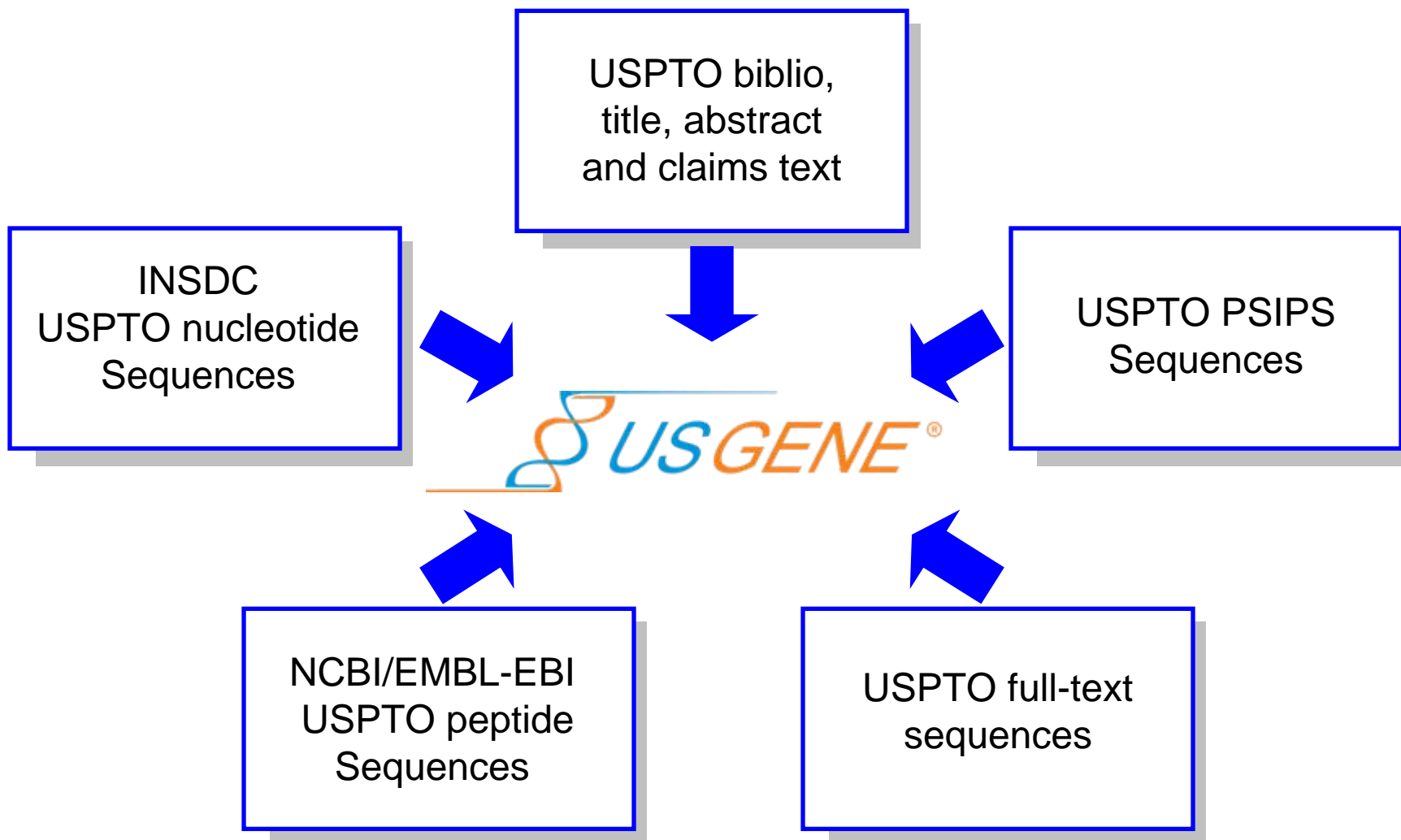
60

.....

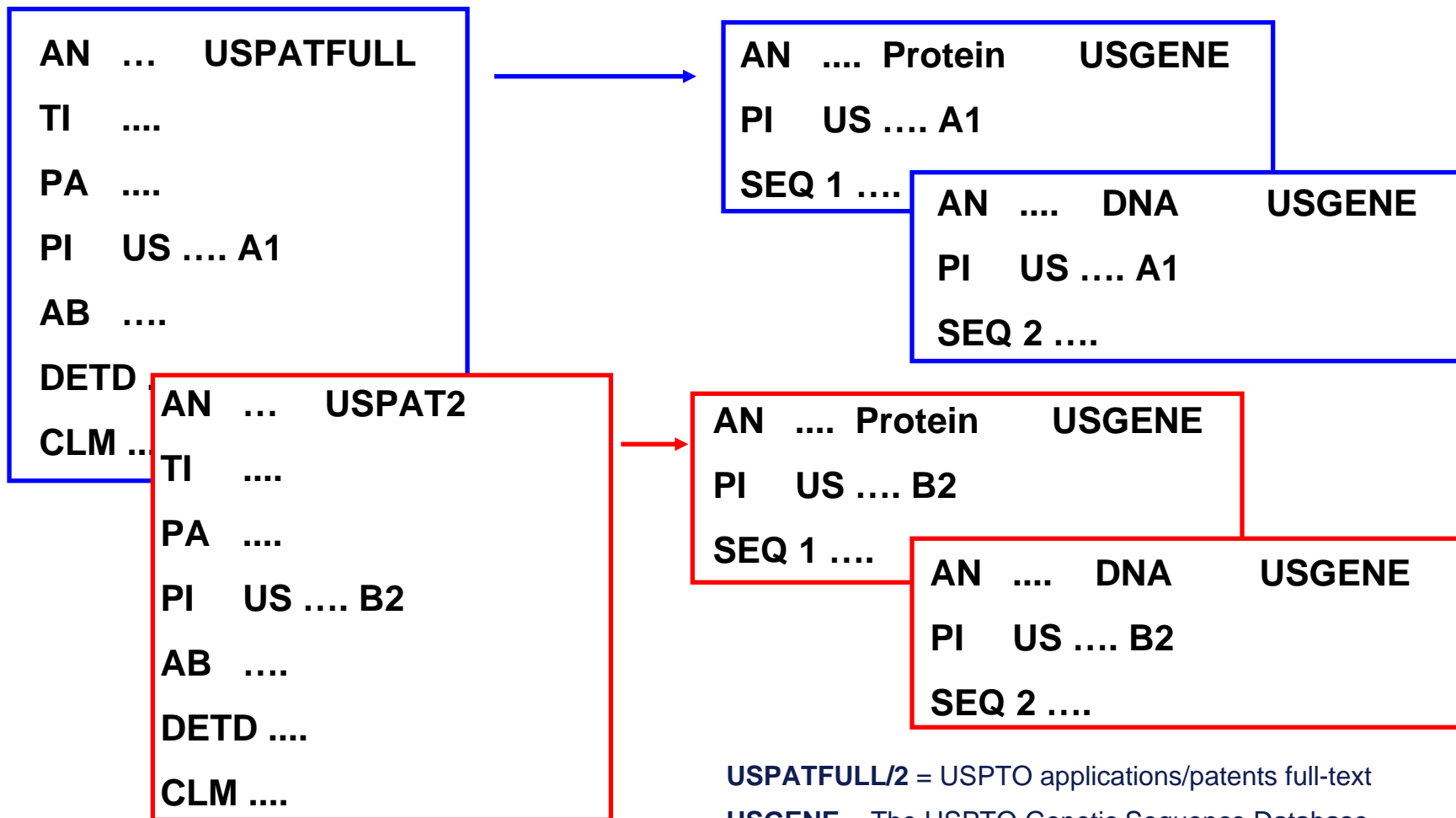
USGENE®

- Produced by the SequenceBase Corporation
- Sequences from all relevant USPTO published patent applications and issued (granted) 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, SEQ ID, molecule type, and feature tables for features/modifications
- Updated weekly – within **7 days** of publication
- 1982 – present

USGENE[®] is compiled technologically from several disparate sources



Relationship between USPATFULL/2 and USGENE[®] databases



USPATFULL/2 = USPTO applications/patents full-text

USGENE = The USPTO Genetic Sequence Database

How to read a USGENE record

ACCESSION NUMBER: 7261897.30 Protein USGENE

TITLE: Fusion proteins of mycobacterium tuberculosis (Patent)

INVENTOR(S): Skeiky Yasir (Bellevue, WA); Reed Steven (Bellevue, WA);
Alderson Mark (Bainbridge Island, WA)

PATENT ASSIGNEE: Corixa Corporation (Seattle WA)

PATENT INFORMATION: US 7261897 B2 20070828
US 20060171965 A1 20060803

APPLICATION INFO: US 2006-396117 20060330

ENTRY DATE: 20070831

DOCUMENT TYPE: Patent

ABSTRACT: The present invention relates to compositions and fusion proteins containing at least two Mycobacterium sp. antigens, and nucleic acids encoding such compositions and fusion proteins. The compositions of the invention increase serological sensitivity of sera from individuals infected with tuberculosis, and methods for their use in the diagnosis, treatment, and prevention of tuberculosis infection.

How to read a USGENE record

CLAIM: US7261897 B2: 1. A fusion polypeptide comprising a first amino acid sequence having at least 95% sequence identity to SEQ ID NO:8 and a second amino acid sequence having at least 95% sequence identity to SEQ ID NO:2.

2. The polypeptide of claim 1, further comprising a third amino acid sequence having at least 95% sequence identity to SEQ ID NO:6.

3. The polypeptide of claim 1, which has at least 95% sequence identity to SEQ ID NO:10 or SEQ ID NO:12.

4. The polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8 and the amino acid sequence of SEQ ID NO:2.

5. The polypeptide of claim 4, further comprising the amino acid sequence of SEQ ID NO:6.

6. The polypeptide of claim 4 or 5, further comprising a histidine tag.

.....

How to read a USGENE record

SEQUENCE SOURCE: PROTEIN; USPTO; GRANTED
ORGANISM NAME: Mycobacterium tuberculosis
SEQUENCE LENGTH: 340

SEQUENCE:

```
1 mtffeqvrrl rsaattlprl vaiaamgavl vyglvgtfgg patagafsrp
51 glpveylqvp sasmgrdikv qfqqgggphav ylldglraqd dyngwdintp
101 afeeyyqsgl svimpvggqs sfytdwyqps qsngqnytyk wetfltremf
151 awlqankgvs ptgnaavgls msggsalila ayypqqfpya aslsgflnps
201 egwwptligl amndsggyna nsmwgpssdp awkrndpmvq iprlvantr
251 iwvycngntp sdlggdnipa kflegltlrt nqtfrdtyaa dggrngvfnf
301 ppngthswpy wneqlvamka diqhvlngat ppaapaapaa
```

FEATURE TABLE:

Key	Location
	85 complex antigen (MTB85 complex antigen)

Sequence searching methods

- GETSEQ for simple sequence queries (Sequence Code Match)
- GETSIM for advanced similarity (homology) searching (based on FastA algorithm)
- BLAST® for advanced similarity (homology) searching (based on Blast® algorithm)

BLAST is a registered trademark of the National Library of Medicine.

BLAST software is used with permission of the National Center for Biotechnology Information (NCBI) of the National Library of Medicine (NLM).

Comparison of homology algorithms

BLAST®	FASTA
Faster than FASTA	Slower than BLAST
Equivalent for highly similar sequences	
Misses some less similar sequences	Better for less similar sequences
Comparison of shorter sequence parts	Comparison of entire sequence length
Less sensitive when using default settings	More sensitive, misses less homologs
Less separation between true homologs and random hits	More separation between true homologs and random hits
Calculates probabilities	Calculates significance “on the fly” from the given dataset

Sequence searching methods

- Offline BATCH homology sequence searching (e.g. for GETSIM searches)
- Current awareness ALERT feature

Interfaces for sequence searching

- DGENE, PCTGEN and USGENE can be searched in two ways
 - In commandline mode via STN Express or STN on the Web (with RUN command and three search options, Getseq, Getsim, Blast)
 - Via Sequence Search Assistant and Results Assistant in STN on the Web (with point & click access)

The Sequence Search Assistant

STN on the web

- About
- First Time User
- Free Search Preview
- Customer Support
 - Get Structure Plug-in
 - Get Sequence Plug-in
 - Standard Login
 - Academic Login
 - Secure Session
 - STN Links
 - Feedback



[New BLAST Plug-in available \(August 12, 2007\)](#)

Login ID

Password

Transcript: OFF
 ON:

**Power of STN —
Convenience of the Web!**

STN's complete functionality for searching an unparalleled collection of essential science and technology databases.

[Browser requirements](#): Windows: Netscape Navigator 4.x and higher or MSIE 4.x and higher; Macintosh: Netscape 4.x or MSIE 4.x. Javascript must be enabled.



[Impressum](#)

The Sequence Search Assistant

STN on the web

- Help
- News
- Search Assistants
 - Search Assistant
 - Patent Search
 - Alert (SDI) Asst.
 - Structure Query
 - Sequence Asst.
 - Upload Sequence
 - Upload Cmd. File
- Results Assistant
- Transcript Assistant
- => Command Line
- Logoff Hold
- Logoff
- Feedback
- Send Break

Transcript: ON

[STN Command List](#)
[File-Specific Help List](#)

```
***** Welcome to STN International *****
NEWS 22 Sep 05 INPADOCDB: New SDI frequency MONTHLY available now
NEWS 23 Sep 07 FORIS renamed to SOFIS
NEWS 24 Sep 10 STN AnaVist, Version 2.0, now available with
Derwent World Patents Index
NEWS 25 Sep 18 CA/CaPlus enhanced with printed CA page images from
1967-1998
NEWS 26 Sep 19 CaPlus coverage extended to include traditional
medicine patents
NEWS 27 Sep 24 EMBASE, EMBAL, and LEMBASE reloaded with enhancements
NEWS 28 Oct 02 CA/CaPlus enhanced with pre-1907 records from Chemisches
Zentralblatt
NEWS 29 Oct 22 BEILSTEIN updated with new compounds
NEWS 30 Nov 18 ICSD (Inorganic Crystal Structure Database) updated
to more than 100,000 records and reloaded with
enhancements
NEWS 31 Nov 18 New display format XMLDOC for WPIX
NEWS 32 Dec 03 New learning file LINPADOCDB available

NEWS EXPRESS 19 Sep 2007: CURRENT WINDOWS VERSION IS
V8.2, CURRENT MACINTOSH VERSION IS V6.0c(ENG)
AND V6.0Jc(JP), AND CURRENT DISCOVER
FILE IS DATED 19 September 2007.

NEWS HOURS Dec 07, 2006 NEW STN OPERATING HOURS

NEWS PRICE Jan 02, 2007 STN International Fees and Prices,
Effective Jan 1, 2007

***** STN Karlsruhe *****
FILE 'HOME' ENTERED AT 17:40:19 ON 03 DEC 2007

=>
```

[Hide session output](#)
[Show session output](#)

The Sequence Search Assistant

STN on the web

- Help
- News
- Search Assistants
 - Search Assistant
 - Patent Search
 - Alert (SDI) Asst.
 - Structure Query
 - Sequence Asst.
 - Upload Sequence
 - Upload Cmd. File
- Results Assistant
- Transcript Assistant
- => Command Line
- Logoff Hold
- Logoff
- Feedback
- Send Break

Sequence Search Assistant

CAS Registry BLAST[®]: [Cost Information](#) [\(Plug-in Required\)](#)

New version of BLAST plug-in available August 12, 2007

Launch CAS Registry BLAST Launch sequence searching and review BLAST Reports ([Help](#)) ([Security](#))

Select STN session status following Launch of CAS Registry BLAST:

- Session will time out in 20 minutes
- LOG-OFF HOLD (session may be resumed within 120 minutes)
- LOG-OFF

Retrieve RNs from BLAST Retrieve a set of RNs previously transferred from a BLAST Report ([Help](#))

DGENE/PCTGEN/USGENE Sequence Searches: Cost Information: [DGENE](#) [PCTGEN](#) [USGENE](#) [\(No Plug-in Required\)](#)

Launch

in DGENE PCTGEN USGENE

Select the Type of Search:

Continue Conduct a menu-driven sequence search using the selected search algorithm in the selected database ([Help](#))

Show Batch Status Show Status Information about Offline Batch Searches ([Help](#))

Show Alert Status Show Status Information about Alert(SDI) Searches ([Help](#))

BLAST is a registered trademark of the National Library of Medicine

Cancel

Transcript:

[STN Command List](#)
[File-Specific Help List](#)

Defining the sequence search

The image displays three sequential screenshots of a web interface for defining sequence searches, with red arrows indicating the progression of steps.

Top Screenshot: The interface is titled "DGENE/PCTGEN/USGENE Sequence Searches:". The "Launch" dropdown menu is open, showing options: "BLAST", "GETSIM", and "GETSEQ". A red arrow points to the "BLAST" option. Below the dropdown, the "in" section has radio buttons for "DGENE", "PCTGEN", and "USGENE", with "DGENE" selected. A "Select the Type of Search:" dropdown is partially visible. Buttons include "Continue", "Show Batch Status", and "Show Alert Status".

Middle Screenshot: The "Launch" dropdown is now closed and set to "BLAST". A red arrow points to the "PCTGEN" radio button, which is now selected. The "Select the Type of Search:" dropdown is open, showing "Protein Search" and "Nucleotide Search". A red arrow points to the "Nucleotide Search" option. Buttons include "Show Batch Status" and "Show Alert Status".

Bottom Screenshot: The "Launch" dropdown is closed and set to "BLAST". The "in" section has radio buttons for "DGENE", "PCTGEN", and "USGENE", with "DGENE" selected. The "Select the Type of Search:" dropdown is closed and set to "Nucleotide Search". A red arrow points to the "Continue" button. Buttons include "Continue", "Show Batch Status", and "Show Alert Status".

Uploading the query sequence

STN on the web

- Help
- News
- Search Assistants
 - Search Assistant
 - Patent Search
 - Alert (SDI) Asst.
 - Structure Query
 - Sequence Asst.
 - Upload Sequence
 - Upload Cmd. File
- Results Assistant
- Transcript Assistant
- => Command Line
- Logoff Hold
- Logoff
- Feedback
- Send Break

Sequence Search Assistant

DGENE BLAST Search

There are three options to enter the sequence query ([Help](#)):

1. Enter the nucleotide sequence (FASTA format recommended):

Continue

(max. sequence length 200 characters)

2. Read from File (Note [valid sequence formats!](#))

Upload File

(max. sequence length 10.000 characters)

3. previously uploaded

Valid Sequence Format!

- For successfully uploading sequences please use the following sequence format:
 - ASCII txt file
 - each line may not exceed 300 characters (essential: carriage return at least every 300 characters)
 - FASTA format
 - no blank lines before the sequence start

Valid Sequence Format!

```
gcacccggca gcggtctcag gccaaagcccc ctgccagcat ggccagcgag ttcaagaaga
agctcttctg gagggcagtg gtggccgagt tcctggccac gacctcttt gtcttcatca
gcatcggttc tgccctgggc ttcaaatacc cggtggggaa caaccagacg gcggtccagg
acaacgtgaa ggtgtcgtg gccttcgggc tgagcatcgc cacgctggcg cagagtgtgg
gccacatcag cggcgccac ctcaaccgg ctgtcacact ggggctgctg ctcagctgcc
agatcagcat cttcctgccc ctcatgtaca tcatcgccca gtgcgtgggg gccatcgtcg
ccaccgcat cctctcagge atcacctcct ccctgactgg gaactcgctt ggccgcaatg
acctggctga tgggtgtgaac tcggggcagg gcctgggcat cgagatcatc gggaccctcc
agctggtgct atgcgtgctg gctactaccg accggaggcg ccgtgacctt ggtggctcag
cccccttgc catcggcctc tctgtagccc ttggacacct cctggctatt gactacactg
gctgtgggat taaccctgct cggtcctttg gctccgcggt gatcacacac aacttcagca
accactggat tttctgggtg gggccattca tcgggggagc cctggctgta ctcactacg
acttcatcct ggccccacgc agcagtgacc tcacagaccg cgtgaagggtg tggaccagcg
gccaggtgga ggagtatgac ctggatgccg acgacatcaa ctccagggtg gagatgaagc
ccaaatagaa ggggtctggc ccgggcatcc acgtaggggg caggggcagg ggcgggcgga
gggaggggag gggtgaaatc catactgtag aactctgac aagctggcca aagtcacttc
cccaagatct gccagacctg catggtcaag cctcttatgg ggggtgtttct atctctttct
.....
```

Uploading the query sequence

Sequence Search Assistant

DGENE BLAST Search

Query: L1 (nucleotide sequence uploaded from a local file)

```
gcacccggca gcggtctcag gccaaagcccc ctgccagcat ggccagcgag tcaagaaga agctctctg gagggcagtg gtggccgagt tctggccac gaccctttt  
gtctcatcagcatcgggtc tgcctgggc tcaaatacc cggtggggaa caaccagacg gcggtccaggacaacgtgaa ggtgtcgtg gccttcgggc tgagcatcg cagctggcg  
cagagtgtgggccacatcag cggcggccac ctaaccgg ctgtcacact gggctgtg ctcagtgccagatcagcat ctccgtgcc ctcatgtaca tcatgcccga gtgcgtggg  
gccatcgtgccaccgccaat cctctcaggc atcaactct cctgactgg gaactcgct ggccgcaatgacctggctga tgggtgaac tcgggccagg gcctgggcat ...
```

Select the search option ([Help](#)):

Select the search mode ([Help](#)): [Advanced parameters](#) (optional)

BLAST is a registered trademark of the National Library of Medicine

New Search


Cancel

Search

Defining search modes

Select the search option ([Help](#)): both strands ▾
Select the search mode ([Help](#)): **both strands**
single strand
compl. strand

[Advanced parameters](#) (optional)



Select the search option ([Help](#)): both strands ▾
Select the search mode ([Help](#)): **online**
offline
Alert (SDI)

[Advanced parameters](#) (optional)



Defining search options

https://stnweb.fiz-karlsruhe.de/html/english/SQ_AdvNucSeaOpt.html - Microsoft Internet Explorer

Advanced DGENE Blast Search Options

Select the **Expect Value**:

Select the **Word Size**:

Select the **Filter**:

Enter the **Gap Cost values**:

Gap Penalty:

Gap Extension:

Penalty for Mismatch:

Reward for Match:

[Reset Form to Defaults](#)
[Reset Form to saved Settings](#)

BLAST is a registered trademark of the National Library of Medicine

Note: Settings will only be saved for current search !

RUN BLAST advanced options

Expectation Value (-E)

Expectation value (E-Value) is the statistical significance threshold for reporting matches against a sequence database. The E-value can be any positive number, and the default value is 10. This means that 10 matches may be expected to be found merely by chance. In general E-value is lowered to make the search more precise and raised to retrieve more answers.

Word Size (-W)

Word Size is the length of the character string fragments of a sequence query which are used as the basis for a BLAST search. For SQN the default is 11 and the range 7-23. For all other BLAST searches the default is 3 and the range 2-3. For short search queries, reducing the default word size can give improved search results.

RUN BLAST advanced options (cont.)

Low Complexity Filtering (on by default) (-F)

The low complexity filter can eliminate biologically uninteresting segments that have low compositional complexity and are statistically significant, as determined by specific programs for peptide or nucleotide sequences in nature. Filtering is applied to the query sequence and is indicated by a series of Xs for peptide sequences and Ns for nucleotide sequences. Low complexity filtering can be turned off (i.e. set to F - false).

Peptide similarity matrices (-M)

For peptide based searches SQP and TSQN the advanced options provide additional scoring matrices to the default BLOSUM62 (next slide)

Guidelines from NCBI on the use of Advanced Settings for peptide sequence searching are as follows:

<u>Query Length</u>	<u>Matrix</u>	<u>Gap costs</u>
<35	PAM-30	(9,1)
35 – 50	PAM-70	(10,1)
50 – 85	BLOSUM-80	(10,1)
>85	BLOSUM-62	(11,1) (BLAST default)

Starting the sequence search

Sequence Search Assistant

DGENE BLAST Search

Query: L1 (nucleotide sequence uploaded from a local file)

```
gcaccggca gcggtctcag gccaaagcccc ctgccagcat ggccagcgag tcaagaaga agctctctg gagggcagtg gtggccgagt tctggccac gacccttt  
gtctcatcagcatcgggtc tgcctgggc tcaaatacc cgggtgggaa caaccagacg gcggtccaggacaacgtgaa ggtgtcgtg gcctcgggc tgagcatgc cacgctggc  
cagagtgtgggcccacatcag cggcggccac ctaaccgg ctgtcact ggggtcgtg ctacgtgccagatcagcat ctccgtgc ctcatgtaca tcacgcca gtgcgtggg  
gccatcgtgccaccgcat ctctcaggc atcactct cctgactgg gaactcgtt gcccgcaatgacctggctga tgggtgaac tggggccagg gctgggcat ...
```

Select the search option ([Help](#)):

Select the search mode ([Help](#)): [Advanced parameters](#) (optional)

BLAST is a registered trademark of the National Library of Medicine

New Search

Cancel

Search 

Processing sequence search results

Sequence Search Assistant

```
L2      RUN STATEMENT CREATED
L2      382 GCACCCGGCAGCGGTCTCAGGCCAAGCCCCCTGCCAGCATGGCCAGCGAG
        TTCAAGAAGAAGCTCTTCTGGAGGGCAGTGGTGGCCGAGTTCTTGCCAC
        GACCCTCTTTGTCTTCATCAGCATCGGTTCTGCCCTGGGCTTCAAATACC
```



You have two options for processing the answer set L-Number:

1. **Results Assistant**

The Results Assistant leads you, step by step, through the command language. See the [instructions](#) how to

After result retrieval you can change into the Results Assistant for further processing of the sequence search answer set.

2. **Command Line**


In command line mode you can sort and display answers by using STN commands. See [dgene_help.pdf](#) about DGENE sequence specific issues in command line mode.

Restart Seq. Assist.


The Results Assistant

Results Assistant

1. Select an answer set: [Choose a Different L-Number](#)

[Current answer set:](#) L2 RUN STATEMENT CREATED 

2. Select the desired process (optional steps):

[Sort](#)  Sort the answer set and display selected answers via the Results Assistant.

~~[Duplicate Detection](#)~~

3. Select the desired output option:

[Display](#)

[Print Offline](#)

[Print Email](#)

For email delivery, you must have an [STN Mail Id.](#)

[Cancel](#)

Processing sequence search results

Results Assistant

Sort Options

L2 RUN STATEMENT CREATED

2 sorting options are valid for sequence search results: 1) sort by field code and 2) sort by patent family.

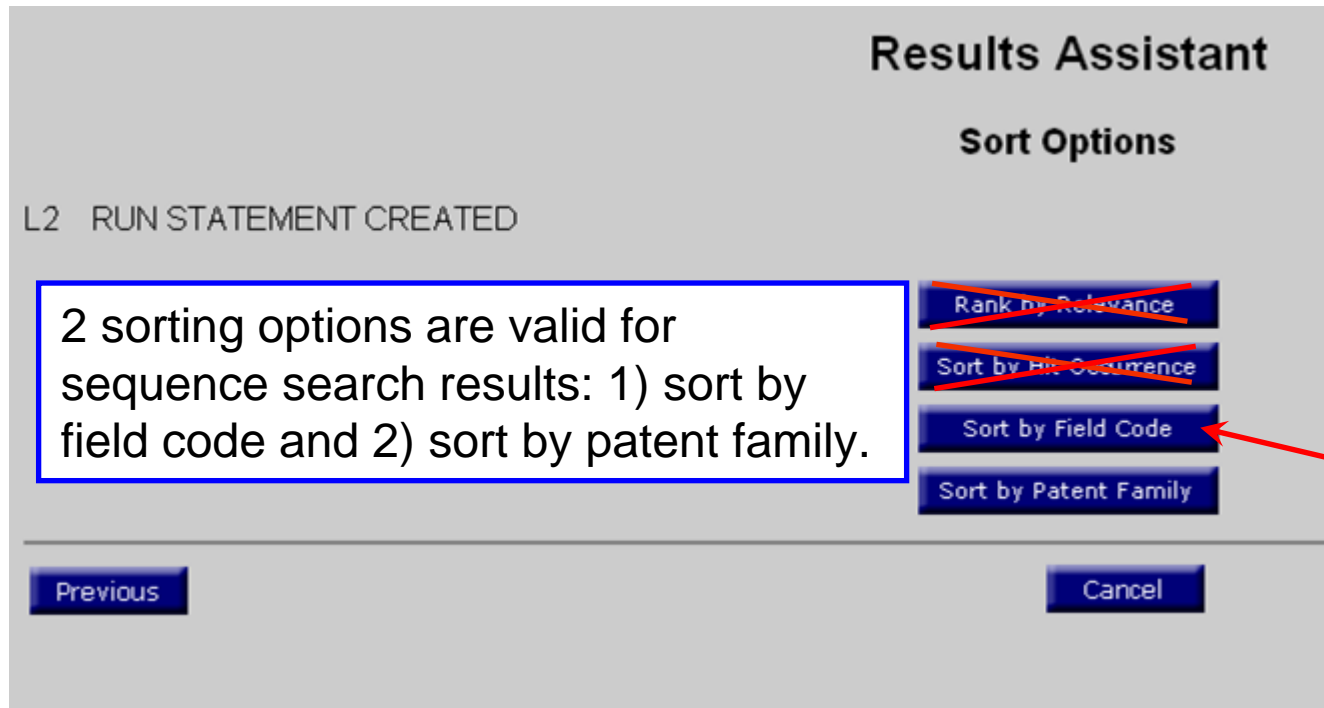
~~Rank by Relevance~~

~~Sort by Hit Occurrence~~

Sort by Field Code

Sort by Patent Family

Previous Cancel



Re-sorting according to SCORE

Results Assistant

L2 RUN STATEMENT CREATED

1. Enter the answer(s):

2. Select the required [sort fields](#):

Sort field 1 (required):

Sort field 2:

Sort field 3:

Sort field 4:

Sort field 5:

Select a field	ASCENDING
PRDF	ASCENDING
PRN	DESCENDING
PRY	ASCENDING
PRYF	ASCENDING
PSL	ASCENDING
SCORE	ASCENDING
SEQ	
SEQ3	
SQL	
TC	
TI	

Previous

Cancel

Continue

Re-sort the complete answer set according to descending SCORE value to get the best answers first.

Re-Sorting according to SCORE

Results Assistant

L2 RUN STATEMENT CREATED

1. Enter the answer(s):


2. Select the required [sort fields](#):

Sort field 1 (required):	<input type="text" value="SCORE"/>	<input type="text" value="DESCENDING"/>
Sort field 2:	<input type="text" value="Select a field"/>	<input type="text" value="ASCENDING"/>
Sort field 3:	<input type="text" value="Select a field"/>	<input type="text" value="ASCENDING"/>
Sort field 4:	<input type="text" value="Select a field"/>	<input type="text" value="ASCENDING"/>
Sort field 5:	<input type="text" value="Select a field"/>	<input type="text" value="ASCENDING"/>

Displaying results

Results Assistant

1. Select an answer set: **Choose a Different L-Number**


Current answer set: L3 382 SORT L2 1- SCORE D 

2. Select the desired process (optional steps):

Sort

~~Duplicate Detection~~

3. Select the desired output option:

Display 

Print Offline

Print Email

Display selected answers in pre-defined formats.

For email delivery, you must have an [STN Mail Id.](#)

Cancel

Displaying results

Results Assistant

Display options

The selected answer set is: L3 382 SORT L2 1- SCORE D

1. Enter answer(s) for Display:

2. Choose the desired [Display format\(s\)](#):

**Free-of-charge display formats:
TRIAL, ALIGN**

```

=> sort L2 1- SCORE D
PROCESSING COMPLETED FOR L2
L3          382 SORT L2 1- SCORE D
=> dis L3 1-20 TRIAL ALIGN

L3  ANSWER 1 OF 382 DGENE  COPYRIGHT 2007 THE THOMSON CORP on STM
AN  AEH10480  cDNA          DGENE
TI  Analyzing differential gene expression associated with histopathologic
     features of colorectal disease, involves detecting overexpression or
     underexpression of pool of polynucleotide sequences in colon tissues.
DESC Aquaporin 1 (AQP1) cDNA SEQ ID 866.
KW  gene expression; prognosis; diagnosis; DNA microarray; colorectal
     disease; colon tumor; colorectal tumor; cytostatic; gastrointestinal
     disease; neoplasm; ss.
SQL  1662
BLASTALIGN
  Query = 1662 letters
  Length = 1662
  Score = 3295 bits (1662), Expect = 0.0
  Identities = 1662/1662 (100%)
  Strand = Plus / Plus

Query: 1   gcacccggcagcgggtctcaggccaagccccctgccagcatggccagcaggttcaagaaga
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 1   gcacccggcagcgggtctcaggccaagccccctgccagcatggccagcaggttcaagaaga

Query: 61  agctcttctggagggcagtggtggccgagttcctggccacgaccctctttgtcttcatca
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 61  agctcttctggagggcagtggtggccgagttcctggccacgaccctctttgtcttcatca

Query: 121 gcatcggttctgccctgggcttcaaataccgggtggggaacaaccagacggcgggtccagg
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 121 gcatcggttctgccctgggcttcaaataccgggtggggaacaaccagacggcgggtccagg

Query: 181 acaacgtgaaggtgtcgctggccttcgggctgagcatcgccacgctggcgcagagtgtgg
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 181 acaacgtgaaggtgtcgctggccttcgggctgagcatcgccacgctggcgcagagtgtgg

Query: 241 gccacatcagcggcgcccacctcaaccggctgtcacactggggctgctgctcagctgcc
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 241 gccacatcagcggcgcccacctcaaccggctgtcacactggggctgctgctcagctgcc

Query: 301 agatcagcatcttccgtgccctcatgtacatcatcgcccagtgcggtgggggccatcgctg

```



Review the sequence search results and check their relevance.

```

L3 ANSWER 20 OF 382 DGENE COPYRIGHT 2007 THE THOMSON CORP on STN
AN ABL64117 DNA DGENE
TI Screening for anti-neoplastic agent involves exposing cells to a chemical
agent to be tested for anti-neoplastic activity, and determining a change
in expression of a gene of a signature gene set -
DESC Breast cancer related gene sequence SEQ ID NO:2454.
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
gene; ds.
SQL 1340
BLASTALIGN
Query = 1662 letters
Length = 1340
Score = 2640 bits (1332), Expect = 0.0
Identities = 1339/1340 (99%), Gaps = 1/1340
Strand = Plus / Plus

Query: 1 gcacccggcagcgggtctcaggccaagccccct
|
Sbjct: 1 gcacccggcagcgggtctcaggccaagccccct

Query: 61 agctcttctggagggcagtggtggccgagttcctggccacgaccctctttgtcttcatca
|
Sbjct: 61 agctcttctggagggcagtggtggccgagttcctggccacgaccctctttgtcttcatca

Query: 121 gcatcgggttctgccctgggcttcaaataaccgggtggggaacaaccagacggcggtccagg
|
Sbjct: 121 gcatcgggttctgccctgggcttcaaataaccgggtggggaacaaccagacggcggtccagg

Query: 181 acaacgtgaaggtgtcgctggccttcgggctgagcatcgccacgctggcgcagagtgtgg
|
Sbjct: 181 acaacgtgaaggtgtcgctggccttcgggctgagcatcgccacgctggcgcagagtgtgg

Query: 241 gccacatcagcggcgcccacctcaaccgggtgtcacactggggctgctgctcagctgcc
|
Sbjct: 241 gccacatcagcggcgcccacctcaaccgggtgtcacactggggctgctgctcagctgcc

Query: 301 agatcagcatcttccgtgccctcatgtacatcatcgccagtgctggggccatcgctcg
|
Sbjct: 301 agatcagcatcttccgtgccctcatgtacatcatcgccagtgctggggccatcgctcg

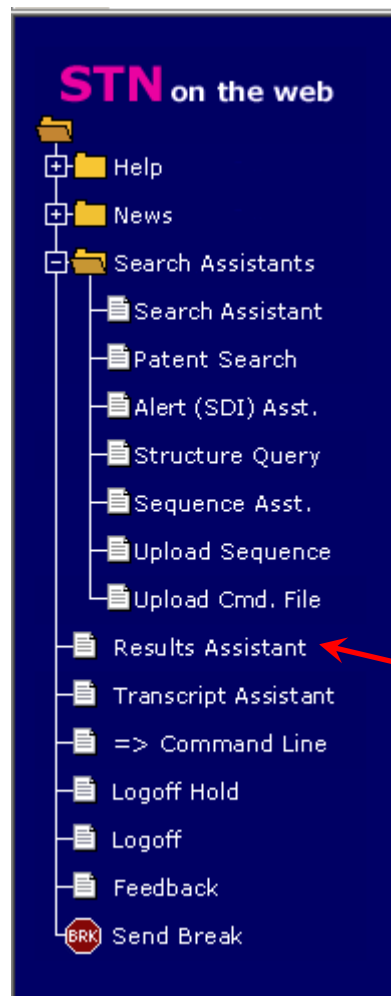
Query: 361 ccaccgccatcctctcaggcatcacctcctcctgactgggaactcgcttggccgcaatg
|
Sbjct: 361 ccaccgccatcctctcaggcatcacctcctcctgactgggaactcgcttggccgcaatg

```



In DGENE the Enhanced Title (TI), Description (DESC) and Keywords (KW) give a clue to the functional relevance of a similar sequence.


Using the Results Assistant




Sorting according to patent family

Results Assistant

1. Select an answer set: [Choose a Different L-Number](#)

[Current answer set:](#) L3 382 SORT L2 1- SCORE D 

2. Select the desired process (optional steps):

[Sort](#)  Add a patent family sort to the SCORE sorted answer set.

~~[Duplicate Detection](#)~~

3. Select the desired output option:

[Display](#)

[Print Offline](#)

[Print Email](#)

For email delivery, you must have an [STN Mail Id.](#)

[Cancel](#)

Sorting according to patent family

L3 382 SORT L2 1- SCORE D


Results Assistant

Sort Options

~~Rank by Relevance~~

~~Sort by Hit Occurrence~~

Sort by Field Code

Sort by Patent Family 

Previous Cancel

Patent family sort

Results Assistant

1. Select an answer set: **Choose a Different L-Number**

Current answer set:

L4 382 FSO L3

https://stnweb.fiz-karlsruhe.de/BASScgi/2/0-01211967000155-1196702954?LN=3&Action=RA_Sorpat - Microsoft Internet Exp...

```
L2 382 SEA FILE=DGENE GCACCCGGCAGCGGTCTCAGGCCAAGCCCCCTGCCAGCATGGCCAGCG
    AGTTCAAGAAGAAGCTCTTCTGGAGGGCAGTGGTGGCCGAGTTCCTGGCCAC
L3 382 SOR L2 1- SCORE D
L4 382 FSO L3
```

70 Multi-record Families Answers 1-306

Family 1	Answers 1-5
Family 2	Answers 6-7
Family 3	Answers 8-10
Family 4	Answers 11-12
Family 5	Answers 13-14
Family 6	Answers 15-16
Family 7	Answers 17-21
Family 8	Answers 22-29
Family 9	Answers 30-32
Family 10	Answers 33-34
Family 11	Answers 35-36
Family 12	Answers 37-38
Family 13	Answers 39-44

https://stnweb.fiz-karlsruhe.de/BASScgi/2/0-01211967000155-1196702954?LN=3&Action=RA_Sorpat

Family 55	Answers 251-260
Family 56	Answers 261-264
Family 57	Answers 265-266
Family 58	Answers 267-268
Family 59	Answers 269-271
Family 60	Answers 272-273
Family 61	Answers 274-275
Family 62	Answers 276-277
Family 63	Answers 278-286
Family 64	Answers 287-288
Family 65	Answers 289-294
Family 66	Answers 295-296
Family 67	Answers 297-299
Family 68	Answers 300-301
Family 69	Answers 302-303
Family 70	Answers 304-306

76 Individual Records Answers 307-382

U Non-patent Records

Displaying patent family sorted answers

Results Assistant

Display options

The selected answer set is: L4 382 FSO L3

Special selection procedure for answers sorted by patent family!

1. Select Family and / or answer(s) to be displayed

Select the Family to be displayed: Family 1 1st member

Enter answer(s) for Display: Family 1 1st member

2. Choose the desired [Display](#)

ABS
ALL
BIB
FAM
IALL

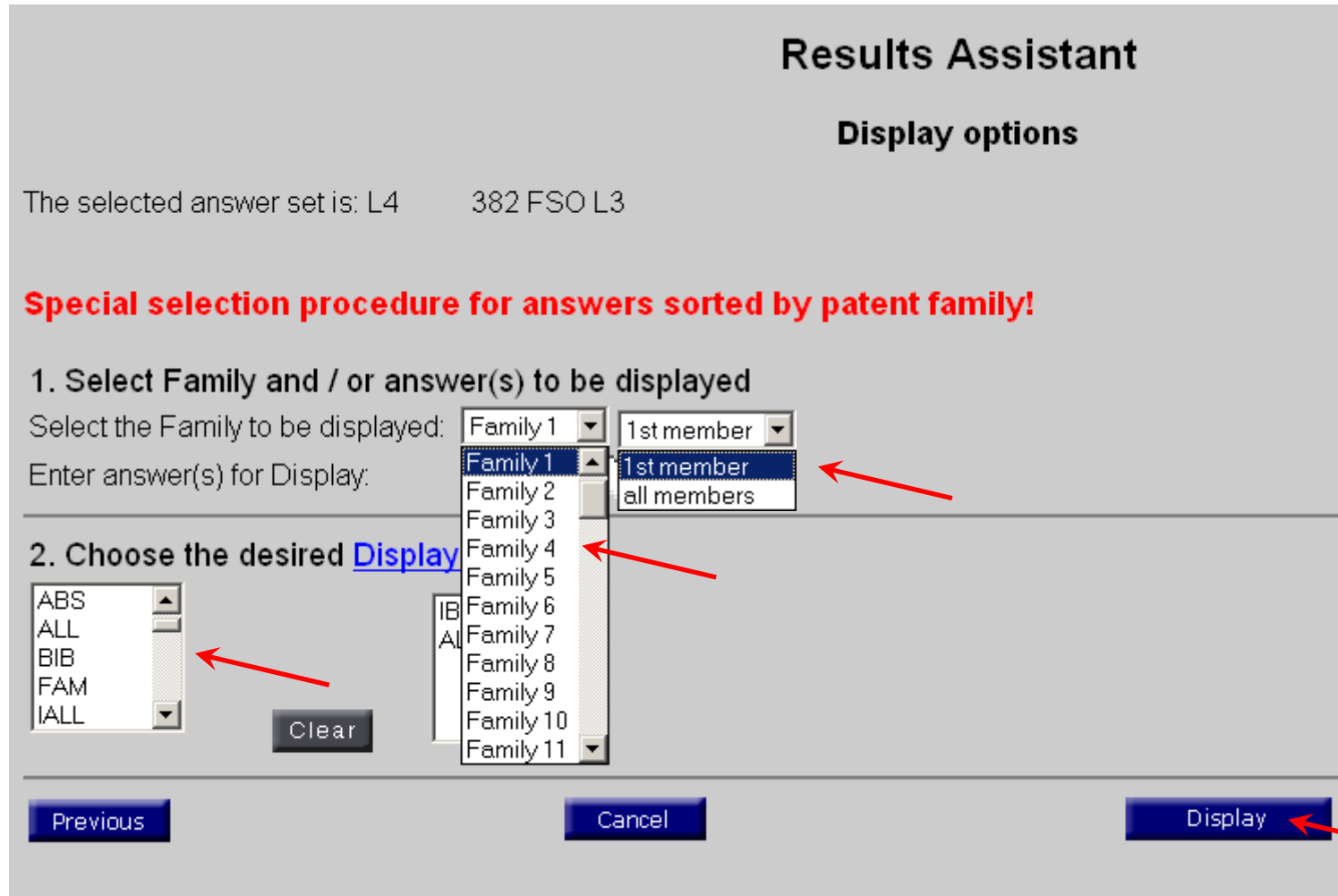
IB
Al

Family 1
Family 2
Family 3
Family 4
Family 5
Family 6
Family 7
Family 8
Family 9
Family 10
Family 11

1st member
all members

Clear

Previous Cancel Display



Displaying answers

```
=> dis L4 pfam=1 1 IBIB ALIGN;
L4 ANSWER 1 OF 382 DGENE COPYRIGHT 2007 THE THOMSON CORP on STN
Full
Text
FAMILY 1
ACCESSION NUMBER: AEH10480 cDNA DGENE
TITLE: Analyzing differential gene expression associated with
histopathologic features of colorectal disease, involves
detecting overexpression or underexpression of pool of
polynucleotide sequences in colon tissues.
INVENTOR: Bertucci F; Houlgatte R; Birnbaum D; Debono S
PATENT ASSIGNEE: (IPSO-N) IPSOGEN.
(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
(PAOL-N) INST PAOLI CALMETTES IPC.
PATENT INFO: WO 2005054508 A2 20050616 154
APPLICATION INFO: WO 2004-IB4323 20041201
PRIORITY INFO: US 2003-525987P 20031201
US 2004-688 20041201
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2005-435408 [44]
DESCRIPTION: Aquaporin 1 (AQP1) cDNA SEQ ID 866.
BLASTALIGN
Query = 1662 letters
Length = 1662
Score = 3295 bits (1662), Expect = 0.0
Identities = 1662/1662 (100%)
Strand = Plus / Plus

Query: 1 gcaccggcagcgggtctcaggccaagccccctgccagcatggccagcgagttcaagaaga
|||||
Sbjct: 1 gcaccggcagcgggtctcaggccaagccccctgccagcatggccagcgagttcaagaaga

Query: 61 agctcttctggagggcagtggtggccgagttcctggccacgaccctctttgtcttcatca
|||||
Sbjct: 61 agctcttctggagggcagtggtggccgagttcctggccacgaccctctttgtcttcatca

Query: 121 gcatcggttctgcacctgggcttcaaataccgggtggggaacaaccagacggcggtccagg
|||||
Sbjct: 121 gcatcggttctgcacctgggcttcaaataccgggtggggaacaaccagacggcggtccagg

Query: 181 acaacgtgaaggtgtcgtggccttgggctgagcatcgccacgctggcgagagtggtg
|||||
```

The first (=best) hit from the first multi-record patent family (answers 1-5).

Displaying answers

```
=> dis L4 pfam=15 1 IBIB ALIGN;
L4 ANSWER 50 OF 382 DGENE COPYRIGHT 2007 THE THOMSON CORP on STN
Full Text
FAMILY 15
ACCESSION NUMBER: AÄC00029 cDNA DGENE
TITLE: New nucleic acid that is a 5' expressed sequence tag (5' EST)
for obtaining cDNAs and genomic DNAs that correspond to
5'ESTs and for diagnostic, forensic, gene therapy and
chromosome mapping procedures -
INVENTOR: Dumas Milne Edwards J; Duclert A; Giordano J
PATENT ASSIGNEE: (GEST)GENSET.
PATENT INFO: EP 1033401 A2 20000906
APPLICATION INFO: EP 2000-200610 20000221
PRIORITY INFO: US 1999-122487 19990226
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2000-500381 [45]
CROSS REFERENCES: P-PSDB: AÄG00023
DESCRIPTION: Human secreted protein 5' EST, SEQ ID NO: 27.
BLASTALIGN
Query = 1662 letters
Length = 382
Score = 673 bits (339), Expect = 0.0
Identities = 357/364 (98%), Gaps = 1/364 (0%)
Strand = Plus / Plus

Query: 1 gcacccggcagcgggtctcaggccaagccccctgccagcatggccagcgagttcaagaaga
|||||
Sbjct: 20 gcacccggcagcgggtctcaggccaagccccctgccagcatggccagcgagttcaagaaga

Query: 61 agctcttctggagggcagtggtggccgagttcctggccacgacctctttgtcttcatca
|||||
Sbjct: 80 agctcttctggagggcagtggtggccgagttcctggccacgacctctttgtcttcatca

Query: 121 gcatcggttctgccctgggcttcaaatacccggtggggaacaaccagacggcggtccagg
|||
Sbjct: 140 gcancggttctgccctgggcttcaaatacccggtgggmacaaccagacggcggtccagg

Query: 181 acaacgtgaaggtgtcgctggccttcgggctgagcatcgccacgctggcgcagagtgtgg
|||||
Sbjct: 200 acaacgtgaaggtgtcgctggccttcgggctgagcatcgccacgctggcgcagagtgtgg
```

The first (=best) hit from mutli-record patent family 15.

```
=>     Hide session output  
Show session output
```

Batch and Alert results

STN on the web

- Help
- News
- Search Assistants
 - Search Assistant
 - Patent Search
 - Alert (SDI) Asst.
 - Structure Query
 - Sequence Asst.
 - Upload Sequence
 - Upload Cmd. File
- Results Assistant
- Transcript Assistant
- => Command Line
- Logoff Hold
- Logoff
- Feedback
- Send Break

Sequence Search Assistant

CAS Registry BLAST®: [Cost Information](#) [\(Plug-in Required\)](#)

[New version of BLAST plug-in available August 12, 2007](#)

Launch CAS Registry BLAST Launch sequence searching and review BLAST Reports ([Help](#)) ([Security](#))

Select STN session status following Launch of CAS Registry BLAST:

- Session will time out in 20 minutes
- LOG-OFF HOLD (session may be resumed within 120 minutes)
- LOG-OFF

Retrieve RNs from BLAST Retrieve a set of RNs previously transferred from a BLAST Report ([Help](#))

DGENE/PCTGEN/USGENE Sequence Searches: Cost Information: [DGENE](#) [PCTGEN](#) [USGENE](#) [\(No Plug-in Required\)](#)

Launch

in DGENE PCTGEN USGENE

Select the Type of Search:

Continue Conduct a menu-driven sequence search using the selected search algorithm in the selected database ([Help](#))

Show Batch Status Show Status Information about Offline Batch Searches ([Help](#))

Show Alert Status Show Status Information about Alert(SDI) Searches ([Help](#))

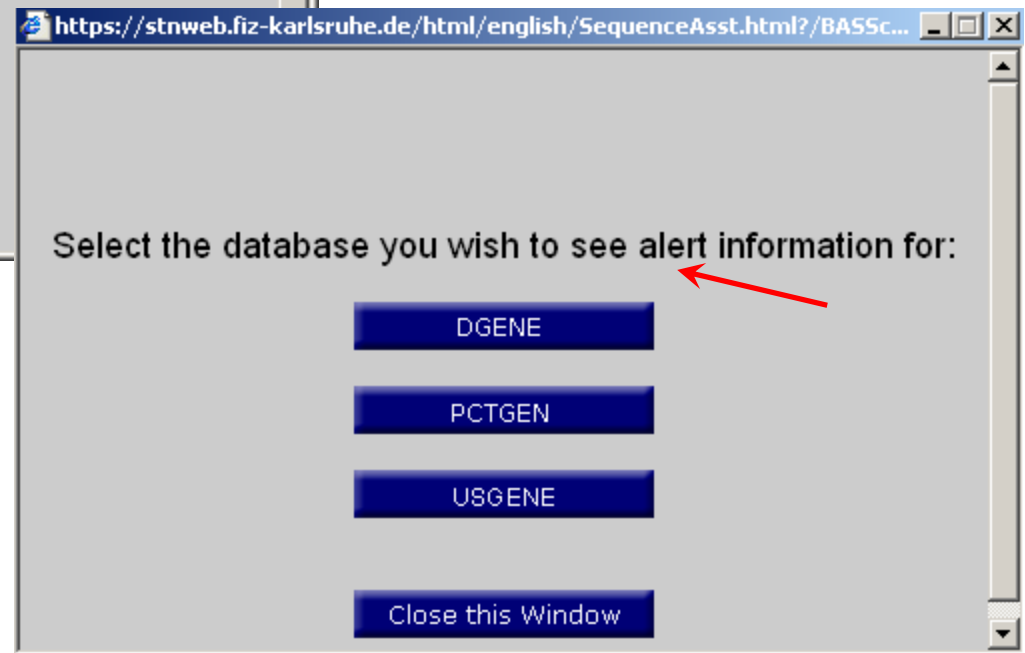
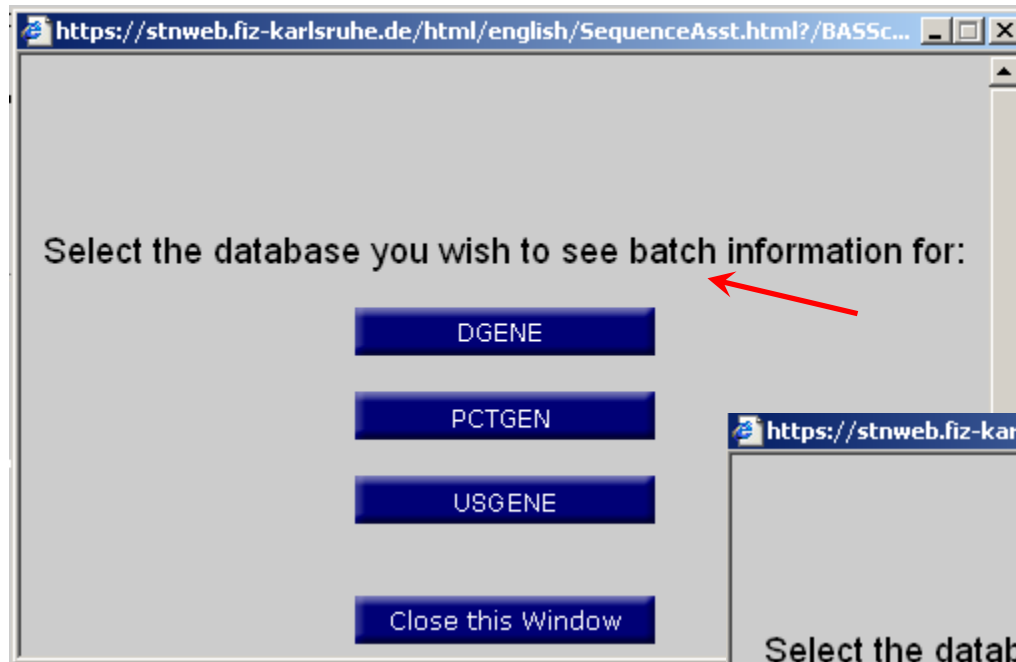
BLAST is a registered trademark of the National Library of Medicine **Cancel**

Transcript:

[STN Command List](#)
[File-Specific Help List](#)

Check your BATCHes and ALERTs via the Sequence Search Assistant.


Retrieving Batch and Alert results








Select the database from which to check the BATCH or ALERT results.

Batch result status



Sequence Search Assistant

Batch results files remaining in USGENE: 

Name	Status	Type	Results	Delete ?
1. EBOLA1	Completed	(blast)	 	<input type="checkbox"/>
2. EBOLA2	Completed	(getsim)		<input type="checkbox"/>
3. SGLT2	Retrieved	(blast)		<input type="checkbox"/>



BLAST is a registered trademark of the National Library of Medicine

'Show Results' will show the search results diagram.

Alert results

Sequence Search Assistant

Alert (SDI) Status:

Current results available in DGENE:

Name	Run Date	Type	Results	Delete ?
1. AQUAP2	2007 11 19	(blast)	Show Results	<input type="checkbox"/>
2. AQUAP2	2007 12 01	(blast)	Show Results	<input type="checkbox"/>
3. TAENIA1	2007 12 01	(getsim)	Show Results	<input type="checkbox"/>
4. TEST1	2007 11 19	(blast)	Show Results	<input type="checkbox"/>
5. TEST1	2007 12 01	(blast)	Show Results	<input type="checkbox"/>
6. TEST2	2007 11 19	(blast)	Show Results	<input type="checkbox"/>
7. TEST2	2007 12 01	(blast)	Show Results	<input type="checkbox"/>
8. TEST3	2007 11 19	(blast)	Show Results	<input type="checkbox"/>
9. TEST3	2007 12 01	(blast)	Show Results	<input type="checkbox"/>
10. TEST5	2007 11 19	(getsim)	Show Results	<input type="checkbox"/>
11. TEST5	2007 12 01	(getsim)	Show Results	<input type="checkbox"/>

[Delete Results](#)

BLAST is a registered trademark of the National Library of Medicine

[Previous](#)

[Cancel](#)

[Queries](#)

Save the session transcript

Transcript Assistant

[Back to Session](#)

- To save transcripts to your desktop, you may download them from the server using this Transcript Assistant. See the [Transcript Capture Notes](#) for more information about the transcript capture process.
- Transcripts may be downloaded in your choice of these formats
 - HTML** -- self-extracting compressed EXE format or standard compressed ZIP format, containing session text in an HTML file and session graphics in JPEG or PNG files
 - RTF (Rich Text Format)** -- self-extracting compressed EXE format or standard compressed ZIP format, containing session text in an RTF file and session graphics in JPEG or PNG files; or a single RTF file containing session text and embedded graphics
 - Adobe Acrobat Format (PDF)** -- session text and graphics in a single file
- Download a transcript by clicking on the appropriate hyperlink in the table below. For additional information about these formats, see [How to get HTML transcripts](#), or [How to get RTF transcripts](#), or [How to get PDF transcripts](#).
- Point your mouse to the "CLOSED" Capture Status for a list of the first few commands from that session.

Transcripts are available for 4 days		Capture Status	Format (all formats include images)		
STNweb20071204X171701 2007 12/04 17:17	Rename	ON	HTML EXE or ZIP	RTF EXE or ZIP or RTF	Adobe PDF
STNweb20071204X114823 2007 12/04 11:48	Rename	CLOSED	HTML EXE or ZIP	RTF EXE or ZIP or RTF	Adobe PDF
aquaporin 2007 12/03 17:40	Rename	CLOSED	HTML EXE or ZIP	RTF EXE or ZIP or RTF	Adobe PDF
STNweb20071203X172904 2007 12/03 17:29	Rename	CLOSED	HTML EXE or ZIP	RTF EXE or ZIP or RTF	Adobe PDF
STNweb20071203X113911 2007 12/03 11:39	Rename	CLOSED	HTML EXE or ZIP	RTF EXE or ZIP or RTF	Adobe PDF

Transcript Capture Notes

- At login time, Transcript capture may be set **OFF** or **ON (default)**, and a **Transcript Name** may optionally be specified, using the fields provided on the Login Page.
- During a session, Transcript capture may be easily toggled **ON** and **OFF** using the convenient Transcript status selection box located in the left-side menu frame.
- During a session, a new Transcript may be started by selecting the **New...** option from the Transcript status selection box. Once a new Transcript has been started, the previous Transcript may not be re-opened.

STN on the web

- Help
- News
- Search Assistants
 - Results Assistant
 - Transcript Assistant
 - => Command Line
- Logoff Hold
- Logoff
- Feedback
- Send Break

Transcript:

[STN Command List](#)
[File-Specific Help List](#)

For price information use HELP COST in each file

=> help cost

STN International Fees and Prices, Effective Jan 1, 2007

DGENE File	Euro
-----	-----
Connect Hour Fee (per hour) .	121,00
SDI Search Fee (every 2 weeks)	29,80
SDI PACKAGE Component Fee 1)	29,80
SDI PACKAGE Component Frequency: every 2 weeks	
Display Fee (per answer) . . .	
- BIB, IBIB	3,03
- ABS	3,03
- SQIDE, SQ3IDE	6,16
- ALL, IALL	12,22
- FAM	9,20
- LS, LS2	0,87
(from the INPADOCDB file)	
- TRIAL (TRI, SAM), SCAN . .	FREE
Print Fee (per answer)	
- BIB, IBIB	3,18
- ABS	3,18
- SQIDE, SQ3IDE	7,52
.....	

- ALL, IALL	13,88
- FAM	11,20
- LS, LS2	0,87
(from the INPADOCDB file)	
- TRIAL (TRI, SAM)	0,81
Offline Print Postage Fee	
(additional per answer) . .	0,16
Sequence Search	
- Sequence Search per RUN GETSEQ	16,48
- Homology Search per RUN GETSIM	21,73
GETSIM Batch Initiation Fee	8,14
GETSIM Batch Collection Fee	26,78
GETSIM Alert Collection Fee	8,14
- Homology search per RUN BLAST	21,73
BLAST Batch Initiation Fee	8,14
BLAST Batch Collection Fee	26,78
BLAST Alert Collection Fee	8,14
.....	

For price information use HELP COST in each file

=> help cost

STN International Fees and Prices, Effective Jul 29, 2007

USGENE File	Euro
-----	-----
Connect Hour Fee (per hour) .	83,00
SDI Search Fee (weekly)	11,00
SDI PACKAGE Component Fee 1)	11,00
SDI PACKAGE Component Frequency: weekly	

Display Fee (per answer)	
- AB	0,83
- BIB, IBIB	1,30
- SQIDE, SQ3IDE	3,00
- ECLM	0,46
- CLM	0,92
- ALL, IALL	6,05
- BRIEF, IBRIEF	5,59
- TRIAL (TRI, SAM), FREE, SCAN	FREE

Select Fees (per record)	
- AN, AP, PN, RLN	0,17
Sequence Search	
- Sequence Search per RUN GETSEQ	15,84
- Homology Search per RUN GETSIM	15,84
GETSIM Alert Collection Fee	5,42
GETSIM Batch Initiation Fee	5,42
GETSIM Batch Collection Fee	18,33
- Homology search per RUN BLAST	15,84
BLAST Alert Collection Fee	5,42
BLAST Batch Initiation Fee	5,42
BLAST Batch Collection Fee	18,33

More information is available

Helpdesk Contact

<http://www.stn-international.com/service/custserv.html>

<http://www.cas.org/supp.html>

Training Material

http://www.stn-international.de/training_center/mat_sea_stn.html#Bioseq

International Workshops

http://www.stn-international.de/training_center/workshops/workshop.html