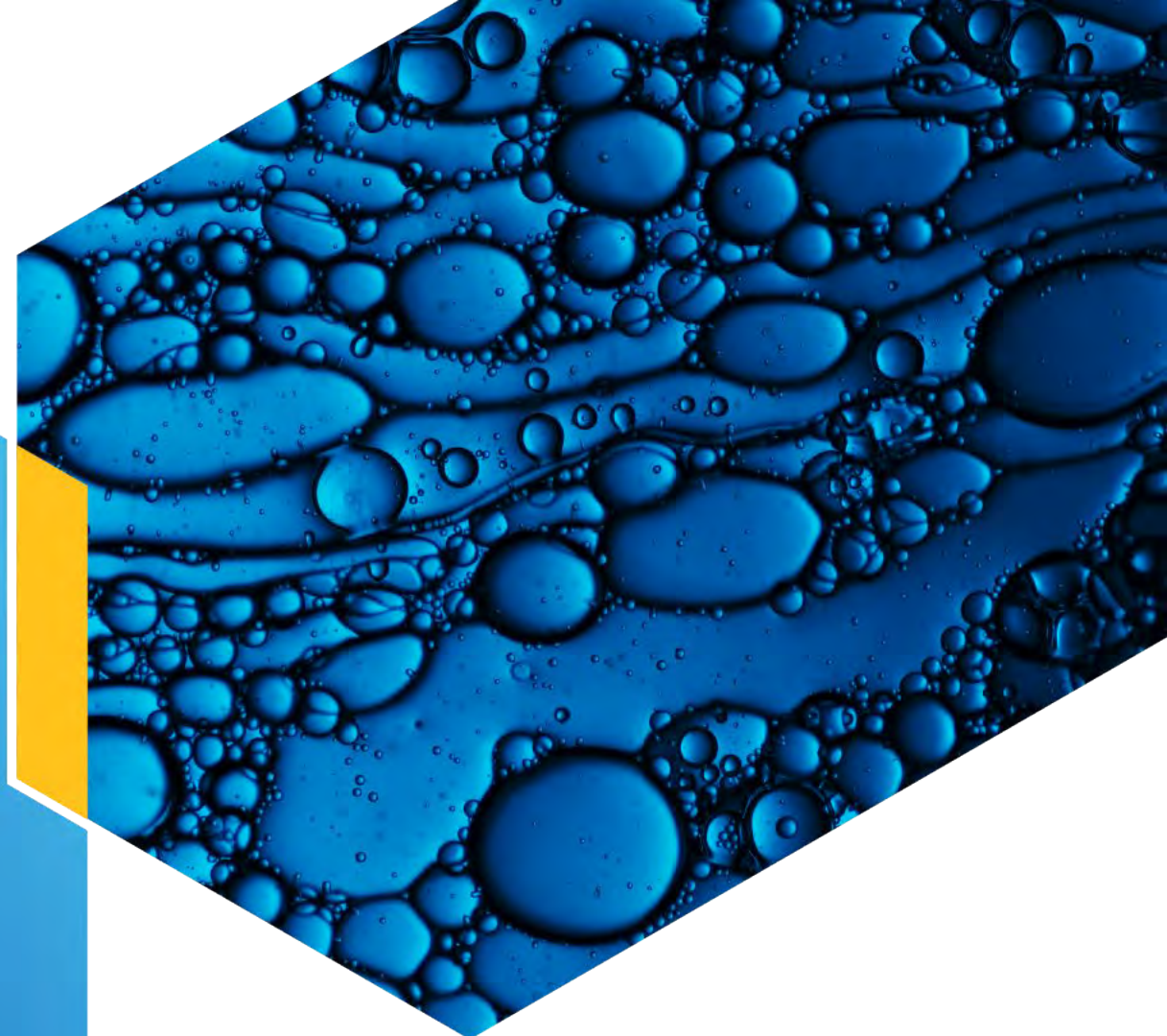


**STN User Meeting - France
October 2021**

Biosequences in STN[®]Next



CAS Biosequences™

Comprehensive collection

- CAS-curated REGISTRY® sequence data and extracted patent sequences from major patent offices
- Over 550M patent-sequence relationships from more than 1.1 M patents and 60+ patent authorities
- 23M sequences with linked non-patent literature references
- Manually curated sequences not found in electronic sequence listings and other databases

New search interface

- Easy access to search modes and parameters
- Sequence search results including alignments and related patents
- Dynamic sorting and filtering
- Crossover results to bibliographic STN databases, e.g. CAplus
- Export to Excel

- <
- Coming soon
 - Inclusion of public sequence databases

CAS Biosequences (Patent Sequence data)

Jurisdiction	Patent Publications	Sequence-patent pairs
Total	1,130,651	627,969,136
US	331,582	260,974,365
WO	225,852	251,937,600
JP	119,735	37,034,978
EP	88,354	36,304,604
CA	62,242	31,470,642
CN	207,767	4,029,203
IN	1,264	12,692
Others	93,855	6,205,052

The data corpus includes biosequence information from the following **62 patent authorities**:

AP, AR, AT, AU, BE, BG, BR, CA, CH, CN, CO, CR, CS, CZ, DD, DE, DK, EA, EE, EG, EP, ES, FI, FR, GB, GR, HK, HR, HU, IE, IL, IN, IP, IT, JP, KR, LT, LU, LV, MC, MD, MX, MY, NL, NO, NZ, PH, 'T, RD, RO, RU, SE, SG, SI, SK, 'R, TW, US, WO, ZA

CAS Biosequences (Non-Patent Literature)

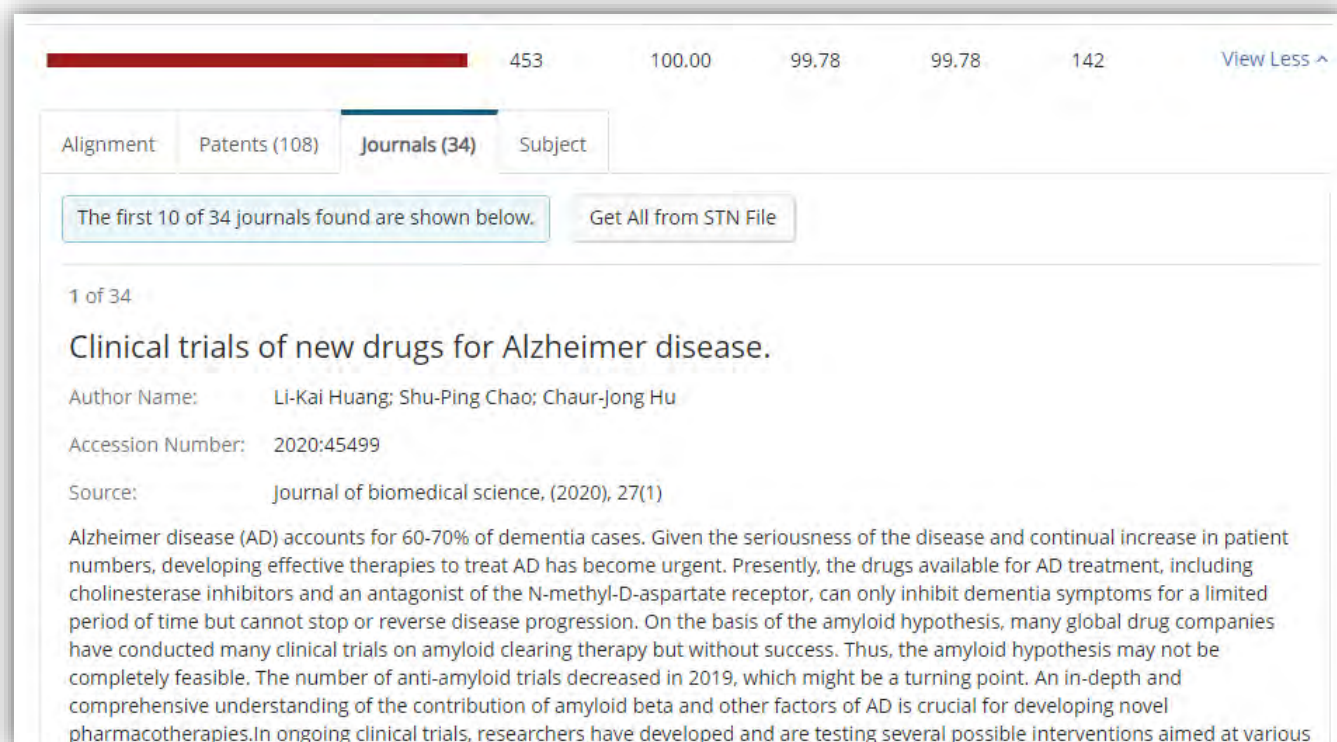
Users can now view the bibliographic information for non-patent sources and cross reference those sources in other STNext files for additional research and investigation.

Literature Types:

- Journals
- Preprint
- Reports

Bibliographic Information:

- CAS Accession Numbers
- Author Names
- Publication Title
- Publication Year, Volume, Issue, Pagination



The screenshot displays a search results page for 'Clinical trials of new drugs for Alzheimer disease'. At the top, there is a progress bar and a 'View Less' link. Below the progress bar, there are tabs for 'Alignment', 'Patents (108)', 'Journals (34)', and 'Subject'. A message indicates that the first 10 of 34 journals are shown, with a 'Get All from STN File' button. The first result is 'Clinical trials of new drugs for Alzheimer disease' by Li-Kai Huang, Shu-Ping Chao, and Chaur-Jong Hu, with an accession number of 2020:45499. The source is 'Journal of biomedical science, (2020), 27(1)'. The abstract text is visible below the source information.

453 100.00 99.78 99.78 142 View Less ^

Alignment Patents (108) Journals (34) Subject

The first 10 of 34 journals found are shown below. Get All from STN File

1 of 34

Clinical trials of new drugs for Alzheimer disease.

Author Name: Li-Kai Huang; Shu-Ping Chao; Chaur-Jong Hu

Accession Number: 2020:45499

Source: Journal of biomedical science, (2020), 27(1)

Alzheimer disease (AD) accounts for 60-70% of dementia cases. Given the seriousness of the disease and continual increase in patient numbers, developing effective therapies to treat AD has become urgent. Presently, the drugs available for AD treatment, including cholinesterase inhibitors and an antagonist of the N-methyl-D-aspartate receptor, can only inhibit dementia symptoms for a limited period of time but cannot stop or reverse disease progression. On the basis of the amyloid hypothesis, many global drug companies have conducted many clinical trials on amyloid clearing therapy but without success. Thus, the amyloid hypothesis may not be completely feasible. The number of anti-amyloid trials decreased in 2019, which might be a turning point. An in-depth and comprehensive understanding of the contribution of amyloid beta and other factors of AD is crucial for developing novel pharmacotherapies. In ongoing clinical trials, researchers have developed and are testing several possible interventions aimed at various

CAS Biosequences

To begin a Sequence search

My Files

Alerts

Transcripts

Structures

Scripts

Biosequences



[Return to Session](#)

Biosequences (1) Sort By: Date Modified: Newest ▾

[Move to Folder](#)

<input type="checkbox"/> 2021_0001_Sequence	DLSKQEEEEAVRLFIEWLKNAGPSSGAPPPS	<input type="button" value="View"/>	<input type="button" value="View Results"/>
2021 May 20 12:41 PM			1000 results

Show per page

CAS Biosequences

Blast: Start a search

Search Biosequences Close

BLAST CDR Motif Tabs for each Search Type

Name

1 DLSKQXEEEAVRLFIEWLK NAGPSSGAPPPS

Query entry

Sequence type
 Nucleotide Protein

Search within
 Nucleotides Proteins

Limit Total Sequence Results to

Up to 20,000 results may be returned

[Advanced Biosequence Search](#)

Advanced search options

CAS Biosequences

BLAST: Output

- Filters adjust parameters
- Get all Patents
- Bioscape analysis
 - See a landscape of similar sequences

Filter By

Query Identity %
No Min to No Max

Query Coverage
No Min to No Max

Subject Coverage
No Min to No Max

Subject Identity %
No Min to No Max

Apply Reset

Biosequence Search Results (1000)

Sort By: Alignment Identity %: Descending



Create Bioscape Analysis

Get All Patents

Show Search Details

Sequence Alignment	Sequence Length	Alignment Identity %	Query Identity %	Subject Identity %	Number of Patents	
	115	100.00	100.00	33.91	14	View More
	964	100.00	100.00	4.05	24	View More
	40	100.00	100.00	97.50	1	View More
	52	100.00	100.00	75.00	1	View More
	205	100.00	100.00	19.02	4	View More
	291	100.00	100.00	13.40	1	View More

CAS Biosequences

CDR: Start a search

Search Biosequences Close

BLAST **CDR** Motif

Name

> CDR 1

> CDR 2

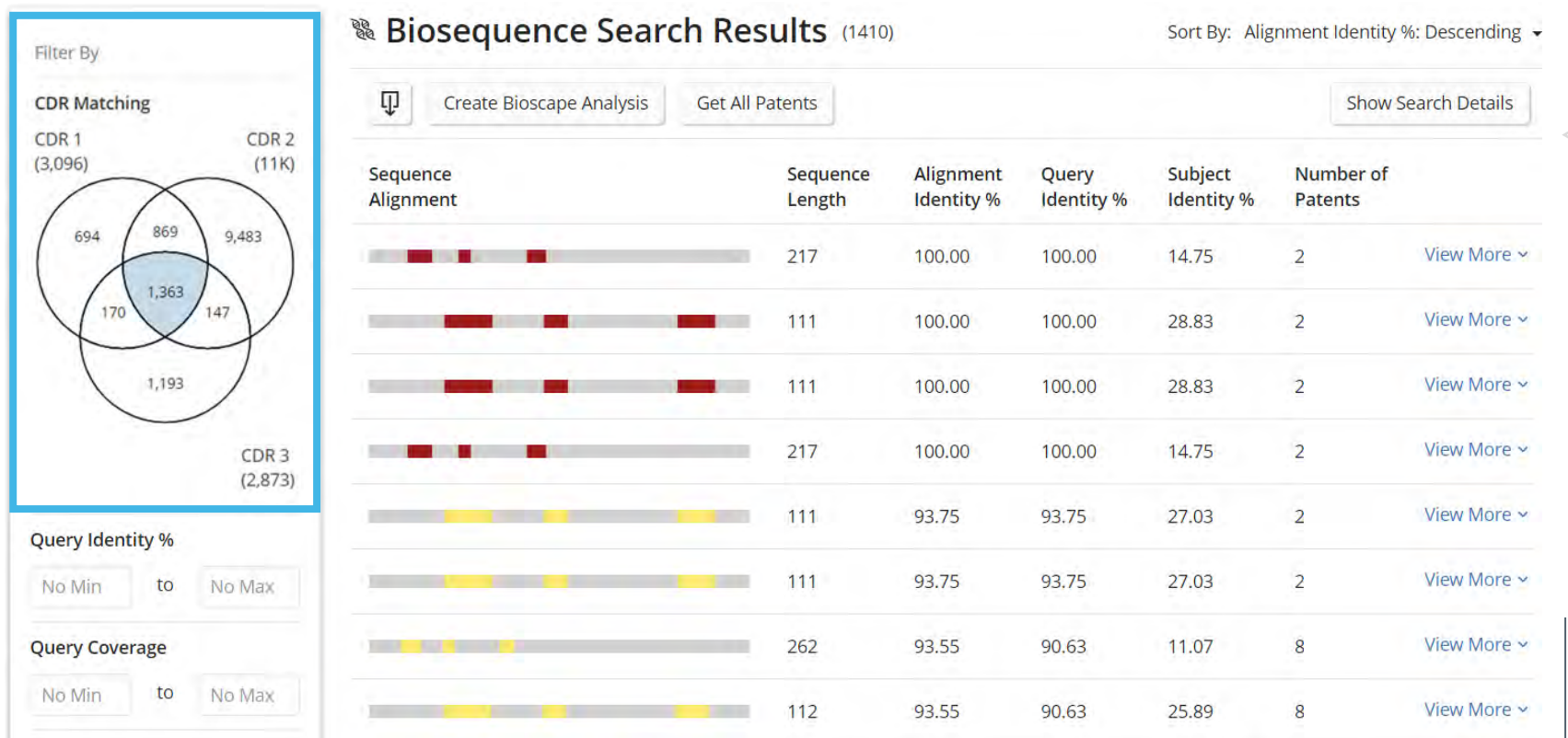
> CDR 3

Limit Total Sequence Results to

CAS Biosequences

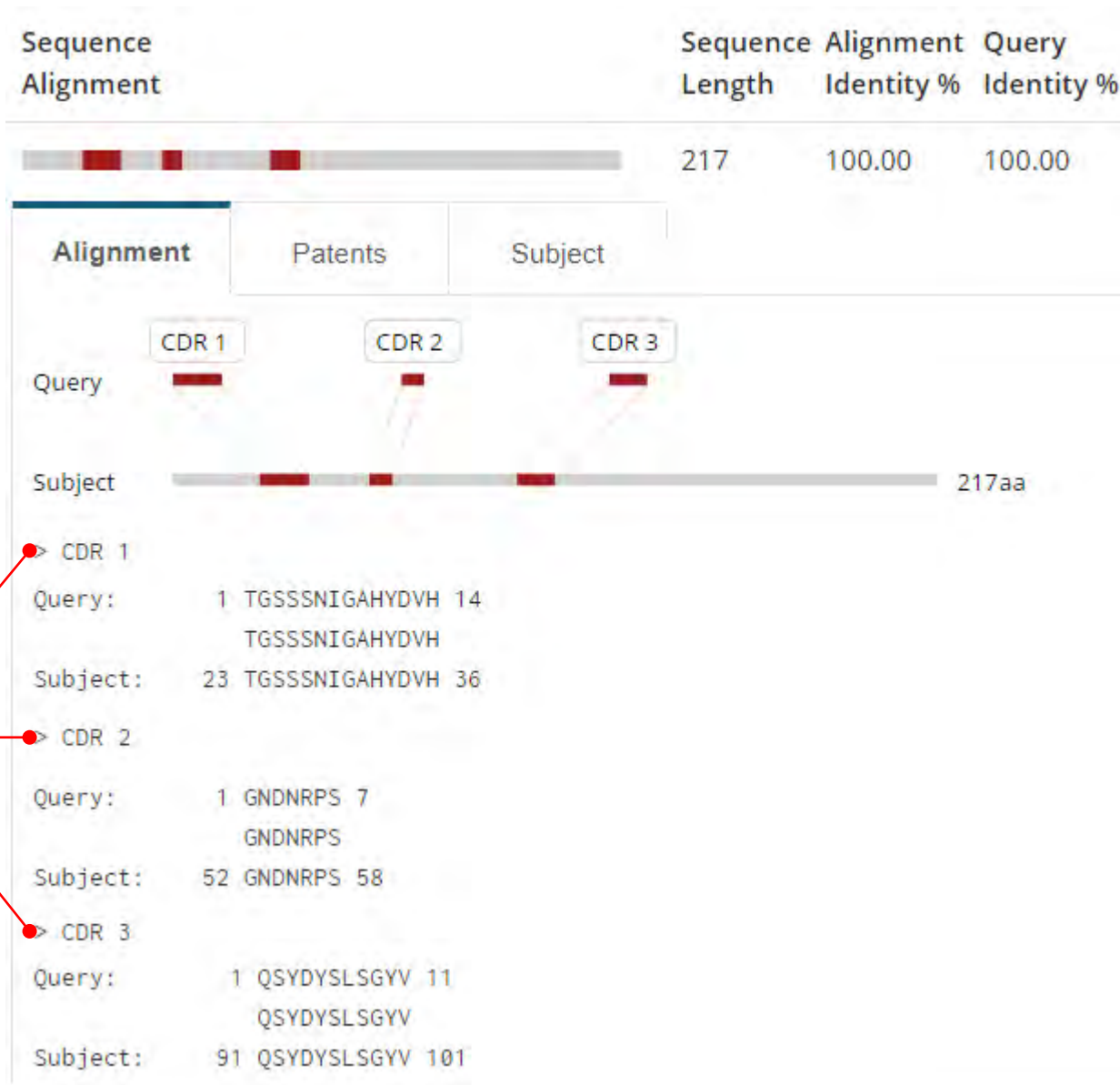
CDR: Results

- Interactive Venn diagram shows distribution of CDRs among alignment results



CAS Biosequences

CDR: Results



CAS Biosequences

Motif: Start a search

Search Biosequences Close

BLAST CDR **Motif** Name 2021_0002_Sequence

Import Sequence Download Clear

1 [LI][ALKV]G[FL][VI][D][AG]DG

Sequence type
 Nucleotide Protein

Limit Total Sequence Results to
1000

Learn more about Motif

Advanced Biosequence Search Reset Parameters

Query Coverage % E-Value
0 to 100 10

Run Search Save Query Cancel

Motif searching allows variable positions

CAS Biosequences

Motif: Results

Query: VGIGGGGGGGGGXGGQGGXGX[SN][NG]XGGNGXGXGSHI

Biosequence Search Results (2119)

Sort By: Alignment Identity %: Descending

View Results

4 of 4 searches complete

8416 results

Query Sequence

Sequence 1: VGIGGGGGGGGGXGGQGGXGXSNXGGNGXGXGSHI

Sequence 1: VGIGGGGGGGGGXGGQGGXGXSNXGGNGXGXGSHI

(2119)

Sequence 2: VGIGGGGGGGGGXGGQGGXGXSGXGGNGXGXGSHI

(2122)

Sequence 3: VGIGGGGGGGGGXGGQGGXGXNNXGGNGXGXGSHI

(2055)

Se

Al

Sequence 4: VGIGGGGGGGGGXGGQGGXGXNGXGGNGXGXGSHI

(2120)

	35	100.00	100.00	100.00
	61	100.00	100.00	57.14
			60.00	95.24
			48.57	100.00

[SN] in the above query finds sequences which contain either Serine (S) or Asparagine (N) at amino acid position 22

The motif search is based on BLAST®, thus also retrieving similar subject sequences

View More

CAS Biosequences

Locate CAS Registry Number(s)

- For BLAST, CDR, and Motif results, users can get registry numbers (RNs) for a particular biosequence in CAS Registry:
 - Click the **Get from STN File** button to **get the single RN**.
 - Click the **Get All from STN File** button to **get RNs from either REGISTRY or ZRegistry (up to 5000)**.

Biosequence Search Results (2119) Sort By: Alignment Identity %: Descending ▾

Query Sequence
Sequence 1: VGIGGGGGGGGGXGGQGGGXSNXGGNGXGXGSHI

Sequence Alignment	Sequence Length	Alignment Identity %	Query Identity %	Subject Identity %	Number of Patents	
	61	100.00	100.00	57.38	0	View Less ▾

Alignment Patents **Subject**

Sequence Length: 61 aa

CAS Registry Number®: 151820-96-1

Sequence:
1 MELKASEFGV VLSVDALKLS RQSP LGVIG GGGGGGGXG GGGGXSNX
51 GGNGXGXGSH I

Alignment Patents **Subject**

Sequence Length: 35 aa

CAS Registry Number®: 1562410-96-1, 183270-54-4, 84286-90-8, 234125-11-2, 234125-12-3, 234125-14-5, 234125-08-7, 234125-09-8

CAS Biosequences

Visualizations

- Adjust parameters
- Get all Patents
- Bioscape analysis

Biosequence Search Results



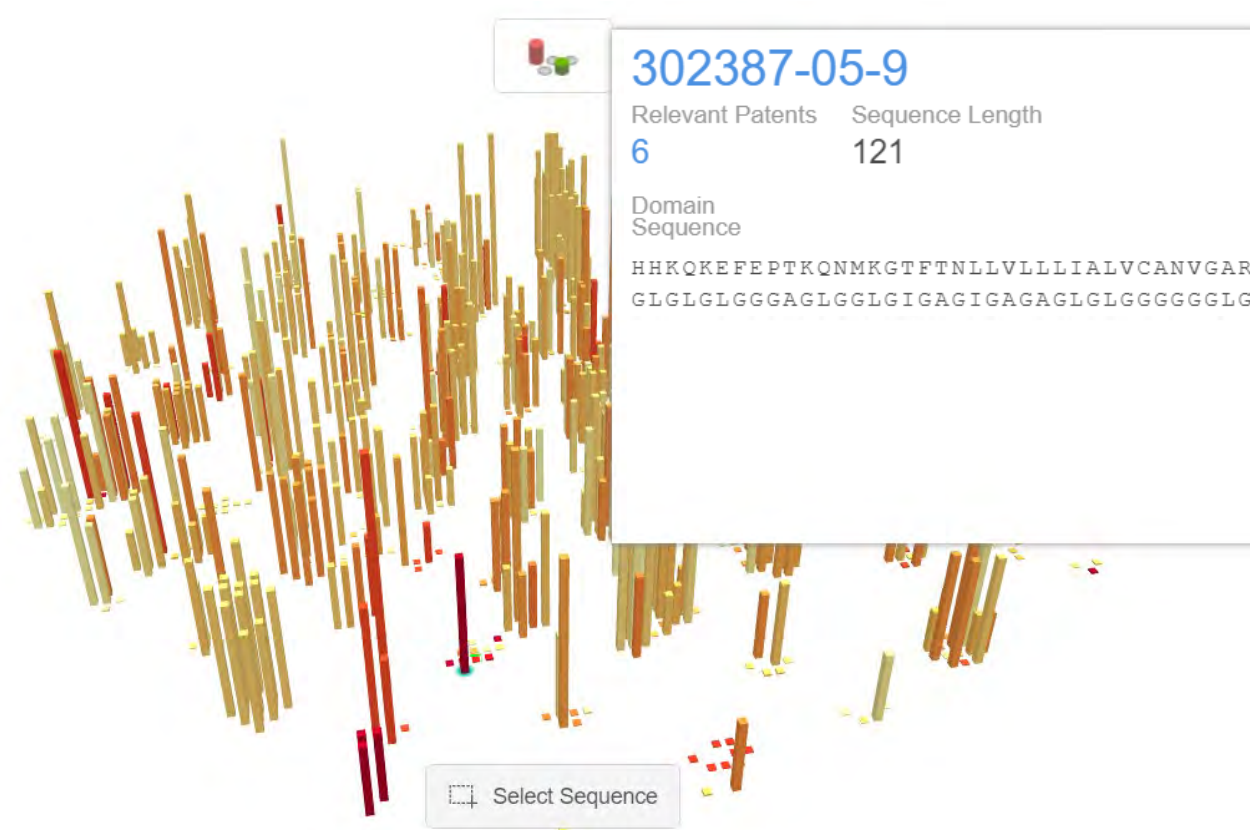
Create Bioscape Analysis

Get All Patents

Query position
Indicated by blue circle

Position
Similarity to Query
(multiple dimensions)

Color
Similarity measured
by Alignment Identity



CONTACT

CAS

help@cas.org
cas.org

FIZ Karlsruhe

helpdesk@fiz-karlsruhe.de
stn-international.de