CAS
A division of the
American Chemical Society

知財戦略＆特許調査に差がつく
STN IP Protection Suite セミナー
&
CAS STNext ユーザーミーティング資料

2022 年 9 月

JAICI
化学情報協会

＊ 目次 ＊

◆ 知財戦略＆特許調査に差がつく STN IP Protection Suite セミナー

1. CAS Scientific Patent Explorer™
   ～ 特許検索・解析をより簡単に ～ ……1

2. Biosequences Search
   ～ CAS STNext® の新しい配列検索機能 ～ ……25

3. CAS FORMULATIONS™ ファイル
   ～ 製剤・配合の調査に最適なデータベース ～ ……41

◆ CAS STNext ユーザーミーティング

1. CAS FILES の強化 ……63

2. WPI ファイルのリロード ……79

3. 配列ファイルのリロード ……97

4. CAS STNext の強化 ……111

知財戦略＆特許調査に差がつく

STN IP Protection Suite セミナー

CAS Scientific Patent Explorer™

～ 特許検索・解析をより簡単に ～

CAS Scientific Patent Explorer™
～特許検索・解析をより簡単に～

化学情報協会　情報事業部

目次

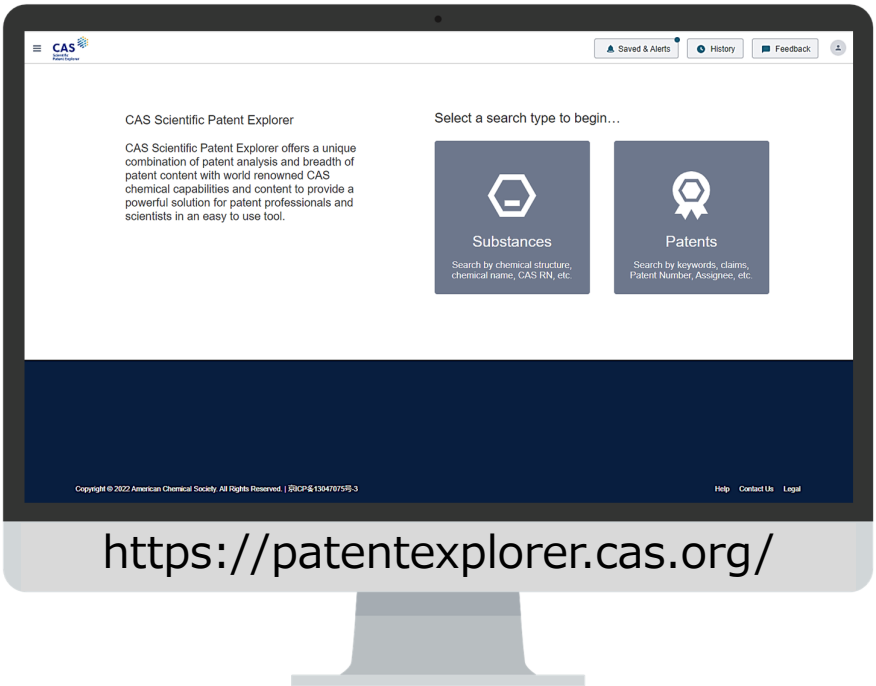
1. CAS Scientific Patent Explorer とは
2. 活用事例 1 - 特許の動向を調査する
3. 活用事例 2 – 物質関連特許の動向を調査する
4. 強化一覧

© 2022 化学情報協会

CAS Scientific Patent Explorer とは

CAS Scientific Patent Explorer とは 3

CAS 独自の化学物質情報と特許情報を組み合わせた
特許調査のための検索ツール



https://patentexplorer.cas.org/

© 2022 化学情報協会

CAS Scientific Patent Explorer の特長 4

簡単な操作で世界中の特許を調査できる

- 検索：広範な特許情報を対象とした検索
  化学物質とその関連特許の検索
- 表示：特許公報・特許ファミリー、法的状況、
  引用情報、CAS 由来の情報をまとめて確認
- 解析：特許・化学物質の検索結果をビジュアル化
  様々な観点で俯瞰できる

© 2022 化学情報協会

広範な特許情報を収録 5

(2022 年 8 月)

	CAS Scientific Patent Explorer	CAS STNext*3		
		CAplus	WPI	INPADOCDB/ INPAFAMDB
特許発行国・機関	115	63	61	約 100
収録分野	全産業	化学および周辺分野	全産業	全産業
特許全文	約 60*1	13	24*2	×
法的状況	○	○	×	○
特許分類	IPC, CPC, F-TERM, FI	IPC, CPC, F-TERM	IPC, CPC, F-TERM, FI	IPC, CPC, F-TERM, FI
引用，被引用情報	○	○	○	×
更新頻度	毎週	毎日	週 2 回	毎週

*1 特許全文の収録は特許発行国の特許種別や年代により異なる
*2 全クレームの収録は 21 特許発行国・機関、メインクレームの収録は 3 特許発行国
*3 CAS STNext では 13 特許発行国・機関の全文データベースを利用できる

© 2022 化学情報協会

CAS 由来の情報を収録 6

(2022 年 8 月)

	CAS REGISTRY (REGISTRY)	CAS Markush (MARPAT)	CAS References (CAplus)	CAS Formulations (CASFORM)	CAS Reactions (CASREACT)
内容	化学物質	特許中の マルクーシュ構造	化学および周辺 分野の文献情報	製剤・配合情報	重要な有機化学 反応
検索	○ ・CAS RN® ・SMILES ・InChI ・Molfile ・化学物質名称 ・構造質問式	○ ・構造質問式	×	×	×
表示	○ ・CAS RN® ・化学物質名称 ・分子式 ・構造図 ・基本的な物性情報	○ ・マルクーシュ構造 ・特許番号	○ ・重要な化学物質 ・主題に関わる 概念の索引	○ ・配合情報 ・CAS Formulus リンク	○ ・反応スキーム

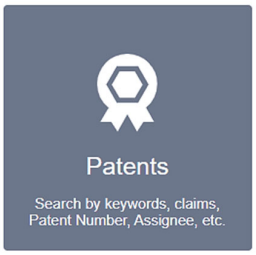
（　）内は CAS STNext のファイル名

© 2022 化学情報協会

活用事例 1
特許の動向を調査する

リチウムイオン電池のリサイクルについて、特許の動向を調査する

- 広範な特許情報を収録
- 検索・絞り込み機能
- 特許検索結果のビジュアル化



© 2022 化学情報協会

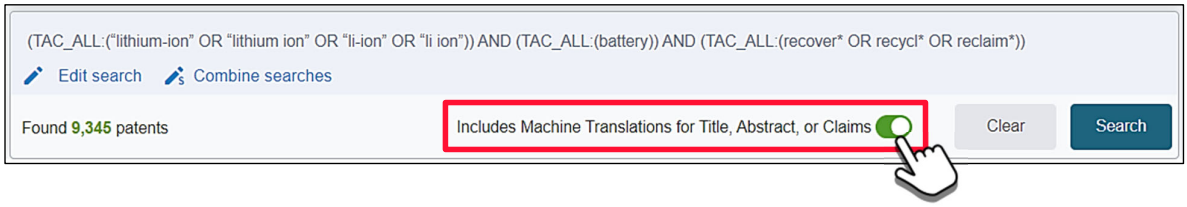
Advanced Patent Search 9

メニュー形式で簡単に複数の語や項目を組み合わせた検索が可能

© 2022 化学情報協会

機械翻訳

Includes Machine Translations for Title, Abstract, or Claims を指定すると標題、抄録、クレームについて翻訳語も含めた検索が実行される



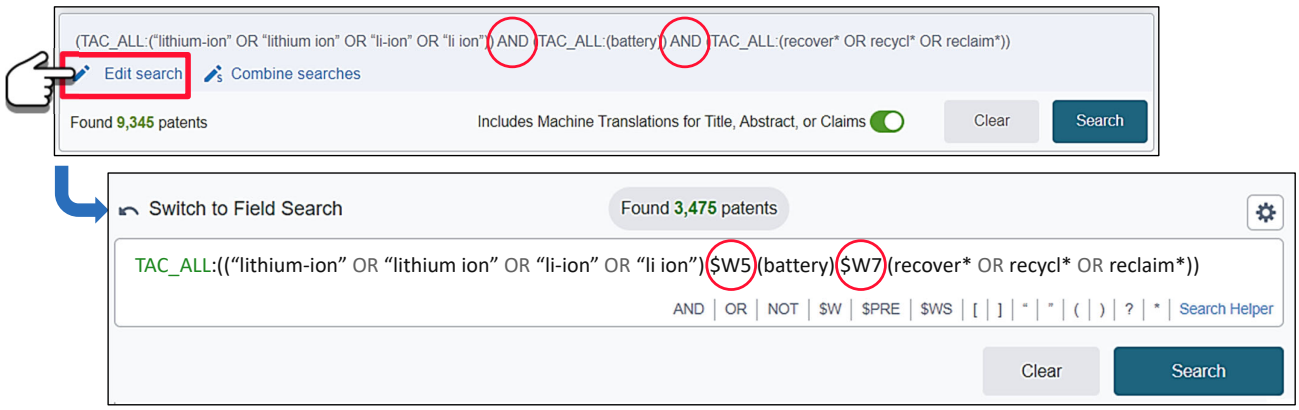
項目	質問式の言語	翻訳される言語	翻訳語の検索対象
標題 抄録	英語 日本語 中国語	英語 日本語 中国語 ドイツ語*1、フランス語*1	オリジナル言語 英語、日本語*2、中国語*2、 ドイツ語*2、　フランス語*2 機械翻訳の英語
クレーム	英語	ー	オリジナル言語の英語 機械翻訳の英語

*1 ドイツ語、フランス語には翻訳されない場合がある
*2 収録されている場合

© 2022 化学情報協会

Edit Search

自動的に作成された検索式は適宜編集できる



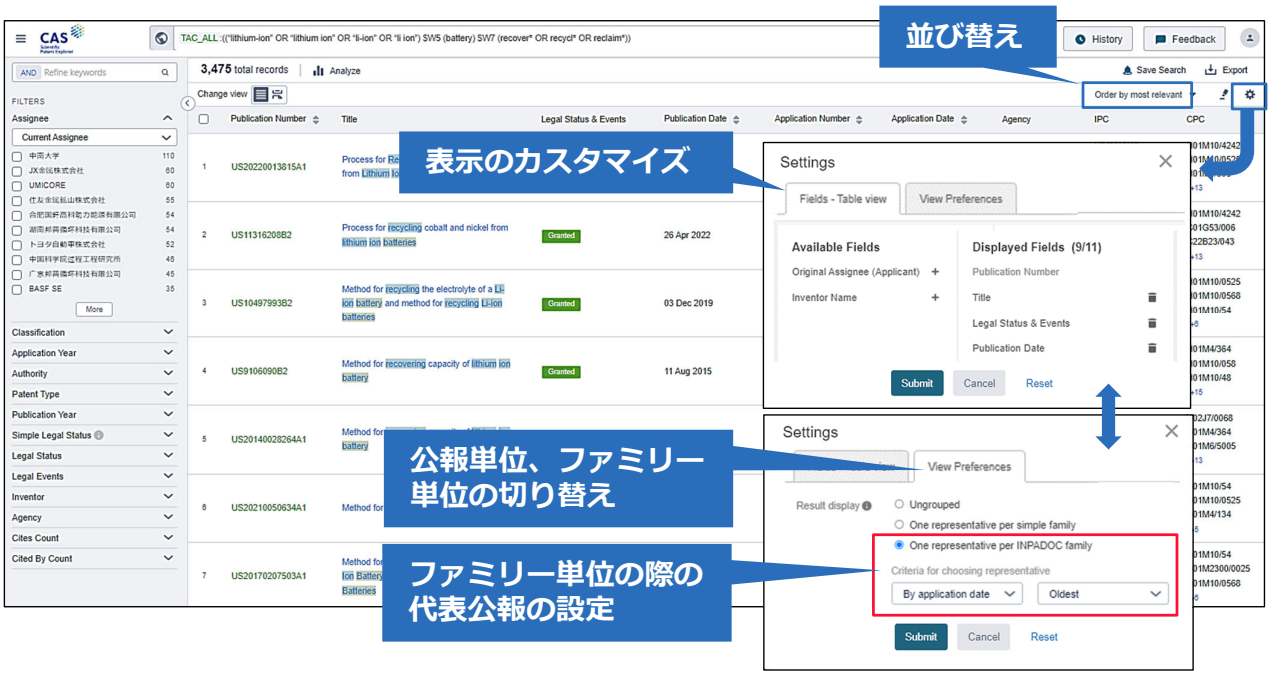
記号	定義
" "	フレーズ検索
*	何文字でもよい * は単語の中間もしくは末尾に使用可能
?	ちょうど一文字 ? は単語の中間もしくは末尾に使用可能
ブール演算子 優先順位	
NOT > AND > OR	

近接演算子	定義
$Wn	入力順序に関係なく検索語の間は n 個以下の語が存在する
$PREn	入力順序通りに検索語の間は n 個以下の語が存在する
$WS	入力順序に関係なく検索語の間は 99 個以下の語が存在する

© 2022 化学情報協会

特許検索の結果 12

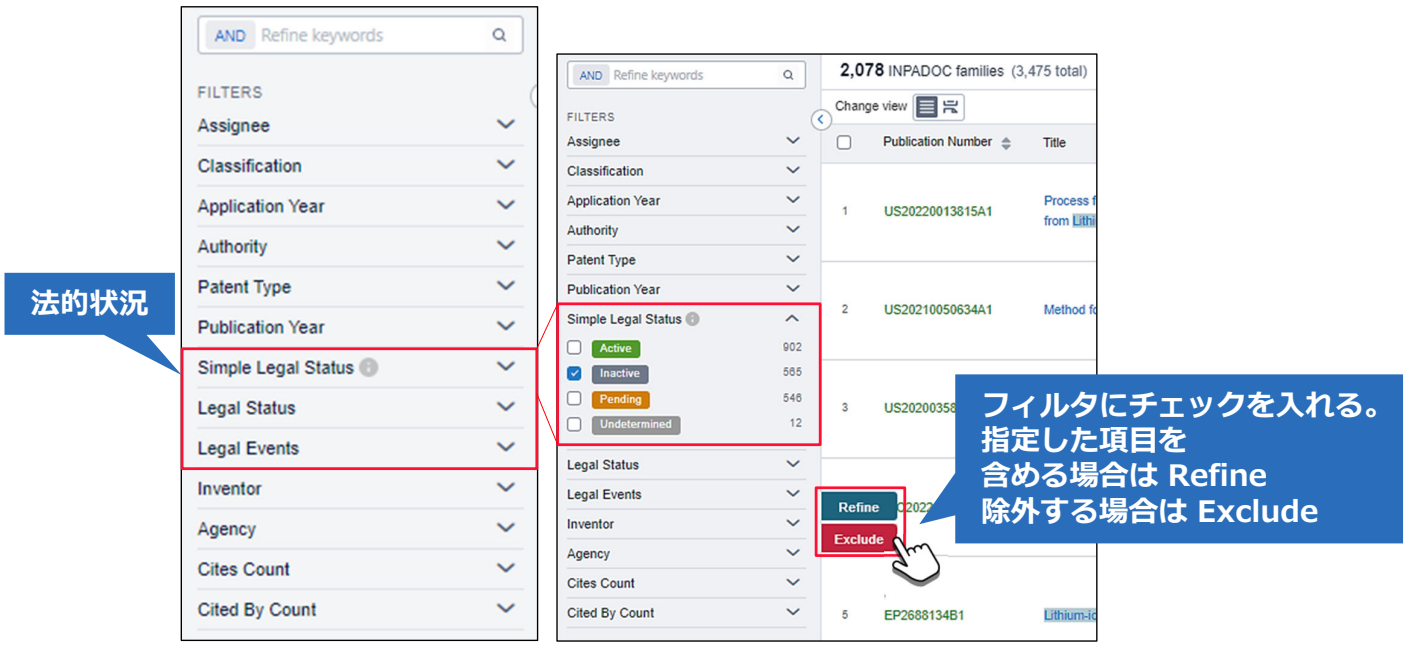
検索結果は見やすい表形式で表示され、カスタマイズできる



© 2022 化学情報協会

絞り込み 13

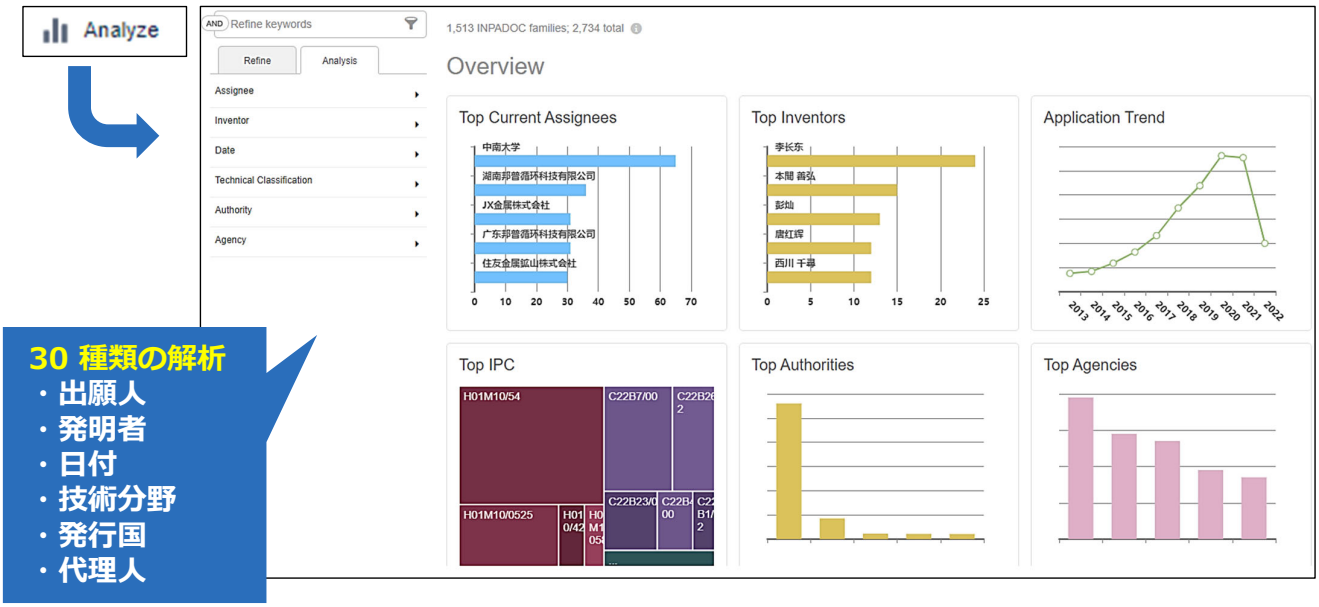
様々な項目のフィルターを利用した絞り込みができ、効率的に回答を確認できる



© 2022 化学情報協会

特許検索結果の解析 14

特許検索の結果をクリック一つで解析でき、様々な観点から特許の動向を把握できる

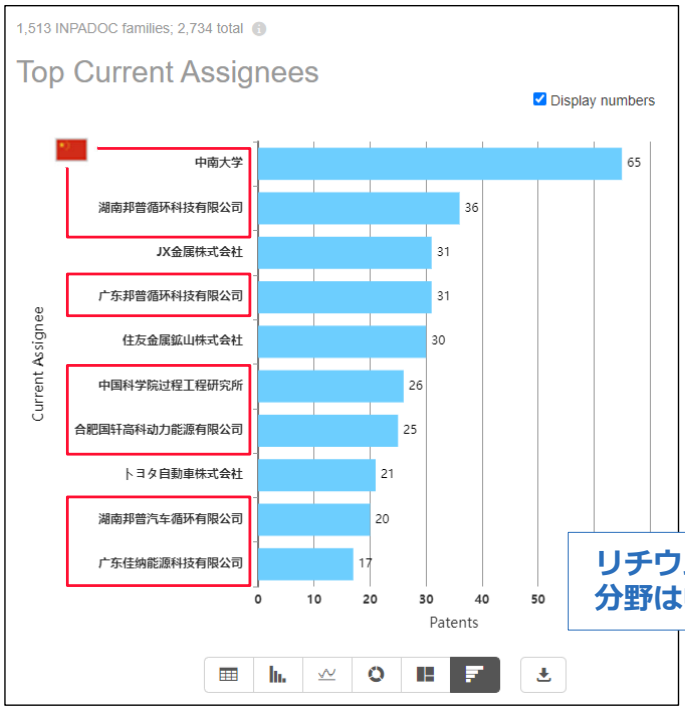


30 種類の解析
- 出願人
- 発明者
- 日付
- 技術分野
- 発行国
- 代理人

© 2022 化学情報協会

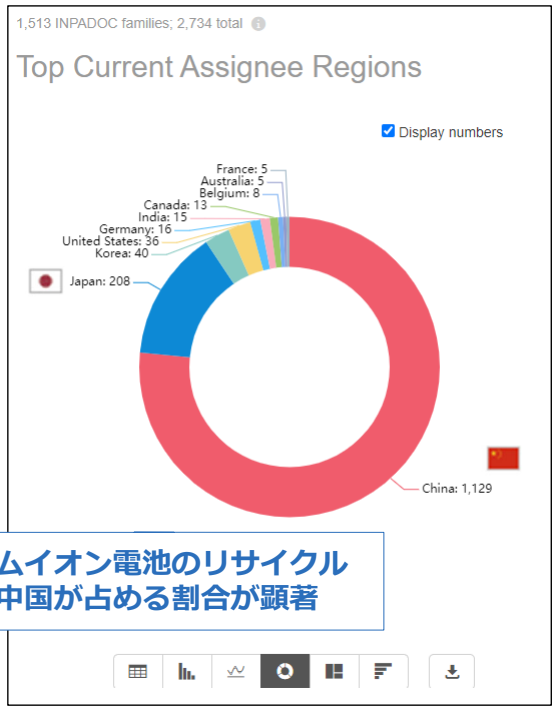
Analyze – Assignee/ Authority 15

上位の特許出願人



Assignee - Top Current Assignees

上位の特許出願人の所属国



リチウムイオン電池のリサイクル分野は中国が占める割合が顕著

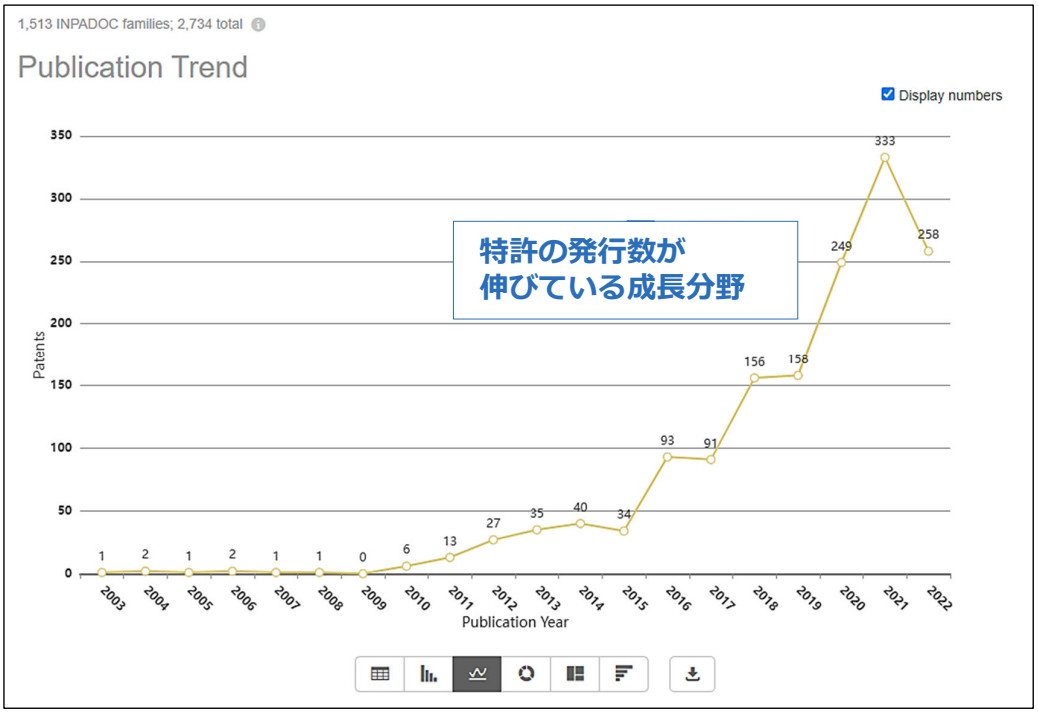
Authority - Top Current Assignee Regions

© 2022 化学情報協会

Analyze – Date

16

特許発行トレンド



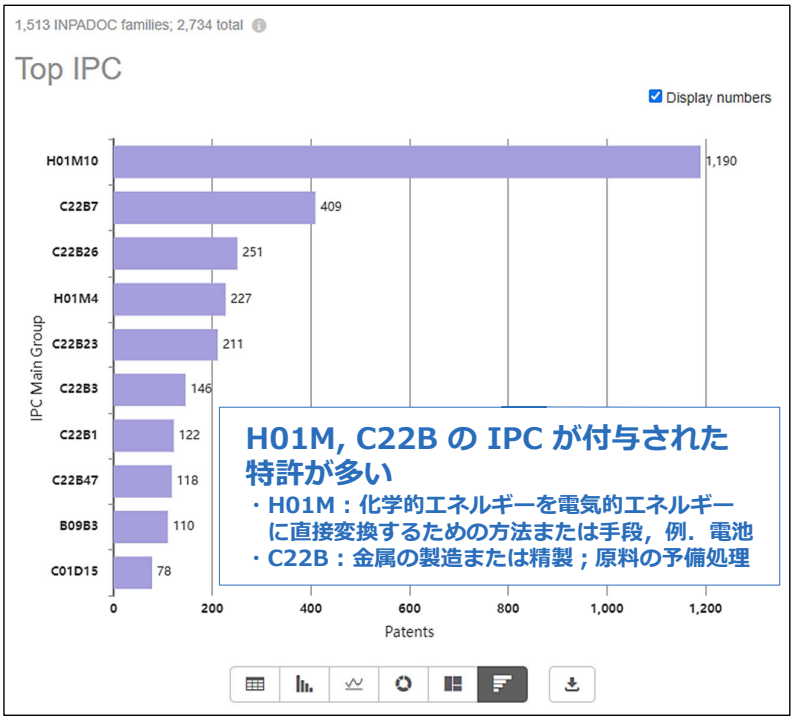
Date – Publication Trend

© 2022 化学情報協会

Analyze – Technical Classification

17

上位 IPC （メイングループで解析）



Technical Classification – Top IPC

H01M10/00
二次電池；その製造

C22B7/00
鉱石以外の他の原材料，例．スクラップ，からの非鉄金属またはその化合物抽出のための処理

C22B26/00
アルカリ金属，アルカリ土類金属またはマグネシウムの採取

H01M4/00
電極

C22B23/00
ニッケルまたはコバルトの採取

C22B3/00
湿式による鉱石または濃縮物からの金属化合物の抽出

C22B1/00
原鉱石またはスクラップの予備的処理一般

C22B47/00
マンガンの採取

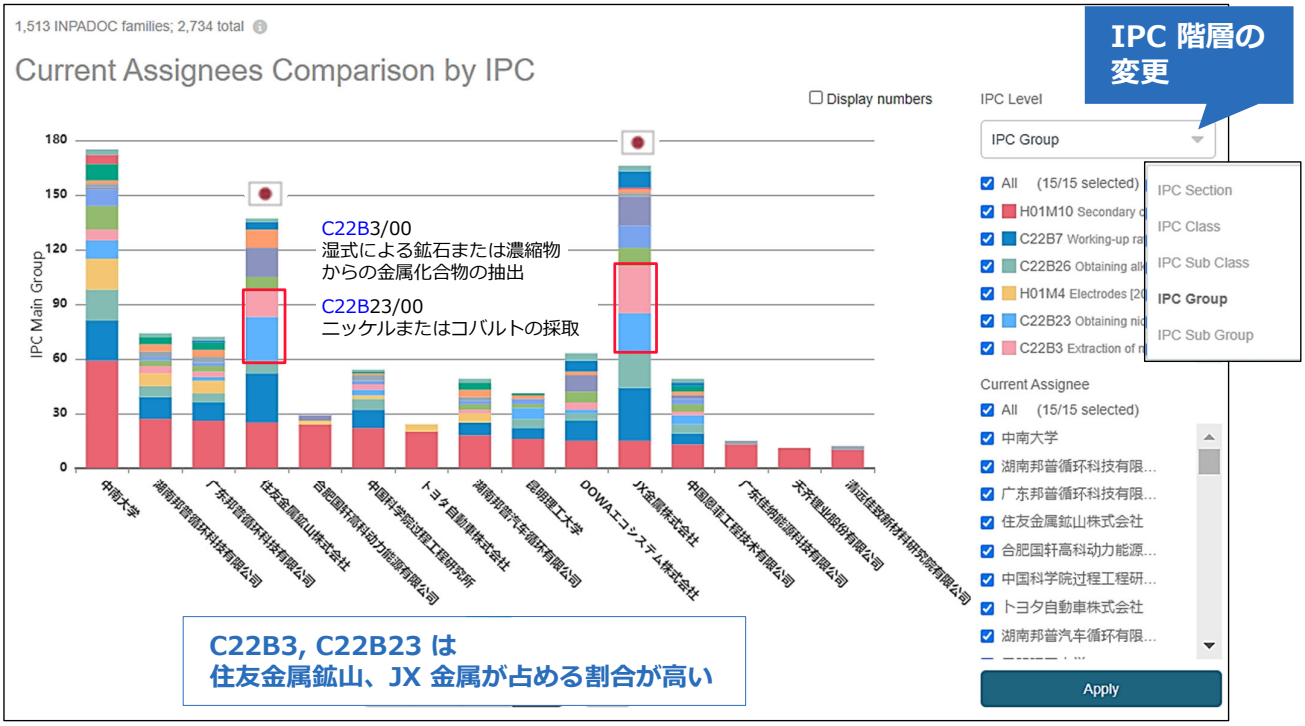
B09B3/00
固体廃棄物の破壊あるいは固体廃棄物の有用物化もしくは無害化

C01D15/00
リチウム化合物

© 2022 化学情報協会

Analyze – Technical Classification

Top IPC による出願人の解析（メイングループで解析）



Technical Classification – Current Assignees Comparison by IPC

© 2022 化学情報協会

活用事例 1 – まとめ

リチウムイオン電池のリサイクルについて、以下の特許動向を確認できた

- 中国がこの分野の中心である
- 直近 10 年で急激に特許が発行されている
- IPC の解析により、日本が強い技術分野がわかる

© 2022 化学情報協会

（参考）

CAS References（CAplus ファイル）は、ベーシック特許の IPC で選別し化学および周辺分野の特許を収録している。CAplus ファイルで選択収録となる分野については、全分野を収録する CAS Scientific Patent Explorer の検索が有効

（例）H01M のリチウムイオン電池関連 IPC

H 電気	
H01	基本的電気素子
H01M	化学的エネルギーを電気的エネルギーに直接変換するための方法または手段，例．電池
H01M10/00	二次電池；その製造
H01M10/05	・非水電解質二次電池
H01M10/052	・・リチウム二次電池
H01M10/0525	・・・ロッキングチェア電池，すなわち両方の電極でリチウムの挿入を伴うもの；リチウムイオン電池

（例） CAplus ファイルの H01M 特許収録範囲 (www.cas.org/support/documentation/references/patentcoverage)

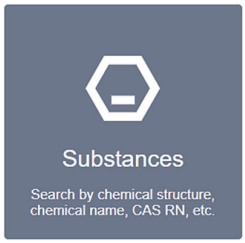
収録保証 IPC (IPC リスト I)	
H01M	4/13-4/1399, 4/24-4/60, 4/86-6/00, 6/04-6/06, 6/14-6/20, 6/24-6/28, 8/06-8/22

選択収録 IPC (IPC リスト II)	
H01M	00/00-4/12, 4/14-4/23, 4/62-4/84, 6/02, 6/08-6/12, 6/22, 6/30-8/04992, 8/24-16/00

活用事例 2
物質関連特許の動向を調査する

イーライリリーの乳がんの治療薬アベマシクリブと
その誘導体について、特許の動向を調査する

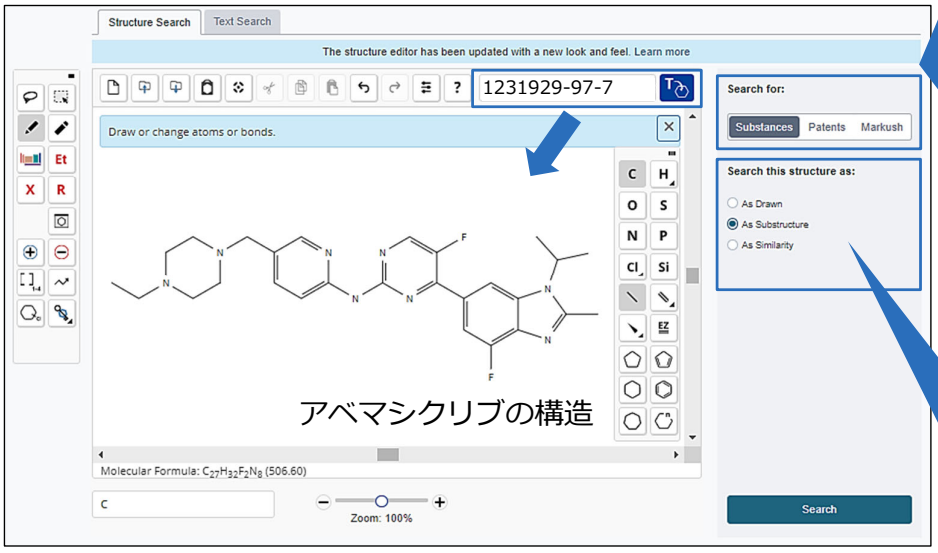
- 化学物質の構造検索
- 化学物質と特許を関連づけた解析



© 2022 化学情報協会

化学物質の検索 23

構造検索は、検索対象と構造検索タイプを選んで
検索する



検索対象
- **Substances**
  化学物質の検索
- **Patents**
  化学物質に関する
  特許情報の検索
- **Markush**
  マルクーシュ構造
  を含む特許検索

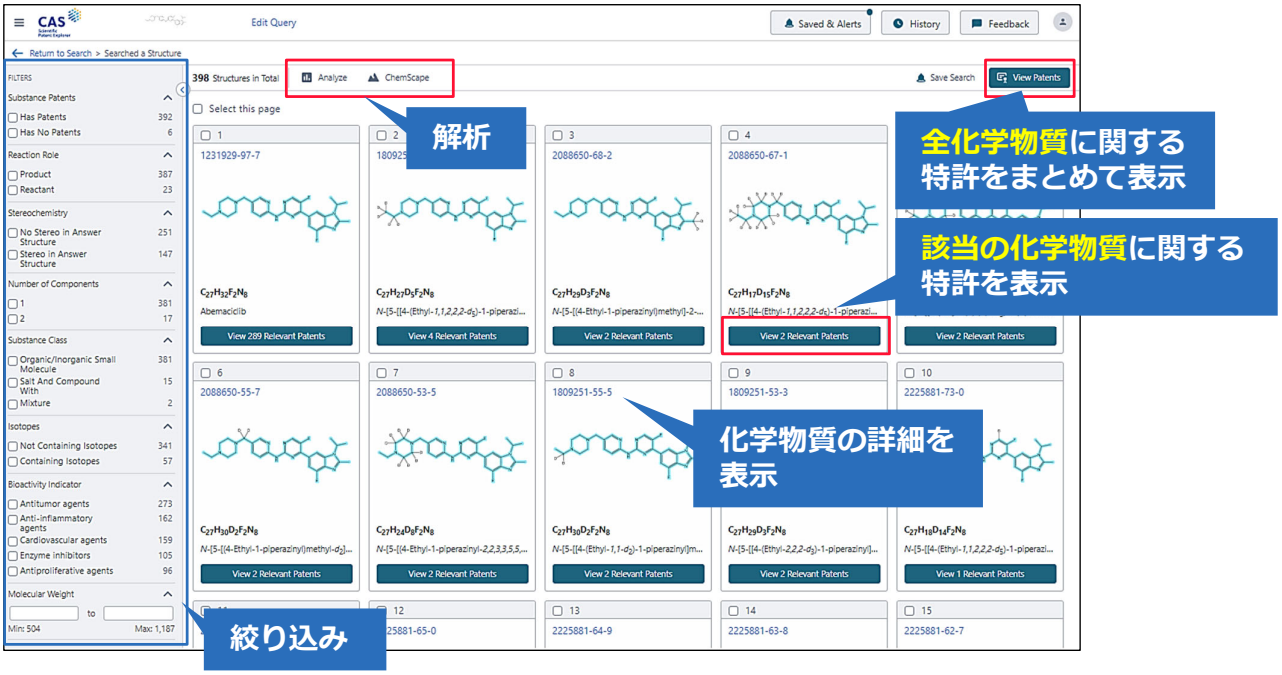
構造検索タイプ
- **As Drawn**
  完全一致検索
- **As Substructure**
  部分構造検索
- **As Similarity**
  類似性構造検索

© 2022 化学情報協会

化学物質検索結果

24

Substances の検索結果は化学物質の一覧で表示される



© 2022 化学情報協会

化学物質検索結果の解析

25

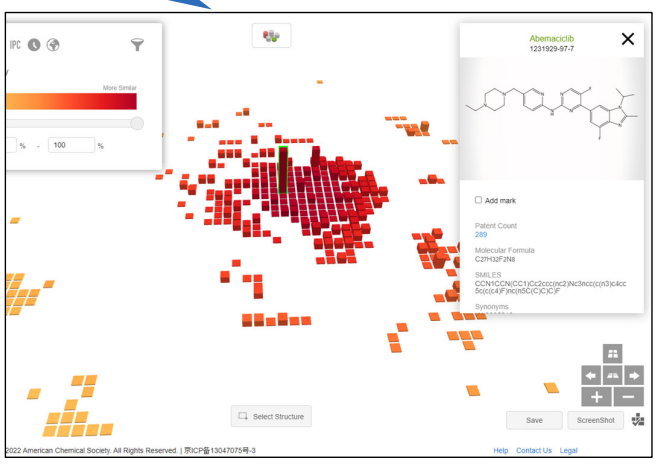
化学物質検索の結果を二通りの方法で解析できる

Analyze

化学物質と関連する特許情報を紐づけて解析

ChemScape

化学物質の構造の類似性と関連する特許情報を紐づけて解析



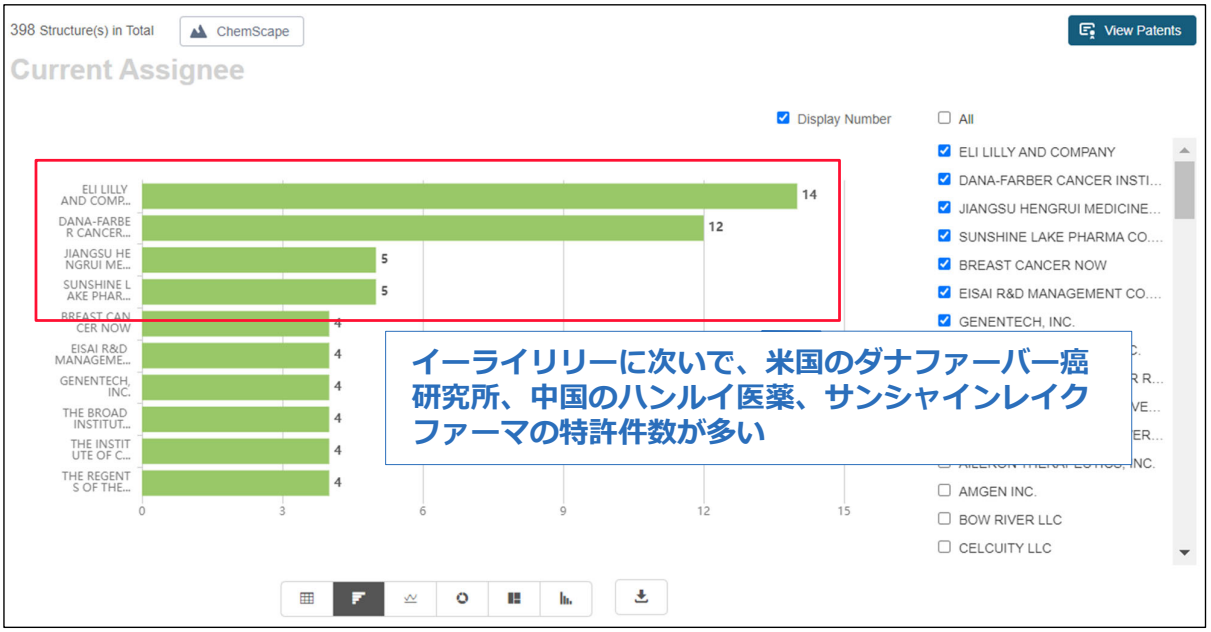
8 種類の解析

© 2022 化学情報協会

Analyze – Current Assignee 26

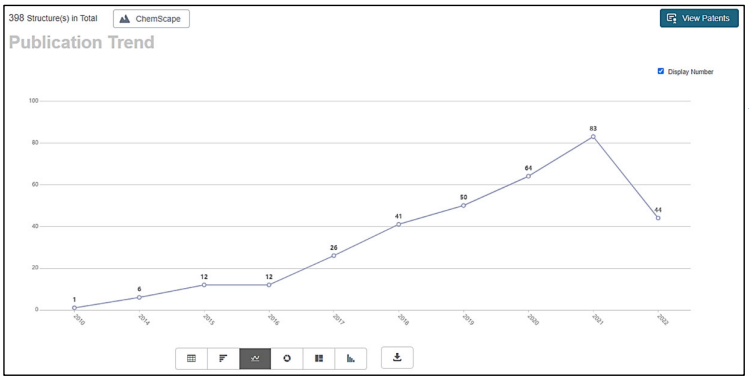
Analyze は化学物質に関する特許情報を解析できる

上位の特許出願人



© 2022 化学情報協会

Analyze - Publication Trend/Assignee Countries 27



特許発行トレンド

特許の発行は増加傾向

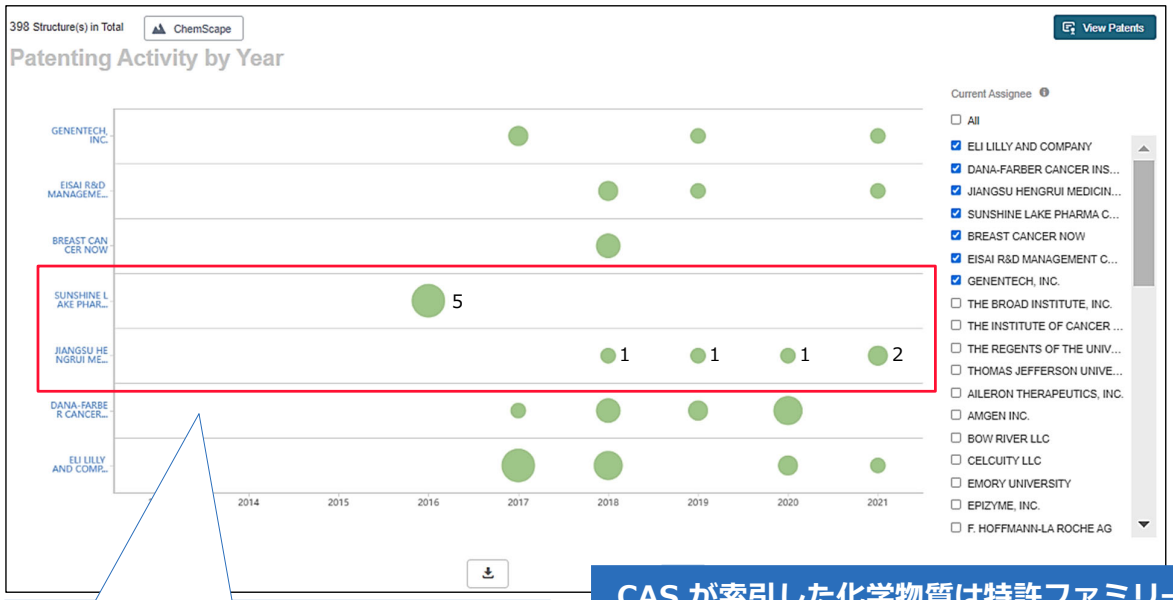
出願人の所属国

米国や中国所属の特許出願人が多い

© 2022 化学情報協会

Analyze – Patenting Activity by Year 28

特許出願人ごとの特許発行年解析



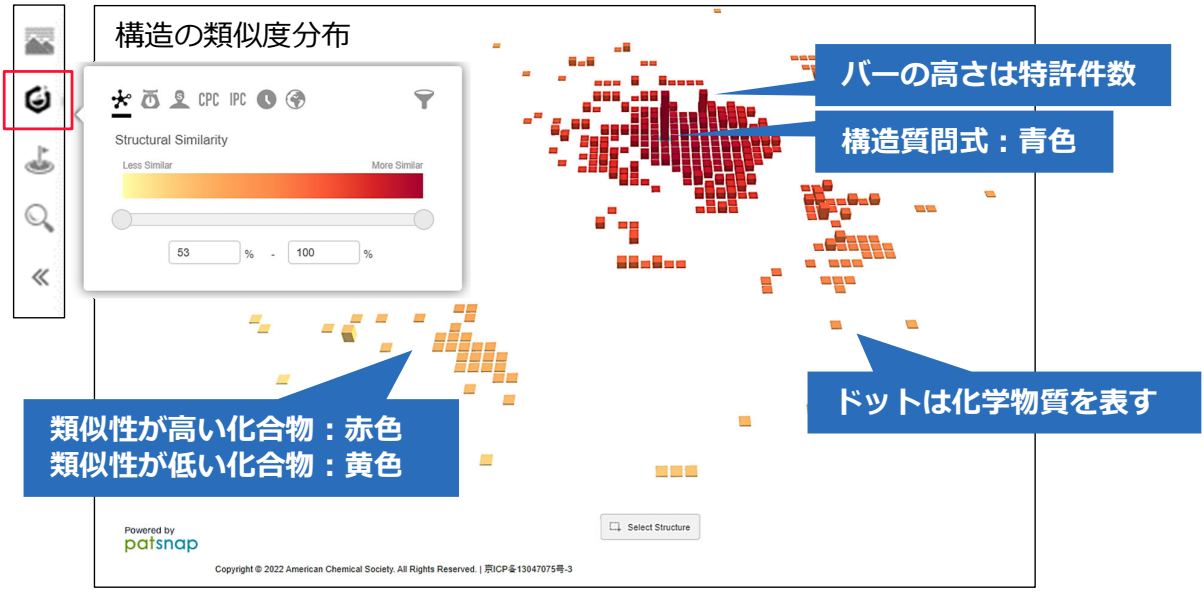
中国のハンルイ医薬、サンシャインレイクファーマは同じ特許件数だが、発行年は異なっていることがわかる

CAS が索引した化学物質は特許ファミリーのうち最初に受け入れた代表の特許（ベーシック特許）に索引されており、この特許で解析している

© 2022 化学情報協会

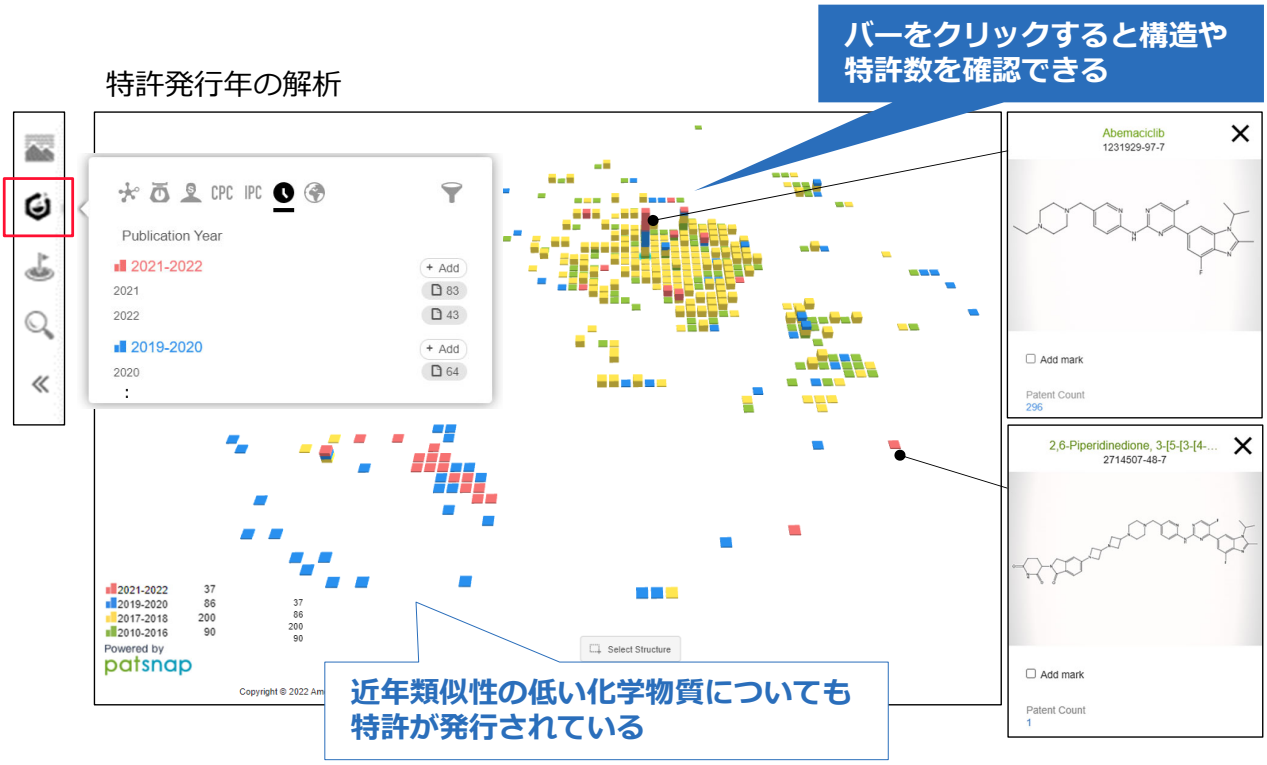
Chemscape Analysis 29

Chemscape Analysis は、構造の類似性と特許情報を紐づけた新しい視点の解析ができる



© 2022 化学情報協会

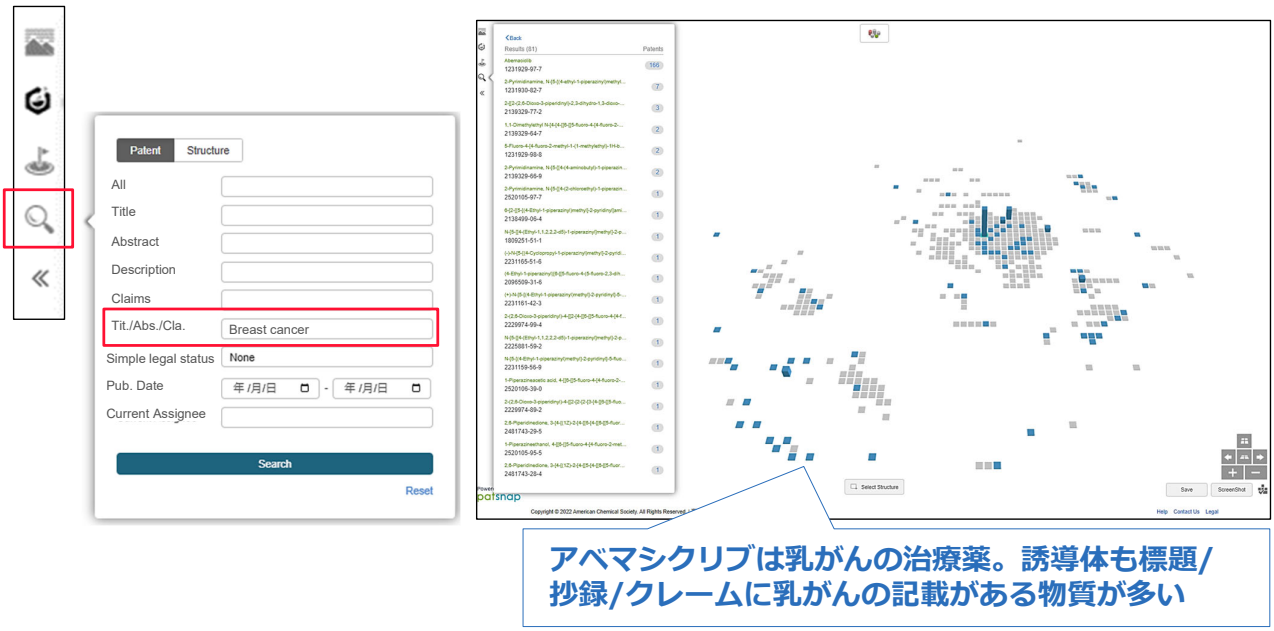
Chemscape Analysis – 特許発行年の解析 30



© 2022 化学情報協会

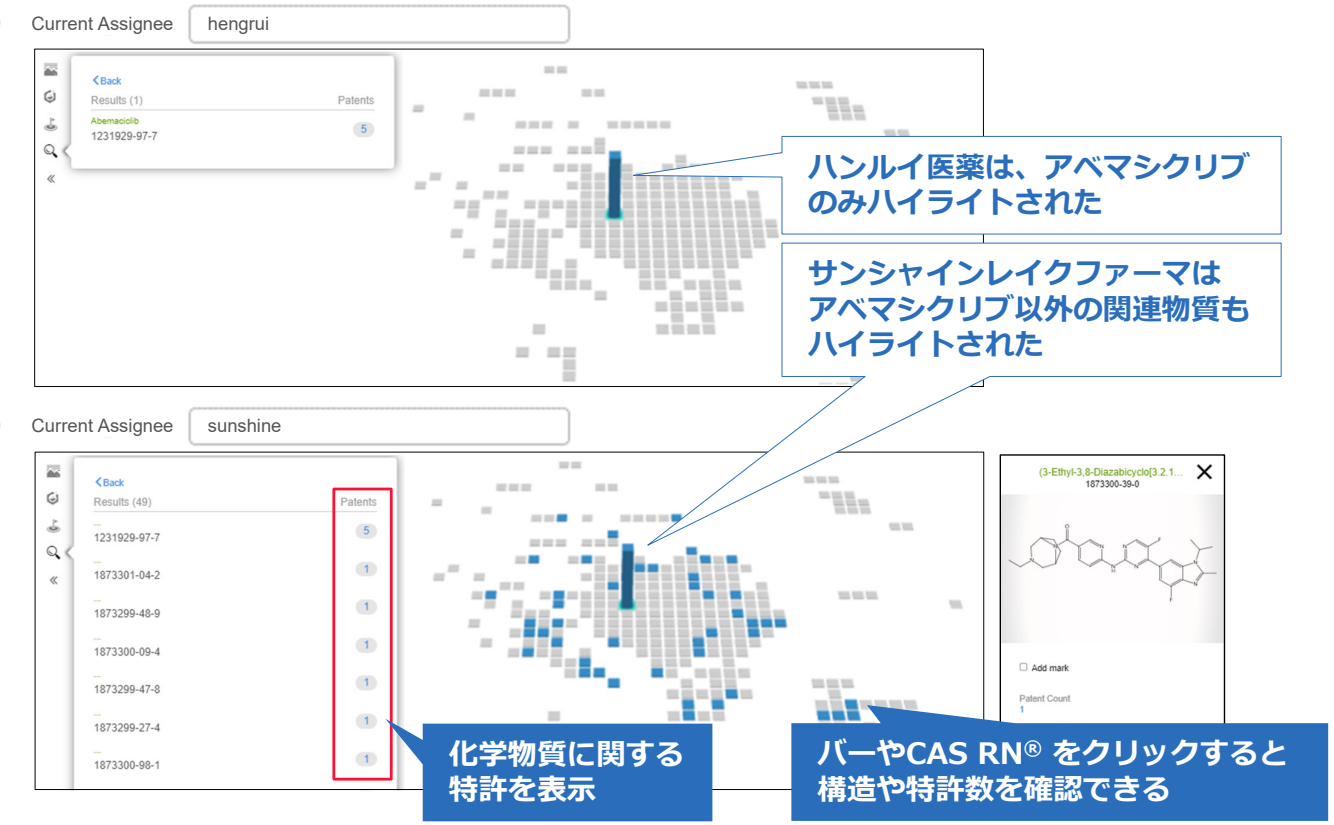
Chemscape Analysis – キーワード検索 31

特許に含まれるキーワードや書誌情報でマップ内を検索できる



© 2022 化学情報協会

Chemscape Analysis – 特許出願人検索

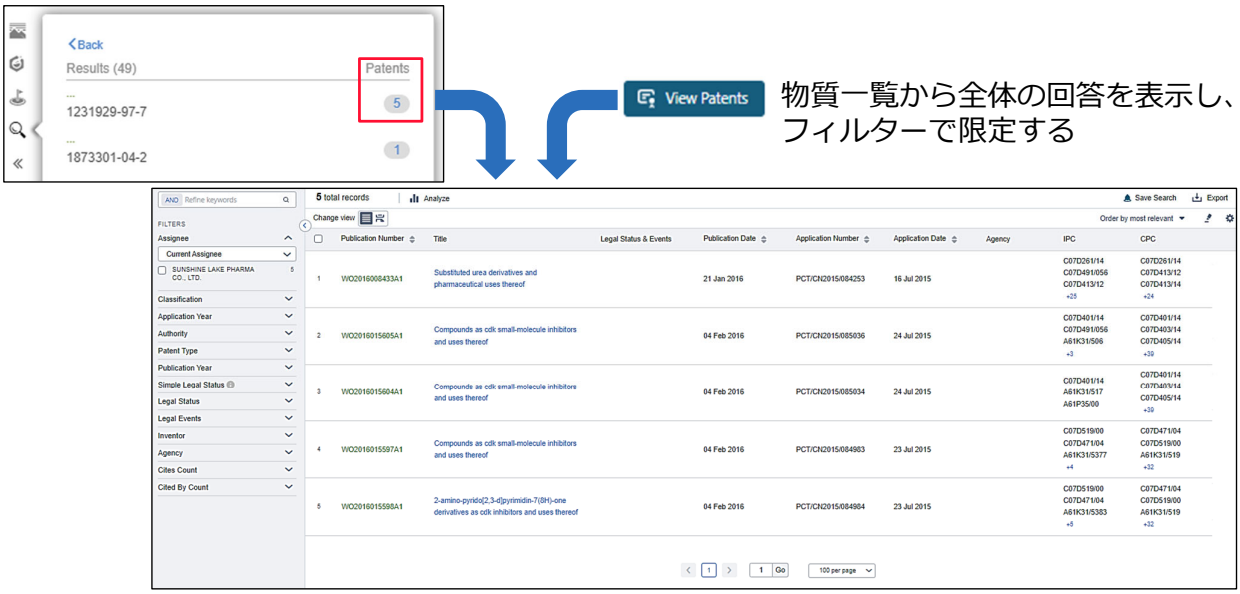


© 2022 化学情報協会

化学物質から文献の検索

検索した化学物質に関する特許は、View Patents や Chemscape Analysis のリンクから確認できる

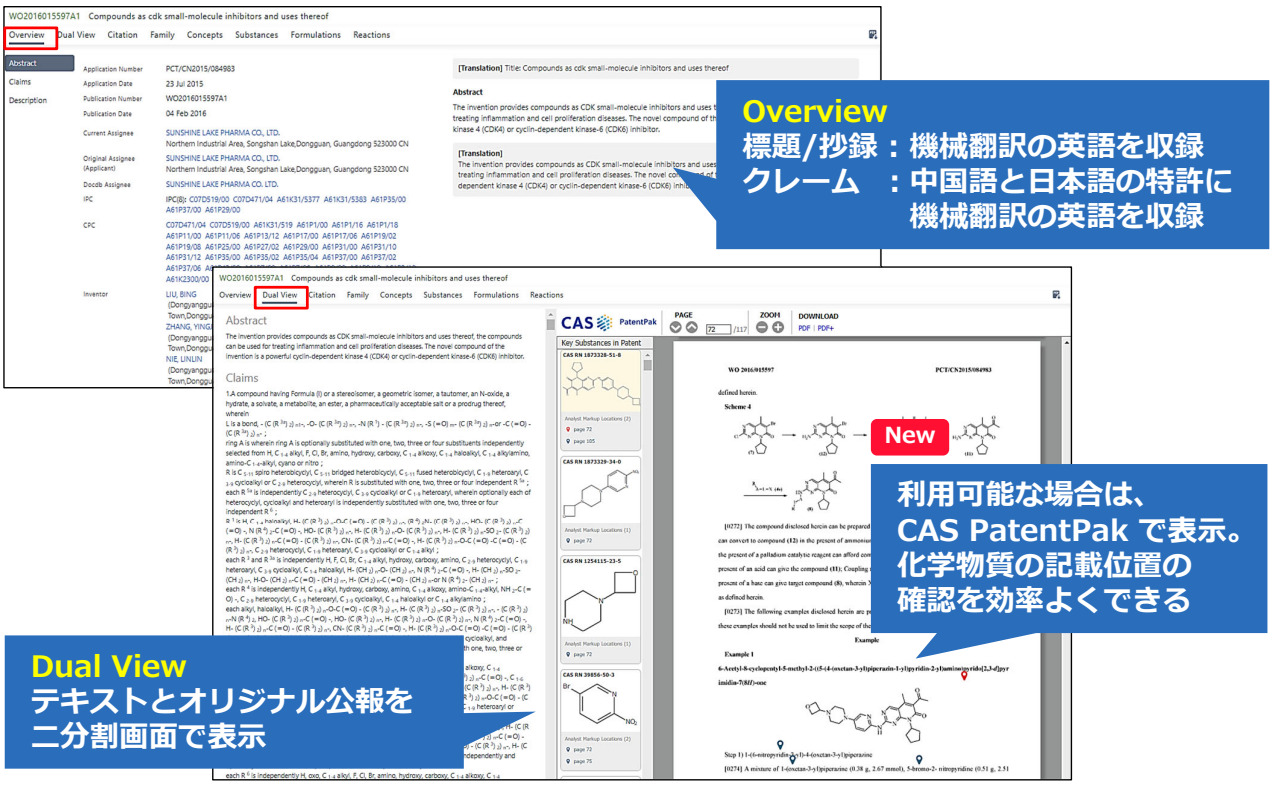
Chemscape Analysis から個別の回答を表示する



© 2022 化学情報協会

Overview, Dual View

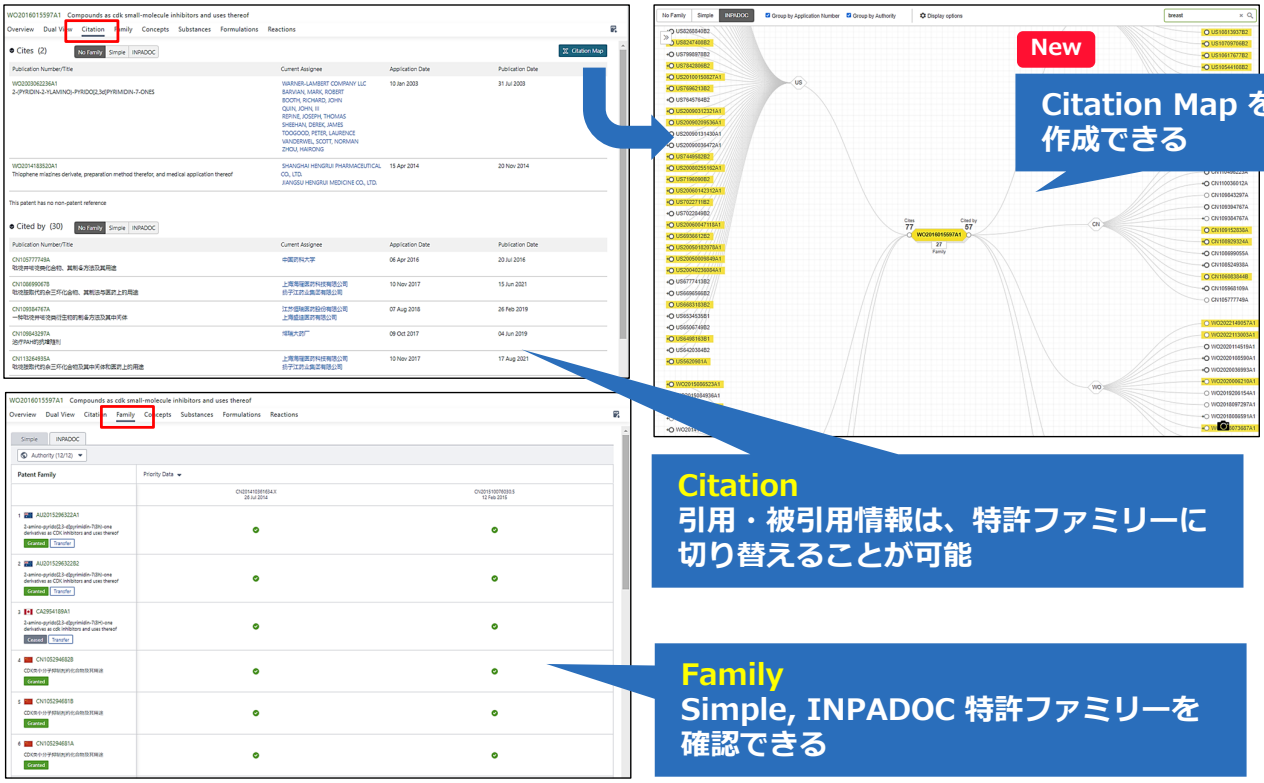
34



© 2022 化学情報協会

Citation, Family

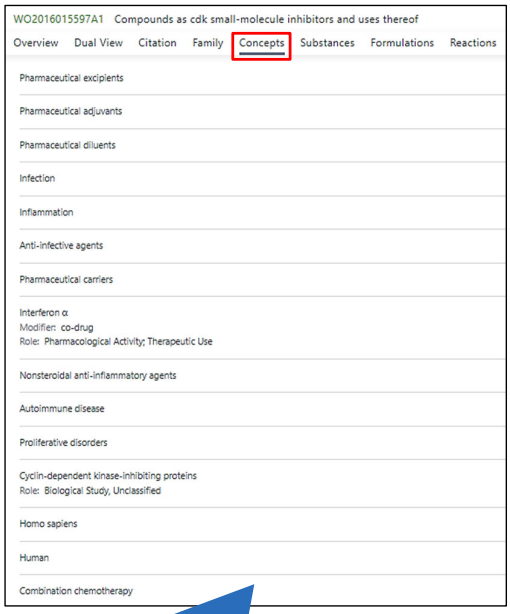
35

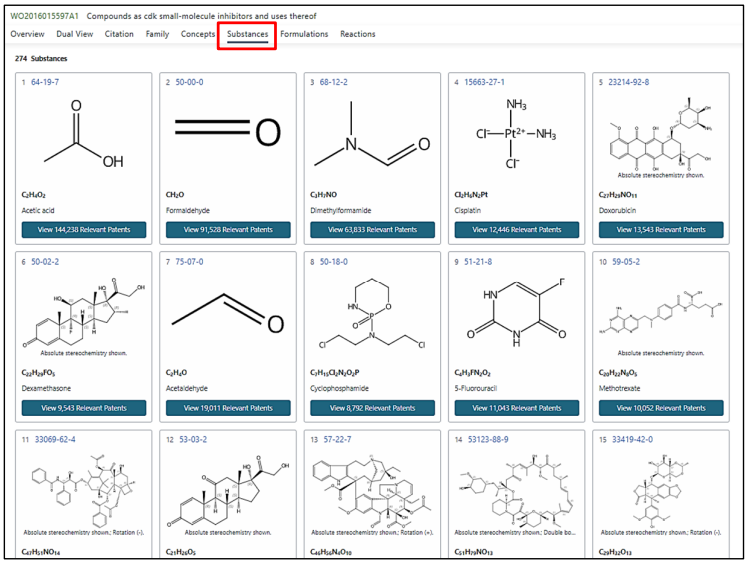


© 2022 化学情報協会

Concepts, Substances

36





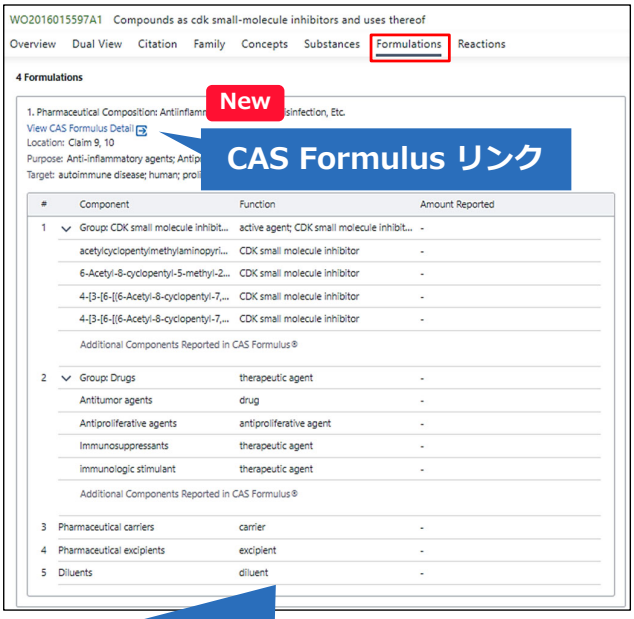
Concepts (CAS References 由来)
特許中で発明者が強調している点や主題に関わる概念を収録

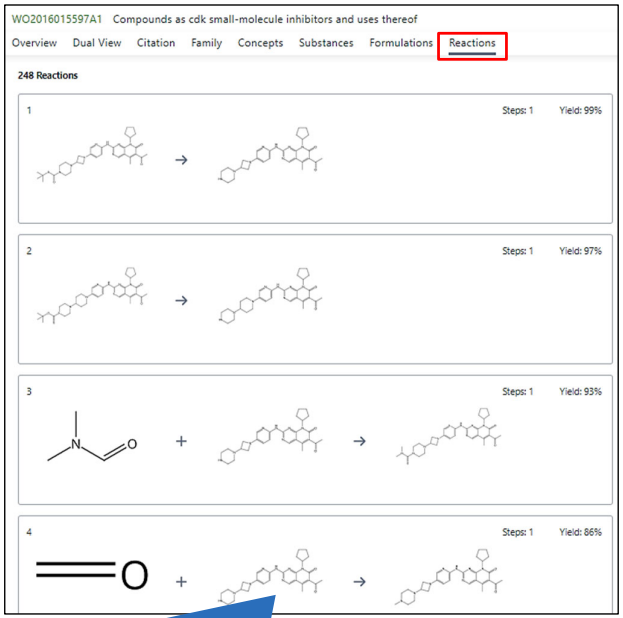
Substances (CAS References 由来)
発明に関する重要な化学物質が索引されており、特許内容の把握が容易になる

© 2022 化学情報協会

Formulations , Reactions

37





Formulations (CAS Formulations 由来)
特許中の製剤・配合の情報が、成分名、機能などと共に表形式でまとめられている

Reactions (CAS Reactions 由来)
特許中の重要な有機化学反応情報を収録

© 2022 化学情報協会

活用事例 2 – まとめ 38

イーライリリーの乳がんの治療薬アベマシクリブと
その誘導体について、以下の特許の動向を確認できた

- 米国や中国所属の特許出願人が多い
- 近年、類似性の低い化学物質についても特許が発行されている
- 特許件数は同等でも、特許発行の傾向や研究対象物質が異なることがわかった

© 2022 化学情報協会

まとめ 39

CAS Scientific Patent Explorer は

- 簡単な操作で特許調査を行える
- 広範な特許情報を対象に検索できる
- キーワードだけでなく化学物質の構造式からも特許調査できる
- 個々の内容を確認するだけでなく、回答全体を解析し、様々な観点から特許の動向を把握できる

© 2022 化学情報協会

強化一覧

強化一覧（ 2021 年 6 月以降） 41

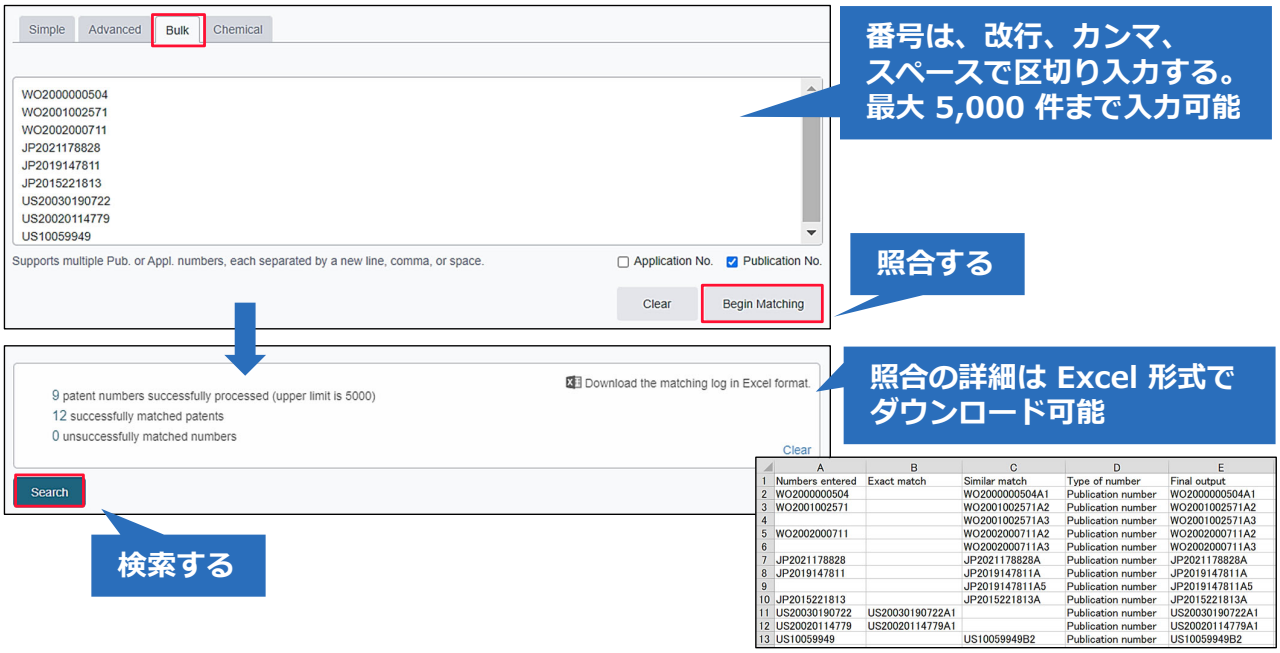
	強化内容	開始時期
特許	ヒットタームのハイライト機能の追加	2021.6
特許	引用・被引用情報のビジュアル化機能 Citation Map を追加	2021.6
特許	INPADOC 特許ファミリーの追加	2021.6
特許	CAS Formulus リンクが利用可能に	2021.6
特許	CAS PatentPak の搭載 CAS PatentPak の索引物質にクレーム記載情報を追加	2022.3 2022.8
特許	★中国語や日本語の特許に機械翻訳の英語クレームを追加	2022.4
特許	★複数の特許番号，出願番号の一括検索機能 Bulk Patent Search を追加	2022.6
物質	検索結果のフィルター追加 - Substance Patents, Bioactivity Indicator	2021.8
物質	構造作図画面の変更	2022.7
特許・物質	データの更新頻度が毎週に	2021.10
特許・物質	★History（検索履歴）ページの追加	2021.11
特許・物質	★質問式の保存とアラート機能の追加 Substance Search の質問式の保存とアラート機能を追加	2022.1 2022.3

★の強化を次ページから紹介

© 2022 化学情報協会

複数の特許番号，出願番号の一括検索機能を追加 42

複数の特許番号、出願番号を一括検索できる機能
Bulk Patent Search が追加された

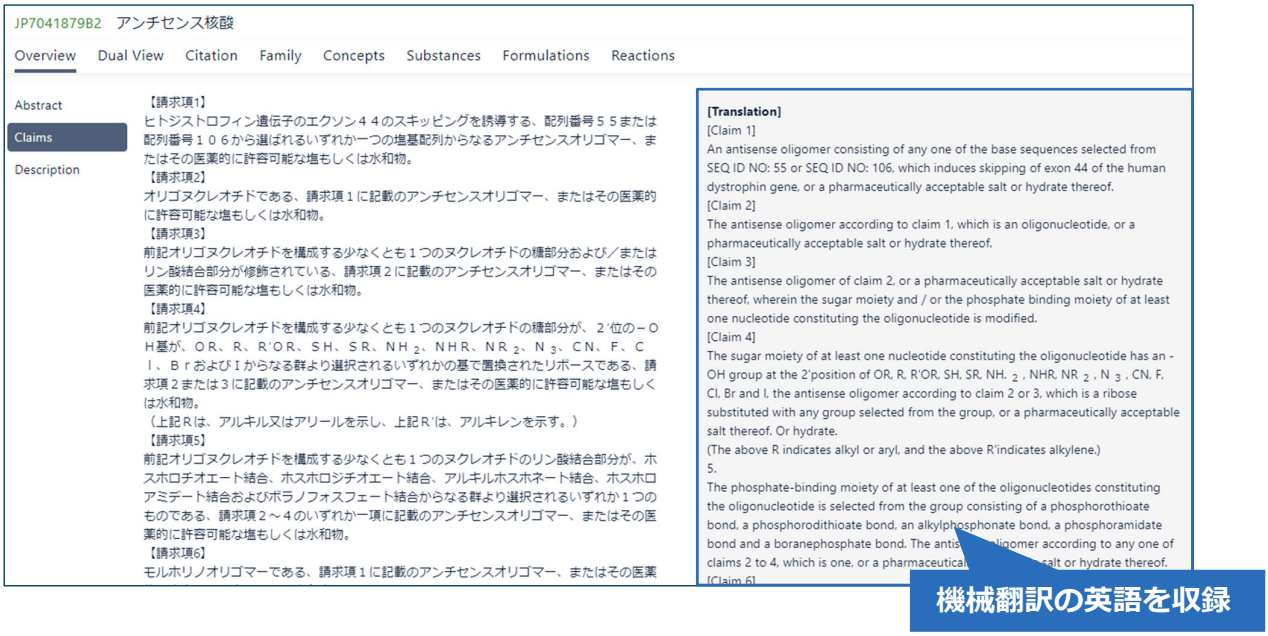


	A	B	C	D	E
1	Numbers entered	Exact match	Similar match	Type of number	Final output
2	WO2000000504		WO2000000504A1	Publication number	WO2000000504A1
3	WO2001002571		WO2001002571A2	Publication number	WO2001002571A2
4			WO2001002571A3	Publication number	WO2001002571A3
5	WO2002000711		WO2002000711A2	Publication number	WO2002000711A2
6			WO2002000711A3	Publication number	WO2002000711A3
7	JP2021178828		JP2021178828A	Publication number	JP2021178828A
8	JP2019147811		JP2019147811A	Publication number	JP2019147811A
9			JP2019147811A5	Publication number	JP2019147811A5
10	JP2015221813		JP2015221813A	Publication number	JP2015221813A
11	US20030190722	US20030190722A1		Publication number	US20030190722A1
12	US20020114779	US20020114779A1		Publication number	US20020114779A1
13	US10059949		US10059949B2	Publication number	US10059949B2

© 2022 化学情報協会

中国語や日本語の特許に英語クレームを追加 43

中国語や日本語の特許に機械翻訳の英語クレームが追加された



© 2022 化学情報協会

質問式の保存とアラート機能の追加 44

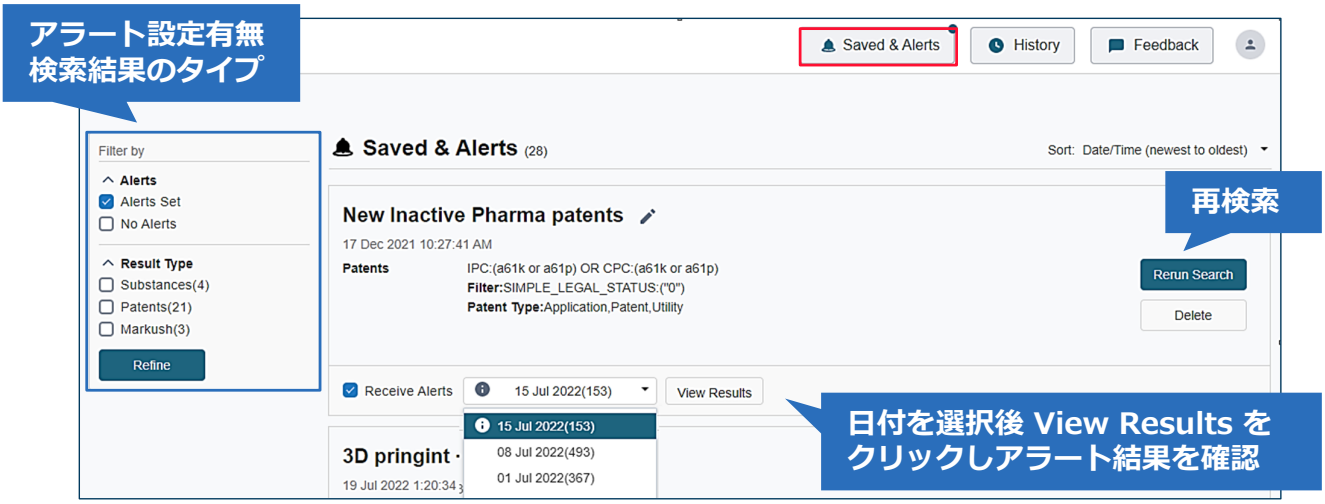
質問式の保存とアラート機能が利用可能になった

（注意）・Substances で化学物質検索を行い、その物質に関連する特許を検索した結果に対してアラート設定すると、設定した時点でヒットした物質に関する新規特許の情報が得られる（アラート実行時に物質は再検索されない）
・Substances でマルクーシュ構造検索し View Patents 後に設定したアラートは 0 件になる

© 2022 化学情報協会

質問式の保存とアラート機能の追加 45

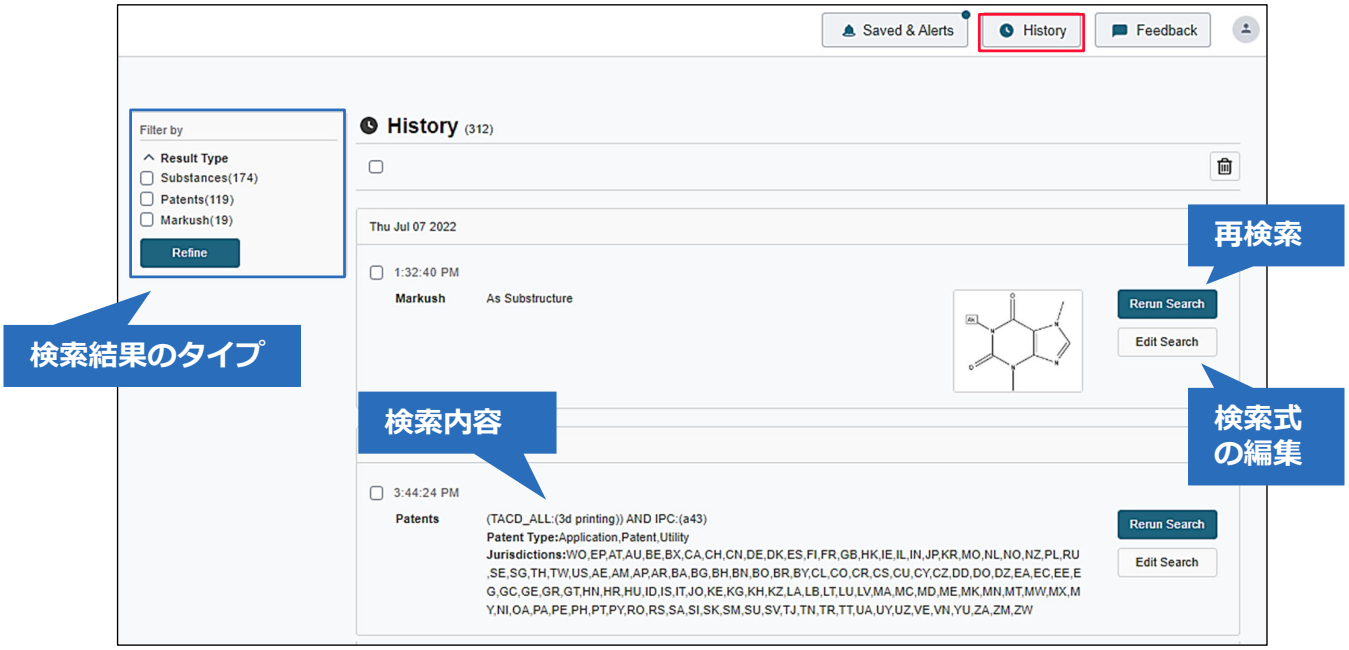
保存した質問式やアラート結果は Saved & Alerts で確認できる



© 2022 化学情報協会

History（検索履歴）ページの追加 46

検索に用いた構造やテキストの質問式が自動的に
History ページに保存されるようになった



© 2022 化学情報協会

知財戦略＆特許調査に差がつく

STN IP Protection Suite セミナー

Biosequences Search

～ CAS STNext® の新しい配列検索機能 ～

Biosequences Search
～ CAS STNext® の新しい配列検索機能 ～

化学情報協会　情報事業部

目次

1. Biosequences Search とは
2. 活用事例
3. 最近の強化

© 2022 化学情報協会

Biosequences Search とは

Biosequences Search 3

既存の配列データベースとは独立した
CAS STNext の新しい配列検索機能

- 膨大な配列コンテンツ
- 3 つの配列検索プログラム
- シームレスな検索

© 2022 化学情報協会

膨大な配列コンテンツ 4

特許・非特許文献由来の配列を収録

(2022 年 8 月)

資料種類	特許	非特許文献
収録件数	7 億 5000 万件以上の配列 (156 万件以上の特許)	16 億 2000 万件以上の配列 (236 万件以上の非特許文献)
収録源	- 主要国を中心にした 73 カ国の特許 (主要国：EP, WO, US, CN, JP, KR, CA) - REGISTRY ファイル	- NCBI New - REGISTRY ファイル

CAS の独自のルールに基づいた収録

参考：REGISTRY ファイルに収録されている配列

- CAplus ファイルに索引された特許・非特許文献に記載の重要な配列
  (特許由来：2210 万件以上、非特許文献由来：2380 万件以上)
- GenBank 由来の配列 (4980 万件以上)

© 2022 化学情報協会

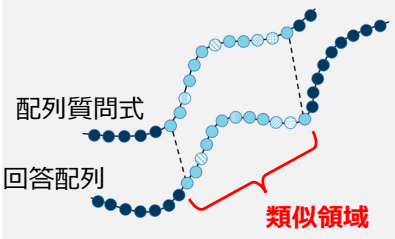
3 つの配列検索プログラム 5

BLAST をベースにした検索プログラム

- 類似配列だけでなく、CDR (相補性決定領域) やパターン配列を対象に検索できる

BLAST 配列検索

局所的に類似した配列の検索



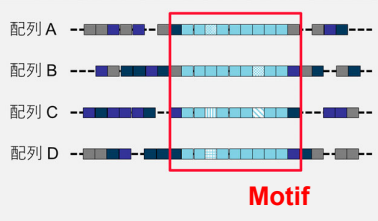
CDR 配列検索

抗体と T 細胞受容体の CDR を指定した検索

CDR : 抗原に特異的な部分

Motif 配列検索

DNA, RNA, タンパク質中の短いパターン配列の検索



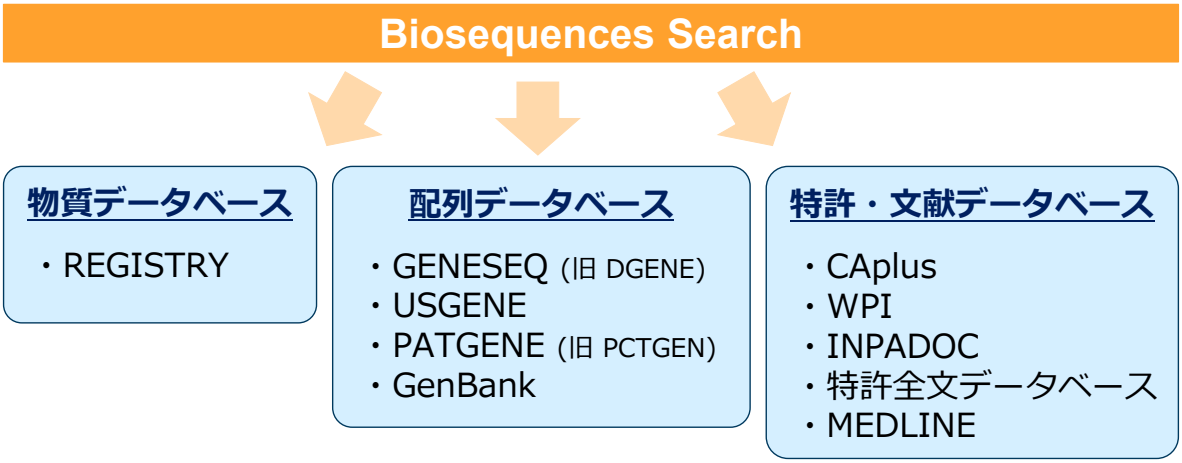
Motif : ある一定の機能を持つと期待される特徴的な共通の配列や構造

© 2022 化学情報協会

シームレスな検索 6

得られた配列情報とその出典情報を CAS STNext のファイルで再現できる

- 配列や出典情報の詳細を表示するだけでなく、キーワードによる絞り込みや出願人による解析などが可能



© 2022 化学情報協会

活用事例

活用事例 8

REGN 10933 の CDR として知られている下記のセグメントを持つタンパク質を検索し、SARS-CoV-2 に関する特許を調査する

- **CDR1 : DYYMS**
- **CDR2 : YITYSGSTIYYADSVKG**
- **CDR3 : DRGTTMVPFDY**

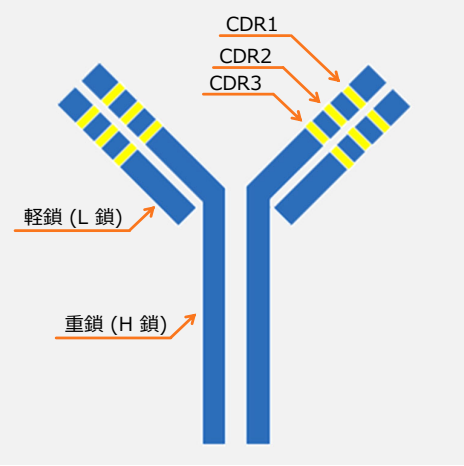
REGN 10933
Regeneron 社によって開発された SARS-CoV-2 の中和抗体。成分名は Casirivimab

© 2022 化学情報協会

抗体と CDR (相補性決定領域) 9

抗体とは、抗原と特異的に結合して生体内から除去するためのタンパク質

- 抗体は 2 本の軽鎖 (L 鎖) と重鎖 (H 鎖) から成る



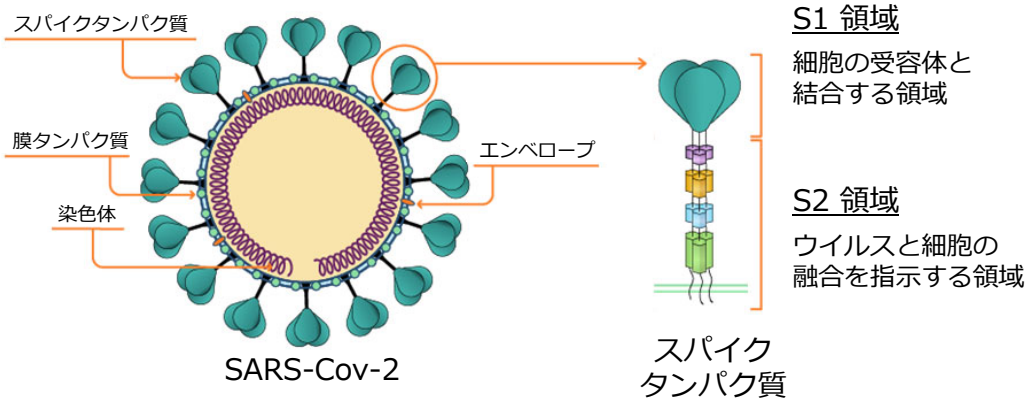
- CDR は抗原結合部位であり、L 鎖と H 鎖にそれぞれ CDR1, CDR2, CDR3 がある
- 抗原ごとに異なる配列を持つ

© 2022 化学情報協会

新型コロナウイルス(SARS-Cov-2) 10

SARS-Cov-2 のスパイクタンパク質の S1 領域が
細胞の受容体結合ドメインと結合することで感染する

- スパイクタンパク質の変異は、抗体への抵抗性や細胞への侵入速度、感染力と関連する



Omicron variant accelerates need for more diverse COVID-19 vaccine strategy
https://www.cas.org/resources/blog/omicron-variant

© 2022 化学情報協会

検索の流れ 11

Biosequences Search

1. CDR 配列検索
2. 回答の絞り込み
   (3 つのセグメントを持つ配列に限定)

REGISTRY ファイル

3. ヒットした配列の名称や修飾の有無を確認する

CAplus ファイル

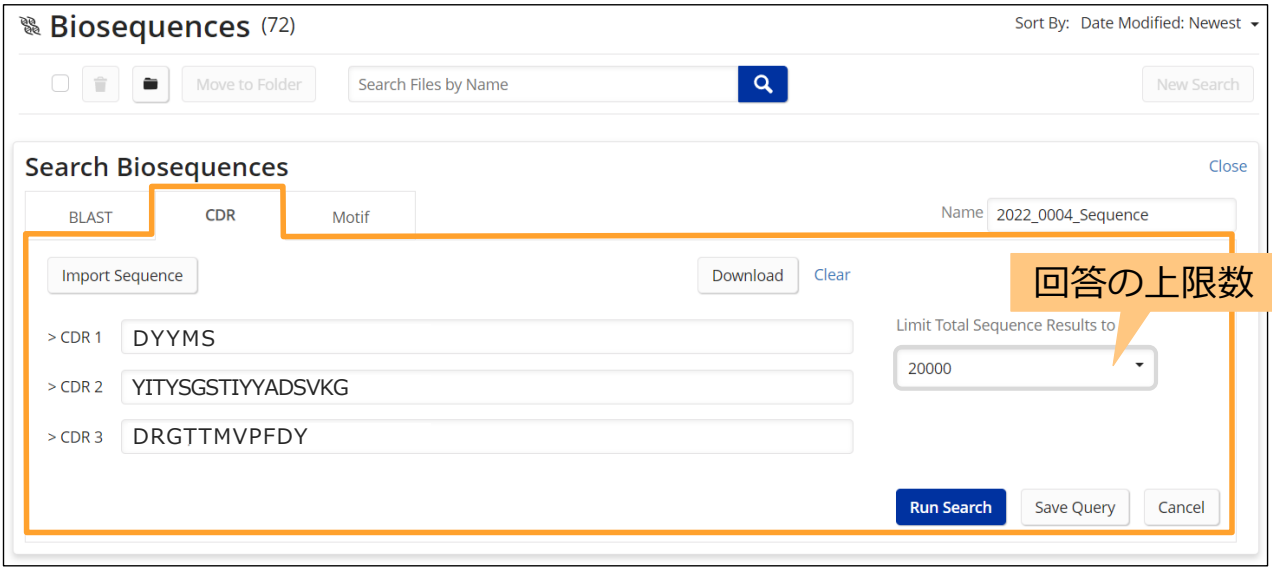
4. ヒットした配列の特許情報を検索する
5. 特許出願人の出願動向を解析する

© 2022 化学情報協会

1. CDR 配列検索 12

CDR タブを選択して配列質問式を入力し、回答の上限数を選択する

- 最大で 3 つの配列質問式を使って検索できる

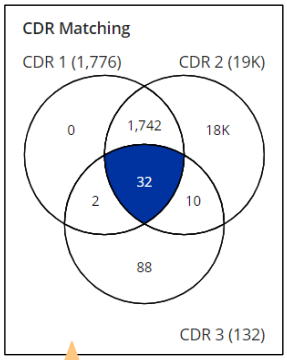


© 2022 化学情報協会

2. 回答の絞り込み 13

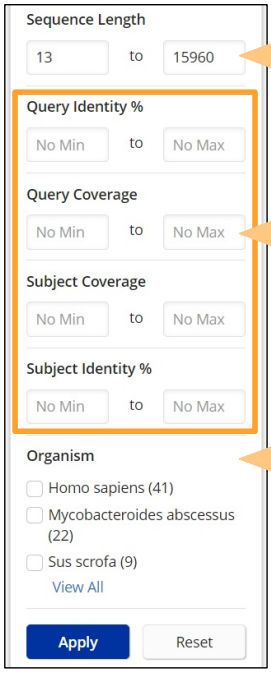
様々なフィルターによる絞り込みが可能

- CDR 配列検索



CDR セグメントの重なり

- BLAST/CDR/Motif 配列検索



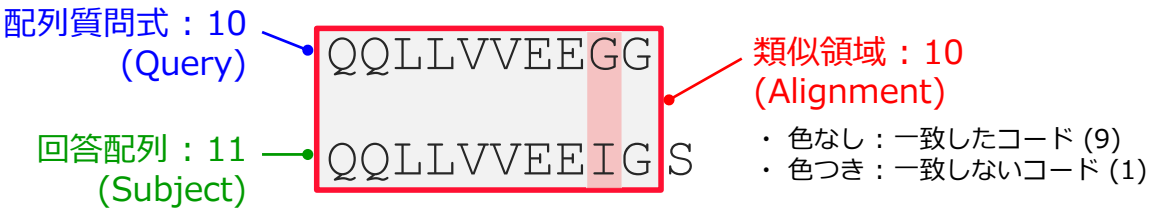
配列長 New

様々な計算値

生物種 (NCBI 由来の配列に限定) New

© 2022 化学情報協会

参考 - 計算値の説明



項目	内容	上記の例
Alignment Identity %*	一致したコード ÷ 類似領域 × 100	9 ÷ 10 × 100 = 90%
Query Identity %	一致したコード ÷ 配列質問式 × 100	9 ÷ 10 × 100 = 90%
Query Coverage %	類似領域 ÷ 配列質問式 × 100	10 ÷ 10 × 100 = 100%
Subject Identity %	一致したコード ÷ 回答配列 × 100	9 ÷ 11 × 100 = 81.82%
Subject Coverage %	類似領域 ÷ 回答配列 × 100	10 ÷ 11 × 100 = 90.91%

* Advanced Biosequences Search では Sequence Identity % と表される

© 2022 化学情報協会

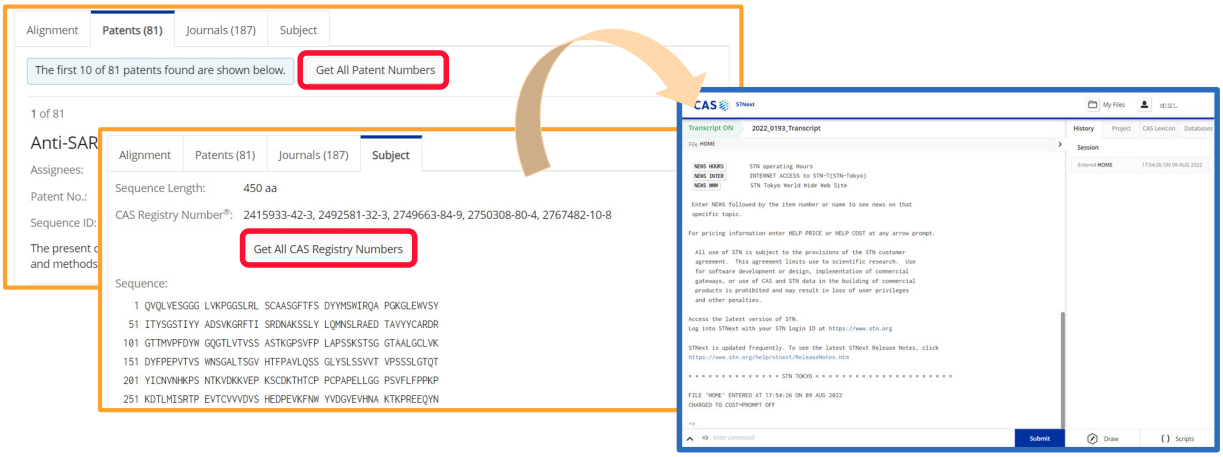
ヒットした配列の確認

Alignment タブで、各配列質問式とヒット配列の類似領域を確認できる

配列関連情報の検索 16

得られた配列や出典情報について CAS STNext のファイルで検索できる

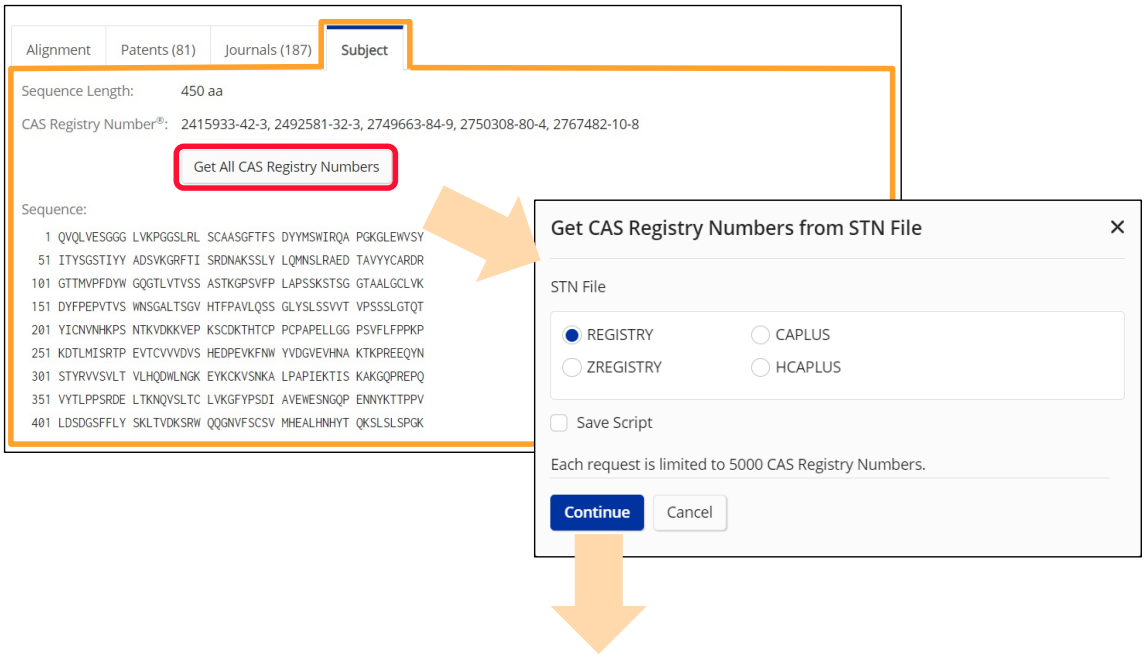
- 絞り込みや解析が可能
  - 配列情報：配列の名称、修飾の有無
  - 出典情報：キーワード、出願人・所属機関、発行年　など



© 2022 化学情報協会

3. ヒットした配列の名称や修飾の有無の確認 17

CAS RN® を持つ配列は、REGISTRY ファイルで配列の名称や修飾の有無について確認できる



© 2022 化学情報協会

3. ヒットした配列の名称や修飾の有無の確認 18

```
=> fil REGISTRY

=> QUE (2415933-42-3 OR 2492581-32-3 OR 2749663-84-9 OR 2750308-80-4 OR 2767482-10-8)
L1    QUE (2415933-42-3 OR 2492581-32-3 OR 2749663-84-9 OR 2750308-80-4 OR 27674
          82-10-8)
=> S L1
L2            5 (2415933-42-3 OR 2492581-32-3 OR 2749663-84-9 OR 2750308-80-4
                OR 2767482-10-8)
```

Script による自動検索

```
=> D 1-5 SQIDE
L2    ANSWER 1 OF 5  REGISTRY  COPYRIGHT 2022 ACS on STN
RN    2767482-10-8   REGISTRY
ED    Entered STN:  10 May 2022
CN    INDEX NAME NOT YET ASSIGNED
OTHER NAMES:
CN    108: PN: WO2022082217 SEQID: 121 claimed protein
FS    PROTEIN SEQUENCE
SQL   450

PATENT ANNOTATIONS (PNTE):
Sequence |Patent
Source   |Reference
=========+=============
Not Given|WO2022082217
         |claimed SEQID
         |121

SEQ      1 QVQLVESGGG LVKPGGSLRL SCAASGFTFS DYYMSWIRQA PGKGLEW
        51 ITYSGSTIYY ADSVKGRFTI SRDNAKSSLY LQMNSLRAED TAVYYCA
                :
```

配列の記載位置

```
                :
L2    ANSWER 5 OF 5  REGISTRY  COPYRIGHT 2022 ACS on STN
RN    2415933-42-3    REGISTRY
ED    Entered STN:  06 May 2020
CN    Immunoglobulin G1, anti-(severe acute respiratory syndrome coronavirus 2
      spike glycoprotein) (human monoclonal REGN10933 γ1-chain), disulfide
      with human monoclonal REGN10933 κ-chain, dimer  (CA INDEX NAME)
OTHER NAMES:
CN    Casirivimab
CN    REGN 10933
CN    RG 6413
FS    PROTEIN SEQUENCE
SQL   1328,450,450,214,214
NTE   multichain
      modified (modifications unspecified)
------------------------------------------------------------
 type            ------ location ------         description
------------------------------------------------------------------------
bridge           Cys-22      - Cys-96      disulfide bridge
bridge           Cys-147     - Cys-203     disulfide bridge
bridge           Cys-223     - Cys-214''   disulfide bridge
                :
```

配列の名称

修飾情報

© 2022 化学情報協会

4. ヒットした配列の特許情報の検索 19

すべての配列の特許情報は画面上部の
Get All Patent Numbers から検索する

© 2022 化学情報協会

4. ヒットした配列の特許情報の検索


```
=> fil CAPLUS

=> QUE (KR20130067314A OR US20150086563A1 OR JP5980384B2 OR KR1346530B1 OR JP2014140372A OR
SG10201400798VA OR CN102388068A OR IL232008A OR US8444981B2 OR CN103382222B OR JP5775615B2 OR
JP2015214563A OR EP2408816A2 OR KR1391472B1 OR KR20110128948A)/PN

L3    QUE (KR20130067314A OR US20150086563A1 OR JP5980384B2 OR KR1346530B1 OR JP
          2014140372A OR SG10201400798VA OR CN102388068A OR IL232008A OR US84449
          81B2 OR CN103382222B OR JP5775615B2 OR JP2015214563A
           KR1391472B1 OR KR20110128948A)/PN
```

作成された質問式

```
=> QUE (CA2754113A1 OR CA2754113C OR JP5498566B2 OR US20120177662A1 OR US9499620B2 OR
CN103382222A OR IL232008D0 OR CN102388068B OR US20170275365A1 OR US8871490B2 OR JP2012520679A
OR SG174344A1 OR WO2010-107752A3 OR US20120183561A1 OR US8454962B2 OR IL232009A OR
IL232009D0)/PN

L4    QUE (CA2754113A1 OR CA2754113C OR JP5498566B2 OR US20120177662A1 OR US9499
           620B2 OR CN103382222A OR IL232008D0 OR CN102388068B OR US20170275365A1
            OR US8871490B2 OR JP2012520679A OR SG174344A1 OR WO2010-107752A3 OR U
           S20120183561A1 OR US8454962B2 OR IL232009A OR IL232009D0)/PN
                  :
=> QUE (CN114395034A OR WO2022-087484A1 OR WO2022-094139A1 OR WO2022-090469A2 OR
US20220153815A1 OR WO2022-101839A1 OR CA3158642A1 OR CA3158742A1 OR WO2022-112392A1 OR WO2022-
119548A1 OR WO2022-133318A2 OR KR20220100011A OR CN114751988A OR CN114787928A)/PN

L11   QUE (CN114395034A OR WO2022-087484A1 OR WO2022-094139A1 OR
           OR US20220153815A1 OR WO2022-101839A1 OR CA3158642A1 O
          R WO2022-112392A1 OR WO2022-119548A1 OR WO2022-133318A
          11A OR CN114751988A OR CN114787928A)/PN
```

得られた配列の特許情報

```
=> S L3 OR L4 OR L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11
L12          67 L3 OR L4 OR L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11

=>
```

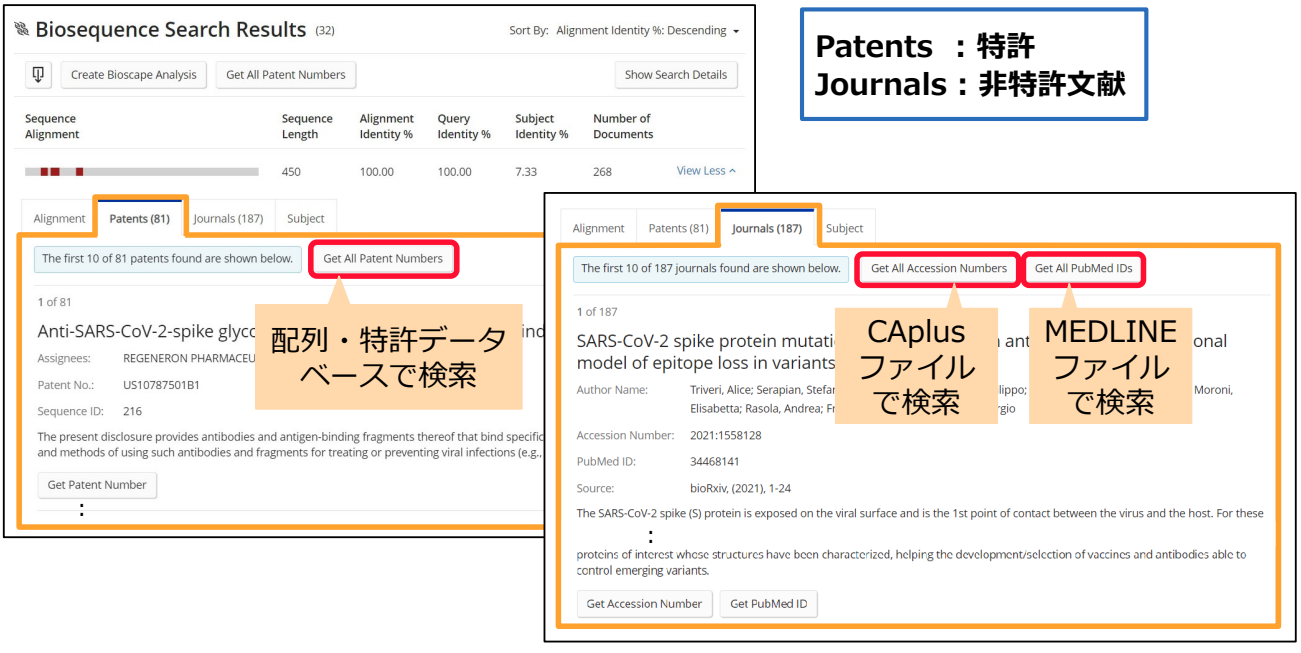
Script による自動検索

配列に紐づいた特許番号が検索される

© 2022 化学情報協会

参考：出典情報の検索

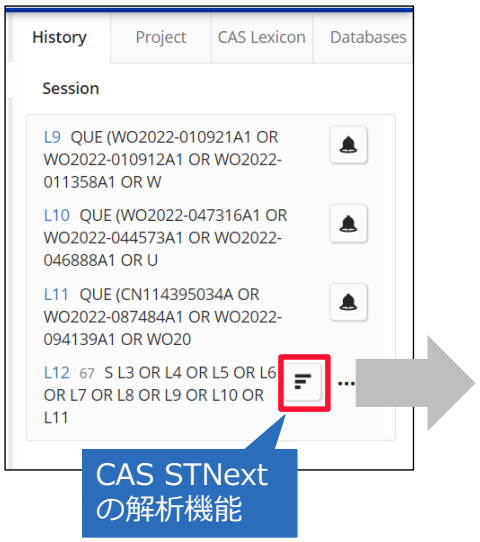
各配列の出典情報は Patents タブと Journals タブの Get All xxx で検索できる

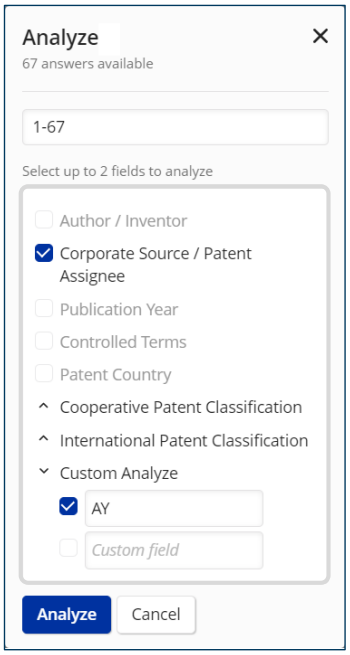


© 2022 化学情報協会

5. 特許出願人の出願動向の解析 22

得られた特許情報の特許出願人を出願年で解析する

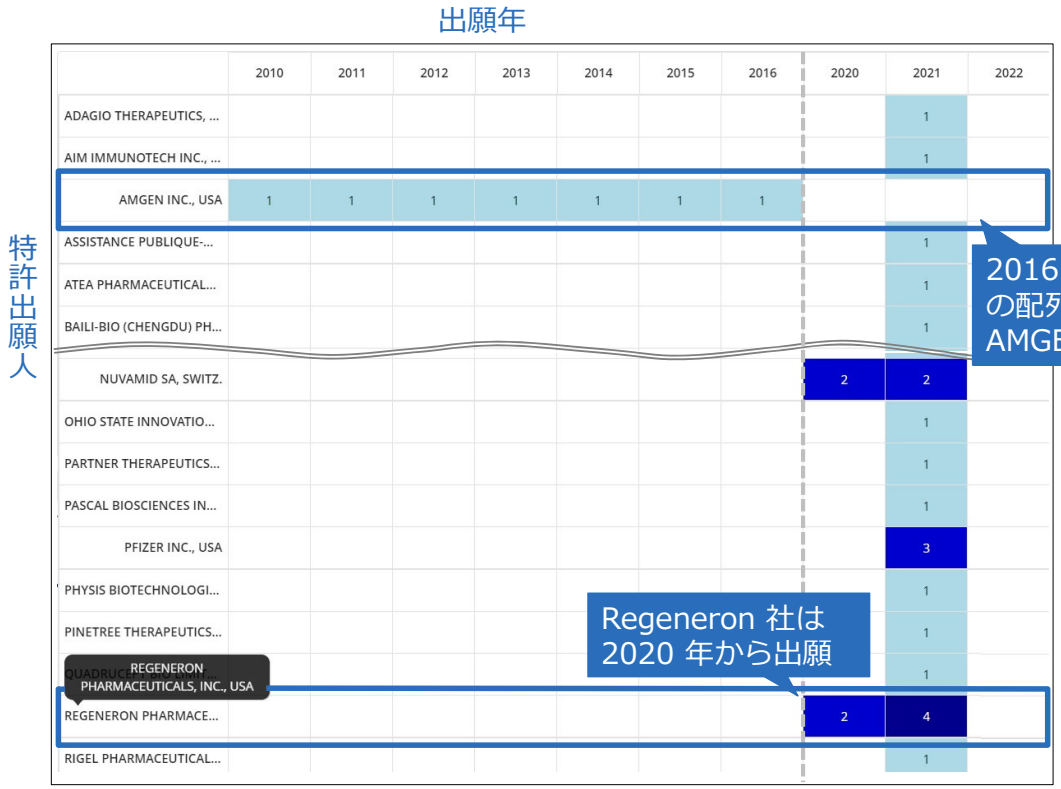




Corporate Source/ Patent Assignee と AY (出願年) を指定して相関を解析する

© 2022 化学情報協会

5. 特許出願人の出願動向の解析 23



© 2022 化学情報協会

まとめ

- 3 つのセグメントを持つ配列：32 件
  - CAS RN® を持たない配列：20 件

- 得られた特許情報：67 ファミリー (CAplus)
  - Casirivimab を開発した Regeneron 社よりも前 (2010～2016 年) に AMGEN 社が類似の配列を出願していた
  - 2020 年から Regeneron 社をはじめとする様々な企業・大学が出願している

© 2022 化学情報協会

最近の強化

強化内容のまとめ（2021 年 6 月以降） 26

強化内容	開始時期
NCBI 由来の配列の追加	2021.12
非特許文献情報の追加	2021.9
Organisms フィルターの追加	2022.2
★ Motif 配列検索の強化 - Combine Motif Results オプションの追加 - ヒットした配列質問式のパターン数の表示	 2021.12 2022.1
配列長フィルターの追加	2022.8

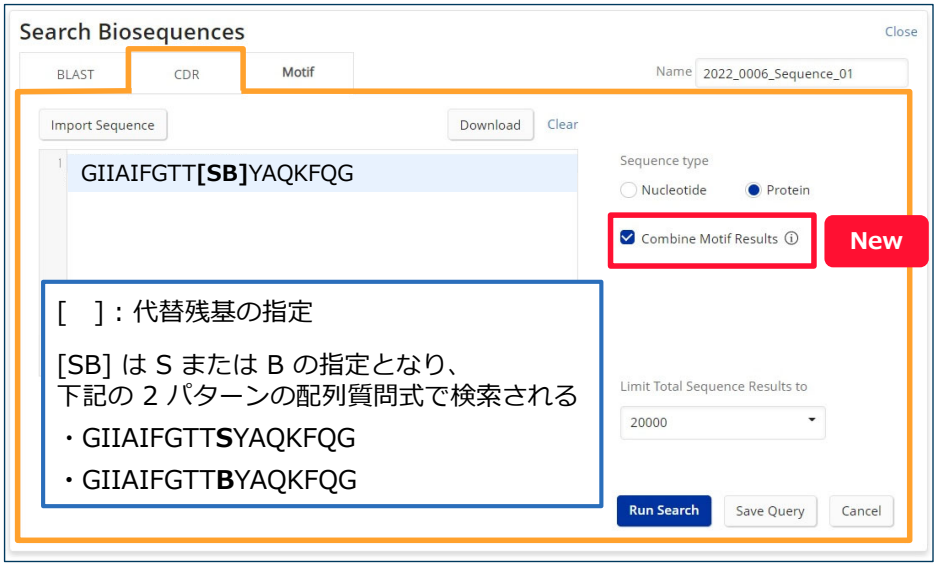
★の強化を紹介

© 2022 化学情報協会

最近の強化 – Motif 配列検索の強化 27

Combine Motif Results オプションの追加

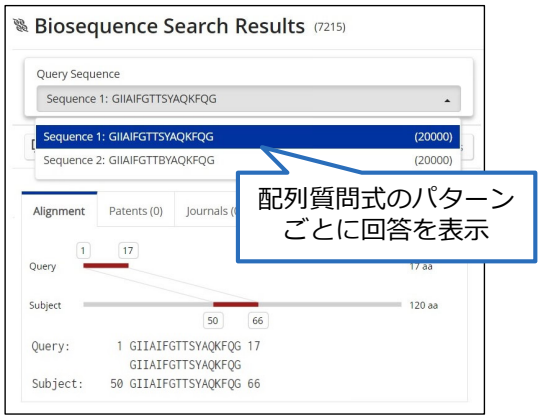
- 代替残基などの指定をした配列質問式を使用した際の回答配列の表示方法を指定する機能



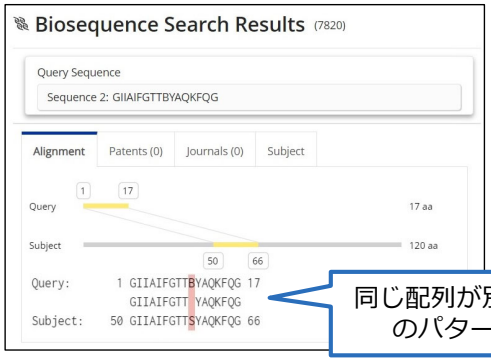
© 2022 化学情報協会

最近の強化 – Motif 配列検索の強化

Combine Motif Results：オフ

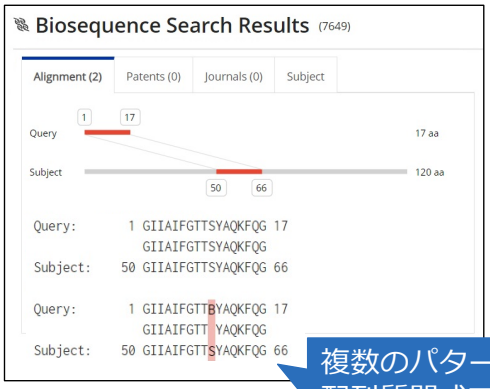


配列質問式のパターンごとに回答を表示



同じ配列が別の配列質問式のパターンでヒット

Combine Motif Results：**オン** New



複数のパターンの配列質問式でヒットした配列は一つにまとめて表示される

© 2022 化学情報協会

知財戦略＆特許調査に差がつく

STN IP Protection Suite セミナー

CAS FORMULATIONS™ ファイル

～ 製剤・配合の調査に最適なデータベース ～

CAS FORMULATIONS™ ファイル

～ 製剤・配合の調査に最適なデータベース ～

化学情報協会　情報事業部

目次

1. CAS FORMULATIONS ファイルの概要
2. 利用例 1 - 複数の成分を指定した検索
3. 利用例 2 – REGISTRY/CAplus を併用した検索
4. まとめ

© 2022 化学情報協会

CAS FORMULATIONS ファイルの概要

CAS FORMULATIONS ファイルとは 3

製剤・配合情報に特化したデータベース

- 製剤・配合に関するデータを人手で収集・整理
- 配合単位のレコード
- 配合中のすべての成分を収録
- 精密で柔軟な検索が可能
- REGISTRY や CAplus ファイルとの連携もスムーズ

© 2022 化学情報協会

ファイル概要 4

(2022 年 8 月)

ファイル名	シーエーエス フォーミュレーションズ CAS FORMULATIONS (CASFORM, CASFORMULTNS)
製作者	CAS
収録源	**特許** (CAplus ファイル由来) ：414 万レコード以上 **雑誌** (CAplus ファイル由来) ：40 万レコード以上 **添付文書*** ：19 万レコード以上
収録分野	重点分野：**医薬品、農薬、化粧品** 一部収録分野：コーティング剤、日用品、食品、材料
レコード構成	製剤・配合単位
収録期間	・英語の特許：1996 年～ ・日本語、中国語、韓国語、ドイツ語、フランス語の特許：2000 年～ (言語による) ・英語の雑誌：2014 年～

8 割以上が特許

* NLM の DailyMed 由来および EMA (欧州医薬品庁) の European public assessment reports (EPAR) 由来

© 2022 化学情報協会

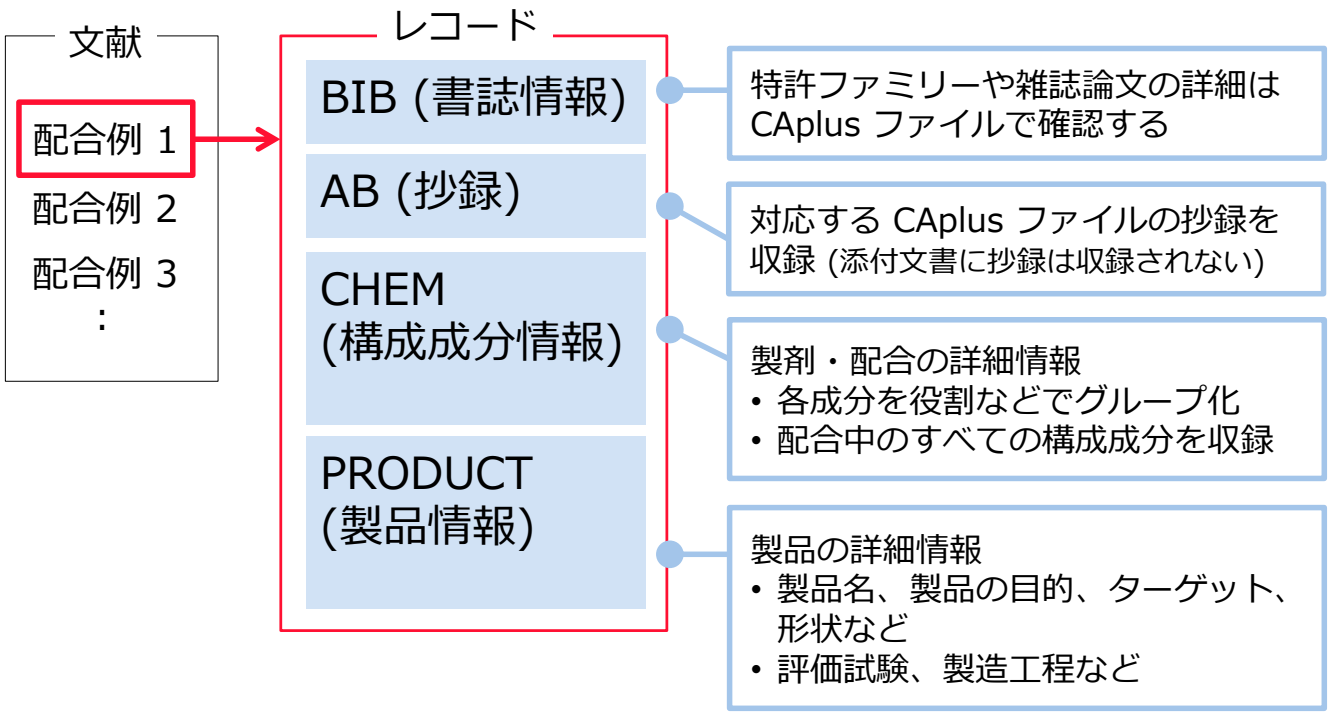
収録の強化 5

- 医薬品は、2019 年のファイルリリース時は経口医薬品がメインだったが、現在は下記も収録対象に
  - 皮膚外用薬、注射剤、点眼剤、点鼻剤
- 界面活性剤化学関連の収録を強化
  - 洗剤、掃除用製品、パーソナルケア製品など

© 2022 化学情報協会

レコード構成 6

1 レコード = 1 製剤・配合

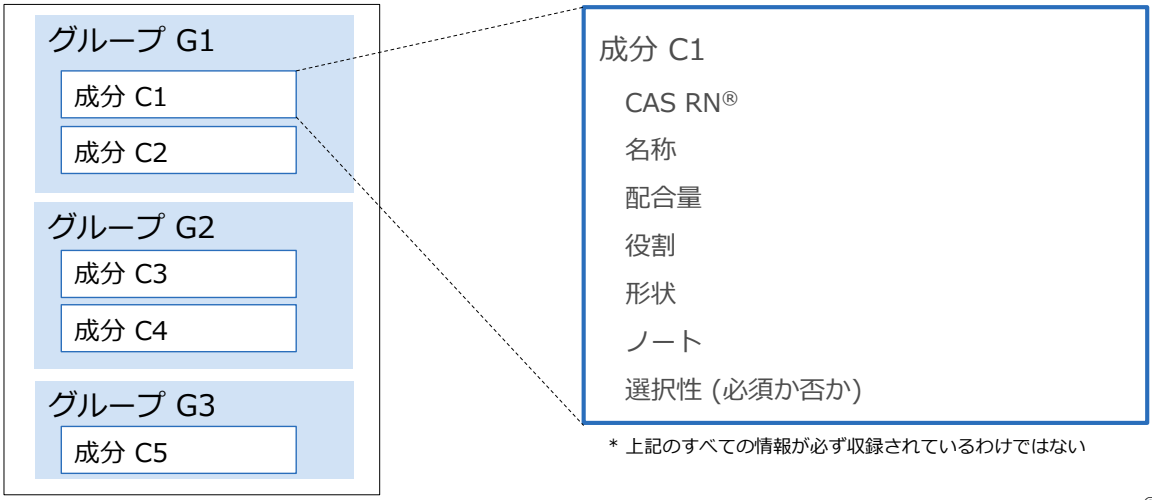


© 2022 化学情報協会

レコード構成 - CHEM (構成成分情報) 7

- 1 つの配合に含まれるすべての成分を収録
- 各成分 (Component) をグループ (Group) に分けて収録
- 名称、CAS RN®、役割、選択性などを収録

CHEM のイメージ



* 上記のすべての情報が必ず収録されているわけではない

© 2022 化学情報協会

レコード例 1 (CHEM 表示形式)

記載位置

```
AN   2019:246295  CASFORMULTNS
LCC  Claim
LCD  Claim 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
```

記載位置情報

グループ (G1) の情報

```
GROUP  G1
GDS  active ingredient
CMPS C1-C2
GOP  Mandatory
GAM  0.1-5 w/w %
```

成分 (C1) の情報

```
COMPONENT  C1
RN   130693-82-2
CNM  dorzolamide hydrochloride
CCN  4H-Thieno[2,3-b]thiopyran-2-sulfonamide,
     4-(ethylamino)-5,6-dihydro-6-methyl-, 7,7-dioxide, hydrochloride (1:1),
     (4S,6S)- (ACI)
CAM  0.1-5 w/w %
CFU  carbonic anhydrase inhibitors
COP  Mutually exclusive alternatives
```

成分 (C2) の情報

```
COMPONENT  C2
RN   120279-96-1
CNM  dorzolamide
CCN  4H-Thieno[2,3-b]thiopyran-2-sulfonamide,
     4-(ethylamino)-5,6-dihydro-6-methyl-, 7,7-dioxide, (4S,6S)- (9CI, ACI)
CAM  0.1-5 w/w %
CFU  carbonic anhydrase inhibitors
CNO  or its salt
COP  Mutually exclusive alternatives
```

:

成分ごとに情報が整理されている

JP 6132968

【請求項１】
ドルゾラミド 又はその塩、セルロース系高分子、及び ホウ酸 又はその塩を含有し、そしてｐＨが６．０以下である、医薬組成物。

【請求項２】
ドルゾラミド 又はその塩が ドルゾラミド塩酸塩 である、請求項１に記載の医薬組成物。

【請求項３】
ドルゾラミド又はその塩の含有量が０．１～５％（ｗ／ｗ）である、請求項１又は２に記載の医薬組成物。

:

© 2022 化学情報協会

レコード例 2 (CHEM 表示形式)

記載位置

```
AN   2022:15007  CASFORMULTNS
LCC  Claim
LCD  Claim 1, 2, 3, 4, 5, 6, 7
```

グループ (G1) の情報

```
GROUP  G1
GDS  Additional ingredients
CMPS C1
GOP  Mandatory
```

成分 (C1) の情報

```
COMPONENT  C1
RN   7631-86-9
CNM  fumed silica
CCN  Silica (9CI, ACI)
COP  Mandatory
```

グループ (G2) の情報

```
GROUP  G2
GDS  Additional ingredients
CMPS C2
GOP  Mandatory
```

成分 (C2) の情報

```
COMPONENT  C2
CNM  alkyl-modified carboxyvinyl polymer
COP  Mandatory
```

:

CAS RN® を持たない成分もある

JP 6995182

【請求項１】
次の成分（Ａ）～（Ｄ）；
（Ａ）煙霧シリカ
（Ｂ）アルキル変性カルボキシビニルポリマー
（Ｃ） 疎水化ヒドロキシプロピルメチルセルロース又はヒドロキシプロピルメチルセルロース
（Ｄ）油剤
を含有し、
前記成分（Ｃ）の含有量は０．０１重量%以上３．０重量%以下であり、
前記成分（Ｄ）が、成分（Ｄ１）油溶性紫外線吸収剤及び成分（Ｄ２）アルキル変性シリコーン及び／又はフェニル変性シリコーンを含有する、
ことを特徴とする水中油型乳化化粧料。

:

© 2022 化学情報協会

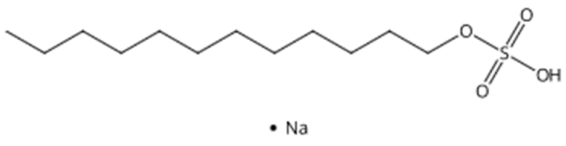
利用例

利用例 1

下記の両方が使われている配合を調べる

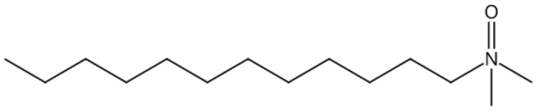
- 硫酸エステル塩系アニオン性界面活性剤 (成分 1)

  例：ラウリル硫酸ナトリウム (CAS RN®：151-21-3)



- アミンオキシド系両性界面活性剤 (成分 2)

  例：ラウリルジメチルアミンオキシド (CAS RN®：1643-20-5)



CAS FORMULATIONS ファイル利用の利点 12

- １ レコード = １ 製剤・配合単位なので
  簡単に “同一配合中” に限定できる
- “必須の成分である“ という限定ができる
- 物質と ”役割” (界面活性剤など) を紐づけた検索
  ができる

© 2022 化学情報協会

検索方針 13

- 成分 1、 2 について REGISTRY ファイルで構造検索し、
  CAS FORMULATIONS ファイルへクロスオーバーする
- 成分 1、 2 について CAS FORMULATIONS ファイルで
  部分名称からも検索する

さらに…

- 成分 1 と成分 2 がどちらも必須 (Mandatory) である配合に限定する
- 成分 1 と成分 2 について役割 (界面活性剤) で限定する
- 成分 1 と成分 2 を両方含む配合に限定する

© 2022 化学情報協会

検索方針 14

- 成分 1、 2 について REGISTRY ファイルで構造検索し、CAS FORMULATIONS ファイルへクロスオーバーする
- 成分 1、 2 について CAS FORMULATIONS ファイルで部分名称からも検索する

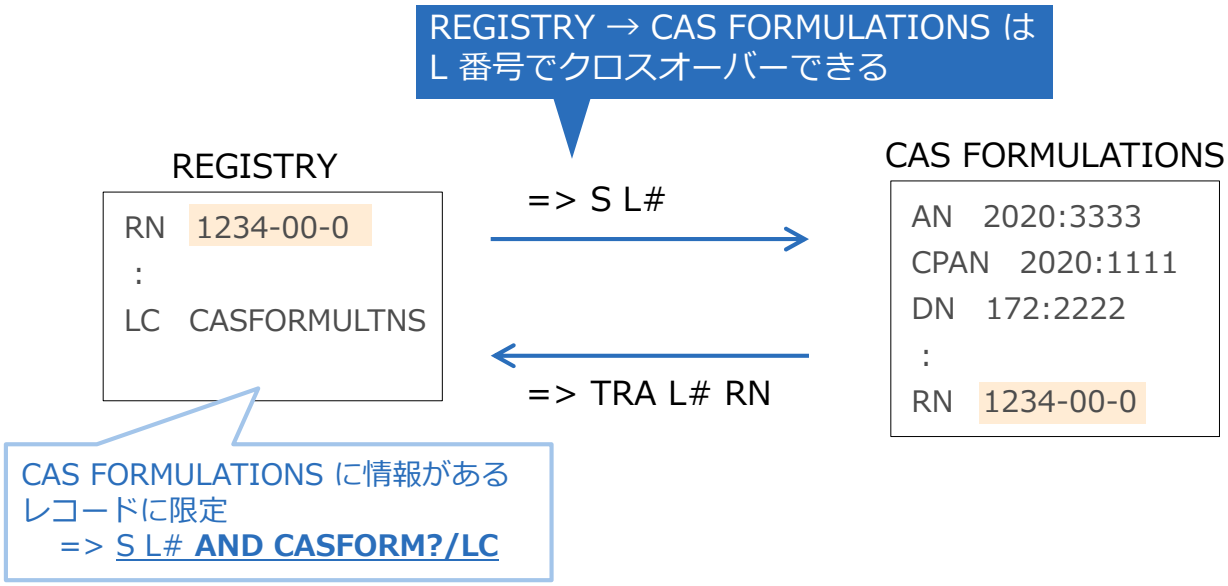
さらに…

- 成分 1 と成分 2 がどちらも必須 (Mandatory) である配合に限定する
- 成分 1 と成分 2 について役割 (界面活性剤) で限定する
- 成分 1 と成分 2 を両方含む配合に限定する

© 2022 化学情報協会

REGISTRY ファイルとの連携 15

- CAS FORMULATIONS ファイルと REGISTRY ファイルは CAS RN® で関連づけられている



© 2022 化学情報協会

CAS FORMULATIONS ファイルの物質名検索 16

- CAS FORMULATIONS ファイルには 2 種類の物質名関連フィールドがある

REGISTRY 由来の名称を最大 50 名称収録

  - CAS RN® に基づく物質名 (GCN, CCN フィールド)
    CAS RN® を持つ物質のみに存在

    => S BENZOIC ACID/CN (完全名、フレーズ) * 前方一致のみ可能
    => S ?HYALURON?/CNS (部分名、単語) * 前方・中間・後方一致可能

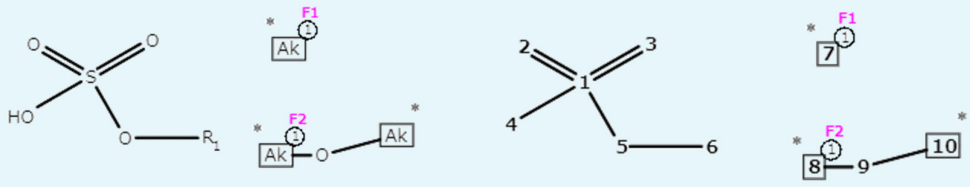
  - 原報の記載に基づく成分名 (CNM フィールド)
    CAS RN® を持たず CNM のみの成分もある

    => S BENZOIC ACID/CNM (完全名、フレーズ) * 前方一致のみ可能
    => S ?HYALURON? (P) CNM/FA (部分名、単語) * 前方・中間・後方一致可能

© 2022 化学情報協会

利用例 1 – 成分 1 の構造検索 17

```
=> FILE REGISTRY      ← REGISTRY ファイルに入る

=>
Uploading structure file: Component1      ← 構造質問式をアップロード
```



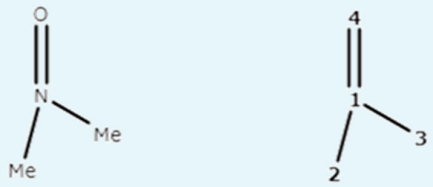

```
      :

L1       STRUCTURE UPLOADED

=> S L1                ← 構造検索（サンプル検索）
L2       50  SEA SSS SAM L1

=> S L1 FUL            ← 構造検索（フルファイル検索）
L3          4002 SEA SSS FUL L1

=> S L3 AND CASFORM?/LC
L4          237 L3 AND CASFORM?/LC
```

CAS FORMULATIONS ファイルに情報があるレコードに限定

© 2022 化学情報協会

利用例 1 – 成分 2 の構造検索


```
=>
Uploading structure file: Component2      ← 構造質問式をアップロード
```




```
L5      STRUCTURE UPLOADED

=> S L5              ← 構造検索（サンプル検索）
L6           50 SEA SSS SAM L5

=> S L5 FUL          ← 構造検索（フルファイル検索）
L7         4361 SEA SSS FUL L5

=> S L7 AND CASFORM?/LC          ← CAS FORMULATIONS ファイルに情報があるレコードに限定
L8          194 L7 AND CASFORM?/LC
```

© 2022 化学情報協会

利用例 1 – クロスオーバー検索、部分名検索

・成分 1

```
=> FILE CASFORM     ← CAS FORMULATIONS ファイルに入る

=> S L4
L9      204088 L4
```

REGISTRY から CAS FORMULATIONS へは
L 番号でクロスオーバー検索する

```
=> S ((SULFURIC ACID OR SULFATE) (P) ALKYL)/CNS OR ((SULFURIC ACID OR SULFATE) (P) ALKYL)(P)CNM/FA
L10      58941 ((SULFURIC ACID OR SULFATE) (P) ALKYL)/CNS OR ((SULFURIC ACID
               OR SULFATE) (P) ALKYL)(P)CNM/FA

=> S L9 OR L10      ← クロスオーバーした結果と
L11     259950 L9 OR L10  部分名検索の結果をまとめる
```

部分名は /CNS と 基本索引 (P) CNM/FA
を併用する

・成分 2

```
=> S L8              ← REGISTRY の回答をクロスオーバー
L12      13058 L8

=> S (?AMINE OXIDE?)/CNS OR (?AMINE OXIDE?) (P) CNM/FA      ← 部分名で検索
L13      27833 (?AMINE OXIDE?)/CNS OR (?AMINE OXIDE?) (P) CNM/FA

=> S L12 OR L13      ← クロスオーバーした結果と部分名検索の結果をまとめる
L14      28509 L12 OR L13
```

© 2022 化学情報協会

検索方針 20

- 成分 1、 2 について REGISTRY ファイルで構造検索し、CAS FORMULATIONS ファイルへクロスオーバーする
- 成分 1、 2 について CAS FORMULATIONS ファイルで部分名称からも検索する

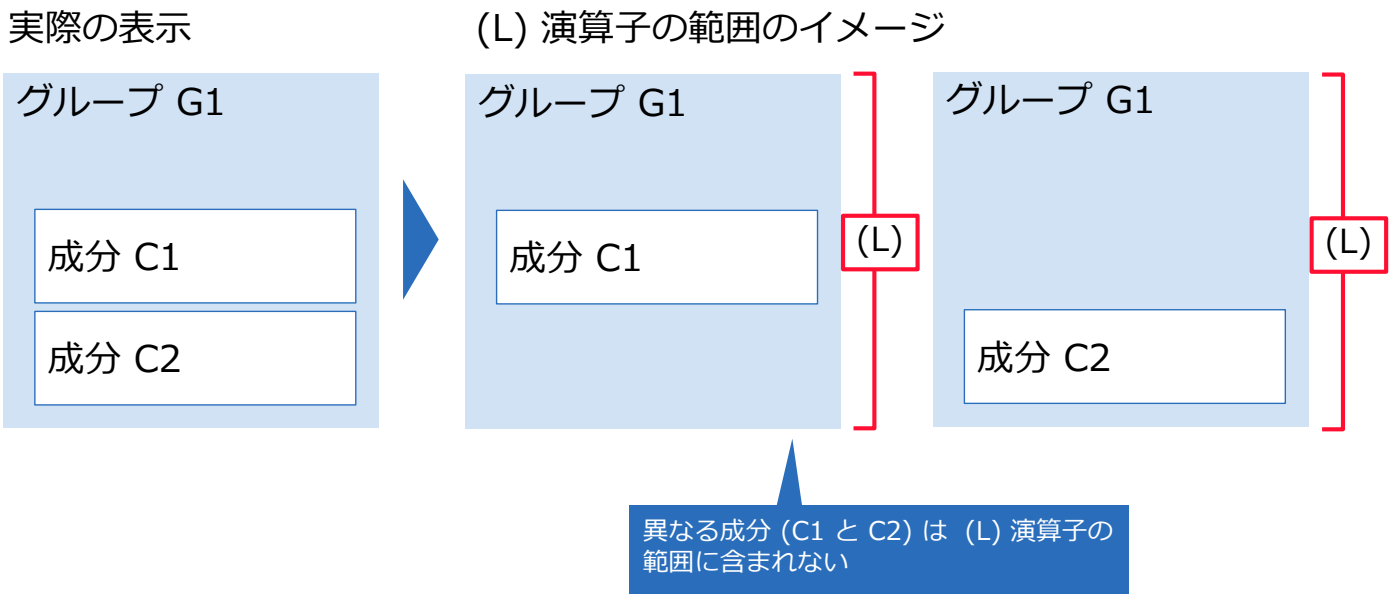
さらに…

- 成分 1 と成分 2 がどちらも必須 (Mandatory) である配合に限定する
- 成分 1 と成分 2 について役割 (界面活性剤) で限定する
- 成分 1 と成分 2 を両方含む配合に限定する

© 2022 化学情報協会

同一成分情報に限定した検索 21

- 同一成分情報およびその成分が属するグループ情報は (L) 演算子でかけ合わせる



© 2022 化学情報協会

選択性での限定 22

- 配合中の成分の選択性は /GOP および /COP フィールドで検索

GOP：グループ (Group) の選択性 (Optionality)
COP：成分 (Component) の選択性 (Optionality)

=> S L1 (L) MANDATORY/GOP,COP (必須成分に限定)
=> S L1 (L) EXCLUDE/GOP,COP (除外成分に限定)

- 選択性には下記の 9 種類がある

EXCLUDE	MANDATORY
MUTUALLY EXCLUSIVE ALTERNATIVES	NON-PREFERRED
ONE OR MORE	OPTIONAL
PREFERRED	SPECIFIED *COP のみ
UNDEFINED	

© 2022 化学情報協会

機能・役割に関するフィールド 23

- 機能・役割は GFU, CFU フィールドに収録

```
GROUP  G1
GDS  surfactants
GFU  surfactants
CMPS C1-C2
GOP  Mandatory
GAM  0.2 mass %
```

```
COMPONENT  C2
RN  151-21-3
CNM  sodium laurylsulfate
CCN  Sulfuric acid monododecyl ester sodium
salt (1:1) (ACI)
CFU  anionic surfactants
CNO  HLB value of 33, molecular weight of 289
COP  Mandatory
```

- さらに GVO, CVO フィールドに類義語を収録

```
GVO  Surfactants; Cleaning agents, surfactant;
Tensides; Surface-active agents; Surfactant;
Surface-active agent; Surface-active substance;
Tenside; Interface-active substance; Surface
active agent
```

```
CVO  Anionic surfactants; Interface-active substances,
anionic; Anionic surfactant; Anionic tensides; Tensides,
anionic; Anion surfactants
```

© 2022 化学情報協会

機能・役割の検索 24

- 機能・役割は /GFU,CFU フィールドで検索

=> S L1 (L) (SURFACTANT? OR ANIONIC SURFACTANT?)/GFU,CFU

SURFACTANT?/GFU,CFU で ANIONIC SURFACTANT はヒットしない！

- フレーズ単位の検索フィールド
- あらかじめ EXPAND で収録されているフレーズを確認

- さらに基本索引 (/BI) を加えると網羅的

=> S L1 (L) (SURFACTANT? OR ANIONIC SURFACTANT?)/GFU,CFU,BI

- 基本索引には GVO, CVO フィールドが含まれる
- 基本索引は単語単位で検索可能

SURFACTANT が含まれていれば ヒット！

© 2022 化学情報協会

利用例 1 – 選択性、役割での絞り込み 25

・成分 1

```
=> S L11 (L) MANDATORY/GOP,COP        ← 成分 1 が必須である配合に限定
L15      256645 L11 (L) MANDATORY/GOP,COP

=> S L15 (L) (ANIONIC SURFAC? OR SURFACTANT? OR SURFACE ACTIVE?)/CFU,GFU,BI
                                   ↑ 成分 1 の役割がアニオン性界面活性剤または界面活性剤である配合に限定
L16      78457 L15 (L) (ANIONIC SURFAC? OR SURFACTANT? OR SURFACE ACTIVE?)/CFU,GFU,BI
```

・成分 2

```
=> S L14 (L) MANDATORY/GOP,COP        ← 成分 2 が必須である配合に限定
L17      28017 L14 (L) MANDATORY/GOP,COP

=> S L17 (L) (AMPHOTERIC SURFAC? OR SURFACTANT? OR SURFACE ACTIVE?)/CFU,GFU,BI
                                   ↑ 成分 2 の役割が両性界面活性剤または界面活性剤である配合に限定
L18      10257 L17 (L) (AMPHOTERIC SURFAC? OR SURFACTANT? OR SURFACE ACTIVE?)/CFU,GFU,BI
```

・成分 1 と成分 2 を両方含む配合に限定

```
=> S L16 AND L18
L19       2747 L16 AND L18
```

1 レコード = 1 配合なので、AND 演算で同一配合中に限定できる

© 2022 化学情報協会

利用例 1 - 回答表示


```
=> D TI FD 1-40

L19  ANSWER 1 OF 2747  CASFORMULTNS  COPYRIGHT 2022 ACS on STN

TI   Liquid hand dishwashing cleaning composition
FD   Liquid Hand Dishwashing Cleaning Composition: Dishwashing Detergents ← FD（配合ディスクリプション）フィールドには
  :                                                                        配合の説明が含まれる
L19  ANSWER 38 OF 2747  CASFORMULTNS  COPYRIGHT 2022 ACS on STN

TI   Cleaning composition with excellent diffusibility and foaming property for
     hard surfaces
FD   Cleaning Composition for Hard Surfaces: Cleaning Agents
  :
```

TI FD 表示形式で回答をチェック

```
=> D HIT 38

L19  ANSWER 38 OF 2747  CASFORMULTNS  COPYRIGHT 2022 ACS on STN

AN   2021:78413  CASFORMULTNS
 :
GROUP  G1
GDS  Additional ingredients
CMPS C1
GOP  Mandatory

COMPONENT  C1
RN   151-21-3
CNM  sodium lauryl sulfate
CCN  Sulfuric acid monododecyl ester sodium salt (1:1) (ACI)
CAM  4.5 parts by mass
CFU  anionic surfactants
COP  Mandatory
```

HIT 表示形式でヒット部分を確認

（続き）

```
AN   2021:78413  CASFORMULTNS
 :

GROUP  G2
GDS  Additional ingredients
CMPS C2
GOP  Mandatory

COMPONENT  C2
CNM  lauryldimethylamine oxide (AO)
CAM  10.5 parts by mass
CFU  amphoteric surfactants
COP  Mandatory
```

協会

利用例 1 - 回答表示


```
=> D BIB CHEM 38

L19  ANSWER 38 OF 2747  CASFORMULTNS  COPYRIGHT 2022 ACS

AN   2021:78413  CASFORMULTNS
CPAN 2021:2383867
DN   176:473618
ED   26 Jun 2022
UP   26 Jun 2022
TI   Cleaning composition with excellent diffusibility and foaming property for
     hard surfaces
AU   Tamura, Nozomu; Handa, Hiroyuki; Maehara, Kunikazu; Egashira, Masatoshi
JT   Jpn. Kokai Tokkyo Koho
PA   Earth Chemical Co.
DT   Patent

PI
     PATENT NO.          KIND  DATE        APPLICATION NO.          DATE
     ---------------     ----  --------    ---------------------    --------
     JP 2021175786        A    20211104    JP 2020-194264           20201124

FN   Form 21
FD   Cleaning Composition for Hard Surfaces: Cleaning Agents
GRPS G1-G6
GRPC 6
CMPC 6
TGTC 1
```

BIB CHEM 表示形式で書誌情報と構成成分情報を確認

© 2022 化学情報協会

利用例 1 - 回答表示


```
AN   2021:78413  CASFORMULTNS
LCC  Example
LCD  Table 6, Example 13

GROUP  G1
GDS  Additional ingredients
CMPS C1
GOP  Mandatory

COMPONENT  C1  ①
RN   151-21-3
CNM  sodium lauryl sulfate
CCN  Sulfuric acid monododecyl ester sodium salt (1:1) (A
CAM  4.5 parts by mass
CFU  anionic surfactants
COP  Mandatory

GROUP  G2
GDS  Additional ingredients
CMPS C2
GOP  Mandatory

COMPONENT  C2  ②
RN   1643-20-5
CNM  lauryldimethylamine oxide (AO)
CCN  1-Dodecanamine, N,N-dimethyl-, N-oxide (9CI, ACI)
CAM  10.5 parts by mass
CFU  amphoteric surfactants
COP  Mandatory
```

成分 1

成分 2

表6

成分名		実施例 1 2	実施例 1 3	
アニオン性界面活性剤	POE(3)ラウリル硫酸エーテル硫酸Na（AES）	5.4		
	ラウリル硫酸ナトリウム		4.5	①
両性界面活性剤	ラウリルジメチルアミンオキサイド ※1	14.4		
	ラウリルジメチルアミンオキサイド（AO）		10.5	②
	ミリスチルプロピルベタイン			
ノニオン性界面活性剤	POE(4)ラウリルエーテル		20	
	ビス(3-アミノプロピル)ドデシルアミン			
	ヤシ油脂肪酸N-メチルエタノールアミド			
	POE(30)ラウリルエーテル			
アルコール	3-メトキシ-3-メチル-1-ブタノール	15	10	
	ジエチレングリコールモノブチルエーテル			
	ラウリルアルコール			
	エタノール			
	ジデシルジメチルアンモニウムメトサルフェート	1.6		
	塩化ベンザルコニウム		2.5	
	プロピレングリコールモノラウレート			
	水	残部	残部	
合計		100	100	

硫酸エステル塩系界面活性剤 (成分 1) とアミンオキシド系界面活性剤 (成分 2) が両方含まれる配合が得られた

© 2022 化学情報協会

利用例 1 - 回答表示


```
GROUP  G3
GDS  Additional ingredients
CMPS C3
GOP  Mandatory

COMPONENT  C3  ③
CNM  POE (4) lauryl ether
CAM  20 parts by mass
CFU  nonionic surfactants
COP  Mandatory

GROUP  G4
GDS  Additional ingredients
CMPS C4
GOP  Mandatory

COMPONENT  C4  ④
RN   56539-66-3
CNM  3-methoxy-3-methyl-1-butanol
CCN  1-Butanol, 3-methoxy-3-methyl- (6CI, 9CI, ACI)
CAM  10 parts by mass
COP  Mandatory
```

表6

成分名		実施例 1 2	実施例 1 3	
アニオン性界面活性剤	POE(3)ラウリル硫酸エーテル硫酸Na（AES）	5.4		
	ラウリル硫酸ナトリウム		4.5	
両性界面活性剤	ラウリルジメチルアミンオキサイド ※1	14.4		
	ラウリルジメチルアミンオキサイド（AO）		10.5	
	ミリスチルプロピルベタイン			
ノニオン性界面活性剤	POE(4)ラウリルエーテル		20	③
	ビス(3-アミノプロピル)ドデシルアミン			
	ヤシ油脂肪酸N-メチルエタノールアミド			
	POE(30)ラウリルエーテル			
アルコール	3-メトキシ-3-メチル-1-ブタノール	15	10	④
	ジエチレングリコールモノブチルエーテル			
	ラウリルアルコール			
	エタノール			
	ジデシルジメチルアンモニウムメトサルフェート	1.6		
	塩化ベンザルコニウム		2.5	
	プロピレングリコールモノラウレート			
	水	残部	残部	
合計		100	100	

© 2022 化学情報協会

利用例 1 - 回答表示


```
GROUP  G5
GDS  Additional ingredients
CMPS C5
GOP  Mandatory

COMPONENT  C5                ⑤
RN   8001-54-5
CNM  benzalkonium chloride
CCN  Quaternary ammonium compounds, alkylbenzyldimethyl,
CAM  2.5 parts by mass
COP  Mandatory

GROUP  G6
GDS  Additional ingredients
CMPS C6
GOP  Mandatory

COMPONENT  C6                ⑥
RN   7732-18-5
CNM  water
CCN  Water (8CI, 9CI, ACI)
COP  Mandatory
```

表6

成分名		実施例 12	実施例 13	
アニオン性界面活性剤	POE(3)ラウリル硫酸エーテル硫酸Na（AES）	5.4		
	ラウリル硫酸ナトリウム		4.5	
両性界面活性剤	ラウリルジメチルアミンオキサイド ※1	14.4		
	ラウリルジメチルアミンオキサイド（AO）		10.5	
	ミリスチルプロピルベタイン			
ノニオン性界面活性剤	POE(4)ラウリルエーテル		20	
	ビス(3-アミノプロピル)ドデシルアミン			
	ヤシ油脂肪酸N-メチルエタノールアミド			
	POE(30)ラウリルエーテル			
アルコール	3-メトキシ-3-メチル-1-ブタノール	15	10	
	ジエチレングリコールモノブチルエーテル			
	ラウリルアルコール			
	エタノール			
	ジデシルジメチルアンモニウムメトサルフェート	1.6		
	塩化ベンザルコニウム		2.5	⑤
	プロピレングリコールモノラウレート			
	水	残部	残部	⑥
合計		100	100	

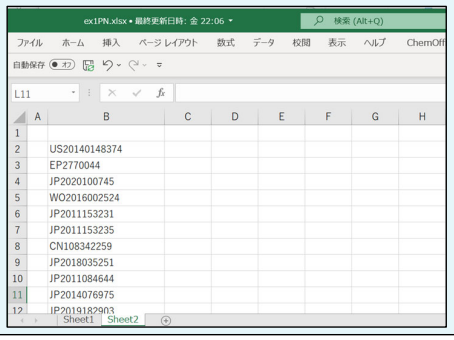
© 2022 化学情報協会

参考：特許全文情報を一括で確認したい場合

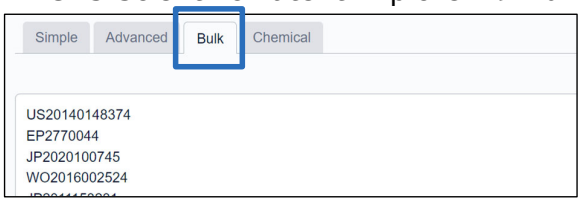
CAS FORMULATIONS の回答から特許番号を抽出して CAS Scientific Patent Explorer で検索することで、特許の全文情報を確認できる

1. CAS FORMULATIONS ファイルでヒットした特許の特許番号 (PN) を ANALYZE コマンドで抽出

```
=> ANA L19 PN
L20          ANALYZE L19 1- PN :      761 TERMS

=> D 1-

L20          ANALYZE L19 1- PN :      761 TERMS

TERM #    # OCC   # DOC   % DOC PN
------ -------- ------ ------ ----------------
     1      108    108   4.04 US20140148374
     2       72     72   2.69 EP2770044
     3       45     45   1.68 JP2020100745
     4       44     44   1.64 WO2016002524
     :
```



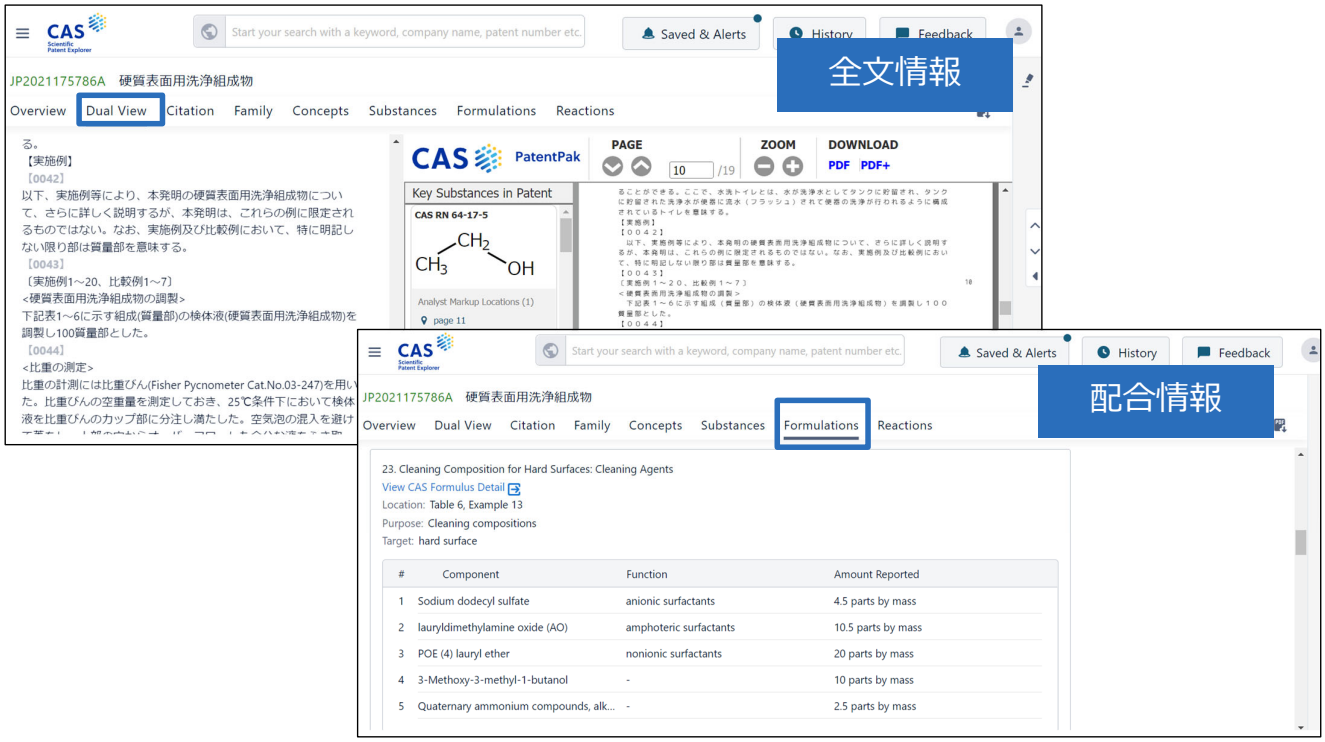
2. CAS Scientific Patent Explorer の Bulk Patent Search*で検索



＊ CAS Scientific Patent Explorer の章参照

© 2022 化学情報協会

参考：特許全文情報を一括で確認したい場合　(続き)　32



© 2022 化学情報協会

利用例 2

L-アルギニン (CAS RN® : 74-79-3) が含まれる注射用製剤に関する日本特許を、REGISTRY、CAplus、CAS FORMULATIONS ファイルを使って網羅的に調べる

CAS FORMULATIONS ファイル利用の利点 34

- 配合中のすべての成分の情報が収録されており、CAplus ではヒットしない回答が得られる可能性がある
- 投与経路や用途に関する情報が収録されている

© 2022 化学情報協会

検索方針 35

REGISTRY で L-アルギニンを検索

CAplus にクロスオーバー

キーワードで注射用製剤に限定

CAS FORMULATIONS にクロスオーバー

キーワードで注射用製剤に限定

CAplus にクロスオーバー

CAplus で回答をまとめる

日本特許に限定

© 2022 化学情報協会

利用例 2 – 検索の流れ

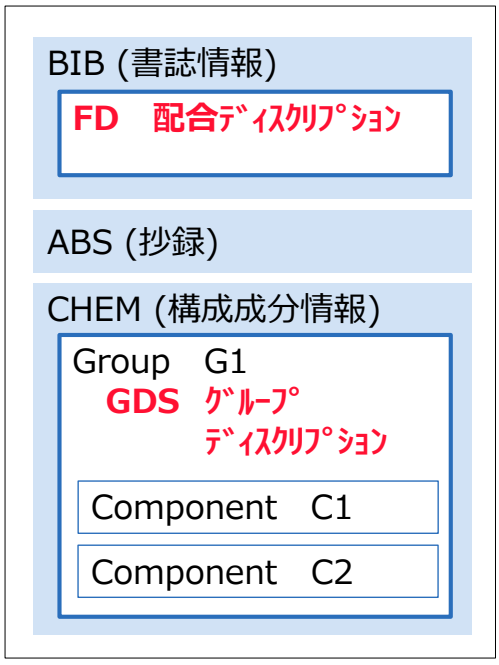

```
=> FILE REGISTRY      ← REGISTRY ファイルに入る

=> S 74-79-3          ← L-アルギニンの CAS RN® で検索
L1             1 74-79-3

=> FILE CAPLUS        ← CAplus ファイルに入る

=> S L1               ← REGISTRY ファイルの回答をクロスオーバー
L2       120280 L1

=> S L2 AND (INJECT? OR INFUS?)          ← 「注射」に関するキーワードで限定
L3         7774 L2 AND (INJECT? OR INFUS?)
```

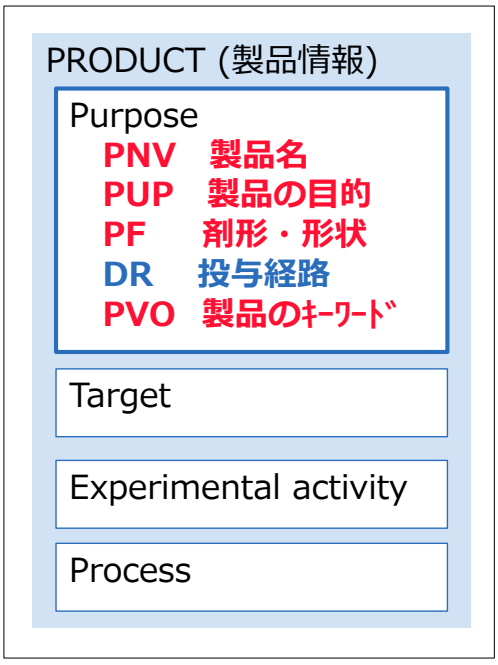
CAplus ファイルで L-アルギニンが含まれる注射用製剤に関する文献を検索

© 2022 化学情報協会

投与経路、用途に関するフィールド

- 投与経路の情報は**投与経路 (DR)** フィールドに含まれる
- **用途に関する情報**は複数のフィールドに含まれる





© 2022 化学情報協会

投与経路、用途の検索 38

- 下記の検索フィールドで検索

	内容	検索フィールド	検索単位
投与経路	投与経路	/DR	単語
用途	配合ディスクリプション*1	/FD	単語
	グループディスクリプション	/GDS	単語・フレーズ
	製品：目的	/PUP	フレーズ
	製品の剤形・形状	/PF	フレーズ
	製品名，概要	/PNV	単語・フレーズ
	製品：キーワード*1	/PVO	フレーズ

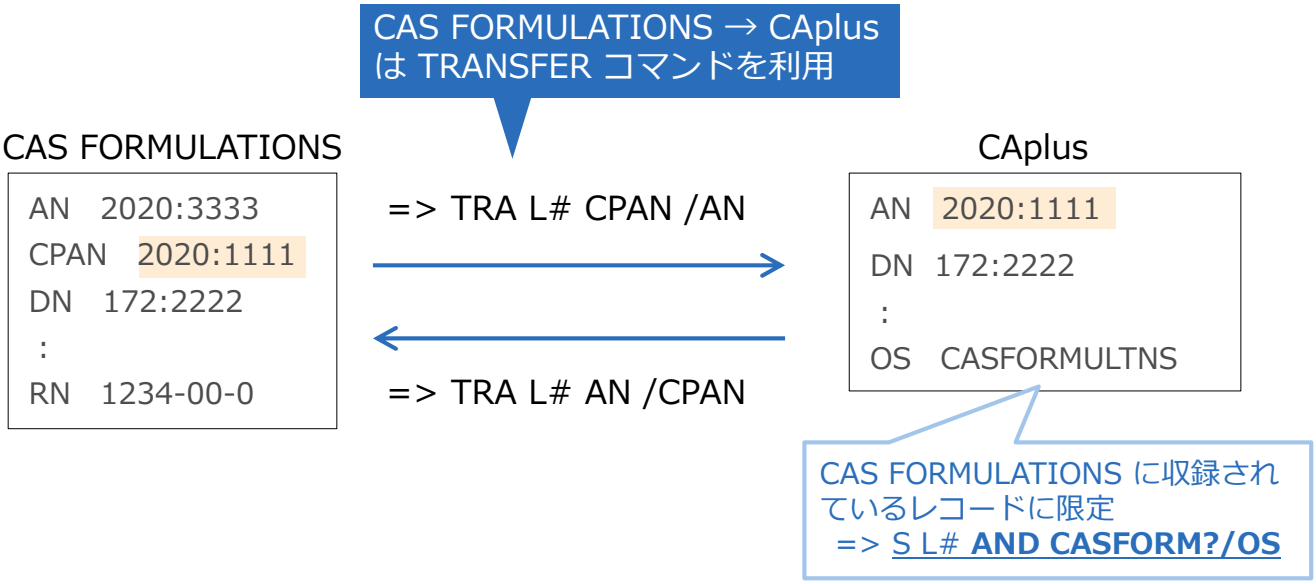
*1　基本索引に含まれるフィールド

- 基本索引では検索されないフィールドが多いため、検索フィールドを指定する
- フレーズ単位のフィールドはあらかじめ EXPAND で収録されているフレーズを確認する

© 2022 化学情報協会

CAplus ファイルとの連携 39

CAS FORMULATIONS ファイルと CAplus ファイルは レコード番号で関連づけられている



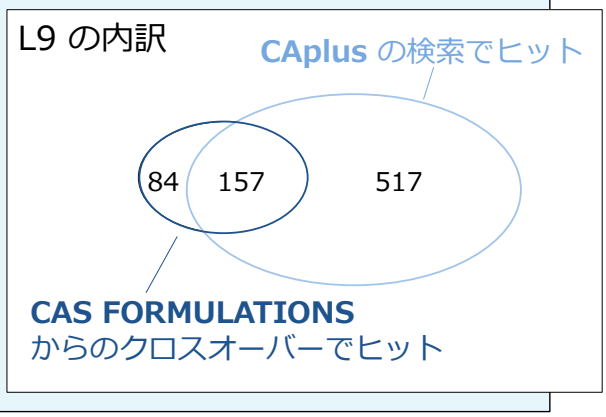
© 2022 化学情報協会

利用例 2 - 検索の流れ


```
=> FILE CASFORM          ← CAS FORMULATIONS ファイルに入る

=> S L1                  ← REGISTRY ファイルの回答をクロスオーバー
L4       30438 L1

=> S L4 AND (INJECT? OR INFUS?)/DR,FD,GDS,PUP,PF,PNV,PVO   ← 投与経路、用途関連のフィールドで
L5        2459 L4 AND (INJECT? OR INFUS?)/DR,FD,GDS,PUP,PF,PNV,PVO   「注射」に関するキーワードで限定

=> FILE CAPLUS           ← CAplus ファイルに入る

=> TRA L5 CPAN /AN
L6          TRANSFER L5 1- CPAN :     690 TERMS
L7         690 L6/AN
ALL TERMS IN L6/AN RETRIEVED.

=> S L3 OR L7            ← CAplus の検索結果と OR でまとめる
L8        7842 L3 OR L7

=> S L8 AND JP/PC        ← 日本特許に限定
L9         758 L8 AND JP/PC
```

CAS FORMULATIONS から CAplus へは TRANSFER コマンドでクロスオーバー

CAS FORMULATIONS には対応特許情報は含まれないため、特許発行国で絞り込む場合は CAplus にクロスオーバーしてから検索する



© 2022 化学情報協会

CAS FORMULATIONS のみでヒットした回答の例


```
AN   2003:150891  CASFORMULTNS
CPAN 2003:1006769
  :
TI   Medicament combinations of sodium channel blockers and fibrinolytics for
     treating ischemic conditions
AU   Banzet, Sophie; Duettmann, Hermann; Mauz, Annerose
JT   PCT Int. Appl.
PA   Boehringer Ingelheim Pharma G.m.b.H. & Co. K.-G.
DT   Patent
PI
     PATENT NO.          KIND  DATE        APPLICATION NO.          DATE
     ----------------    ----  --------    ---------------------    --------
     WO 2003105844        A1   20031224    WO 2003-EP5813           20030604

FN   Form 19
FD   Anti-Ischemic Infusion  Solution: Anti-Ischemic Agents   ← 配合ディスクリプション
  :
LCC  Example
LCD  Example

GROUP  G1
  :
COMPONENT  C1
RN   191588-94-0
CNM  tenecteplase   ①
  :
```

対応特許 (JP2005536478) で確認

【００２７】
D.3.1)成分２の医薬製剤実施例
成分２の製剤は従来技術に知られており、商業上得られてもよい。本発明に従って使用し得る幾つかの市販の製剤が例として、また説明として以下に挙げられる。
アルテプラーゼ：（注射／注入用の溶液を調製するための粉末及び溶媒）
組成：　アルテプラーゼ10mg/20mg/50mg/100mg
その他の成分：アルギニン、リン酸、ポリソルベート80
注射用の水
テネクテプラーゼ：（注射溶液を調製するための粉末及び溶媒）
組成：　テネクテプラーゼ8000U/10000U(40mg/50mg) ← ①
その他の成分：アルギニン、リン酸、ポリソルベート20
注射用の水8ml/10ml

© 2022 化学情報協会

CAS FORMULATIONS のみでヒットした回答の例

42

```
GROUP  G2
  :
COMPONENT  C2
RN   74-79-3          ②
CNM  arginine
  :
COMPONENT  C3
RN   7664-38-2        ③
CNM  phosphoric acid
  :
COMPONENT  C4
RN   9005-64-5        ④
CNM  polysorbate 20
  :
COMPONENT  C5
RN   7732-18-5        ⑤
CNM  water for injection
  :
PNC  author
PNV  Anti-ischemic infusion  solution   ← 製品名

PRODUCT  PA1
PUP  Anti-ischemic agents
PF   Desc: pharmaceutical solutions
DR   infusion drug delivery systems     ← 投与経路
  :
```

対応特許 (JP2005536478) で確認

【００２７】
D.3.1)成分２の医薬製剤実施例
成分２の製剤は従来技術に知られており、商業上得られてもよい。本発明に従って使用し得る幾つかの市販の製剤が例として、また説明として以下に挙げられる。
アルテプラーゼ：（注射／注入用の溶液を調製するための粉末及び溶媒）
組成：　アルテプラーゼ10mg/20mg/50mg/100mg
その他の成分：アルギニン、リン酸、ポリソルベート80
注射用の水
テネクテプラーゼ：（注射溶液を調製するための粉末及び溶媒）
組成：　テネクテプラーゼ8000U/10000U(40mg/50mg)
その他の成分：アルギニン、リン酸、ポリソルベート20
注射用の水8ml/10ml

②　③　④　⑤

実施例中の製剤例にアルギニンが記載されている特許がヒット

CAplus ファイルは重要な化学物質のみが索引されるためアルギニンは索引されていなかった
(クレームにはアルギニンについての言及なし)

© 2022 化学情報協会

まとめ

43

CAS FORMULATIONS ファイルは

- 製剤・配合単位なので、同一配合中に限定した検索ができる
- 役割や用途などの情報で絞り込みができる
- REGISTRY や CAplus ファイルと組み合わせて検索することでより便利に利用できる

© 2022 化学情報協会

CAS STNext ユーザーミーティング

CAS FILES の強化

CAS FILES の強化

化学情報協会　情報事業部

目次

1. CAplus/CA ファイルの強化
2. MARPAT ファイルの強化
3. CHEMLIST ファイルの強化

© 2022 化学情報協会

CAplus/CA ファイルの強化

CAplus/CA ファイルの強化 3

- 特許請求項の収録拡大
- Prior Art Analysis 機能の搭載
- CAS PatentPak 関連の強化
- 特許ステータス関連の強化

© 2022 化学情報協会

CAplus ファイルの概要 4

世界中の科学技術分野の学術論文、単行本、61 特許発行機関の特許、2 技術公開誌を収録する文献データベース

(2022 年 8 月)

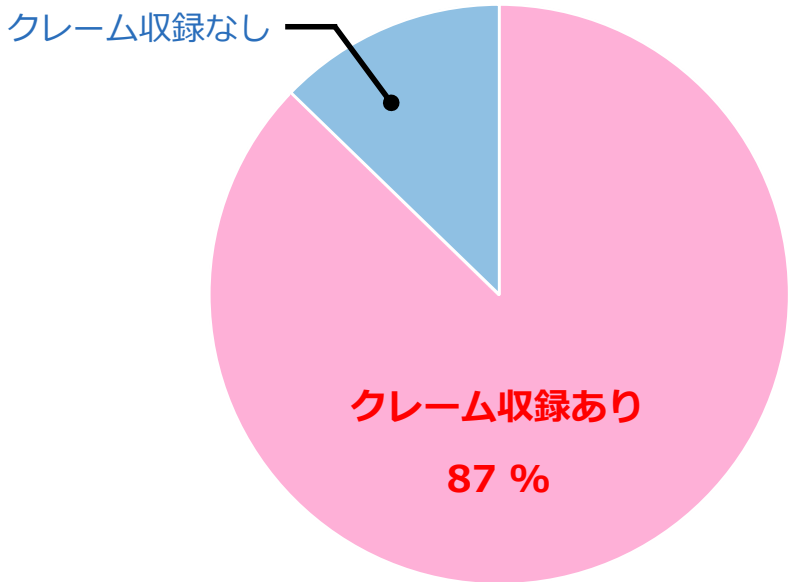
製作者	CAS
レコード構成	文献単位，特許は発明単位
収録期間	1808 年～
収録件数	5,800 万件以上
特長	• 統制語による索引が付与されており、特に化学物質に関する文献検索が的確かつ容易に実行できる • 特許レコードは同一発明単位 (ファミリー単位) で構成されており、対応特許情報が容易に得られる • CAS PatentPak で物質情報付き明細書を表示できる

© 2022 化学情報協会

特許請求項の収録拡大 5

特許請求項の収録対象が 13 の特許発行機関に拡大

- CAplus ファイルの特許レコードの 87 % に収録 (2022 年 8 月)



© 2022 化学情報協会

特許請求項の収録状況 6

ベーシック特許発行国が下記の場合に収録

収録期間拡大

New

ベーシック特許発行国	収録期間	レコード数 (CAplus)
PCT 出願 (WO)	1979 -	176 万件 (98.9%)
米国 (US)	1906 -	184 万件 (84.1%)
中国 (CN)	1985 -	653 万件 (99.5%)
イギリス (GB)	1927 -	7.6 万件 (35.1%)*
ドイツ (DE)	1997 -	37 万件 (99.2%)
ロシア (RU)	1994 -	28 万件 (99.4%)
ヨーロッパ (EP)	1979 -	44 万件 (99.8%)
韓国 (KR)	1995 -	100 万件 (98.9%)
オーストラリア(AU)	2000 -	2 万件 (96.9%)
ブラジル(BR)	2000 -	2 万件 (89.0%)
スイス(CH)	1975 -	0.5 万件 (93.4%)
インド(IN)	2007 -	9.8 万件 (96.2%)
日本 (JP)	1983 -	282 万件 (89.6%)

* 1970 年以降は 7.5 万件 (89.1%)

© 2022 化学情報協会

特許請求項の検索方法 7

- 請求項を検索

=> **S 検索語/CLM**

=> **S 検索語/BIEX**

CAplus/CA ファイルでは、請求項は /CLM または /BIEX で検索できる
(どちらも請求項のみが検索対象)

- 基本索引および請求項を検索

=> **S 検索語/BI,CLM**

=> **S 検索語/BI,BIEX**

© 2022 化学情報協会

特許請求項の検索 - ALL CLM 表示 8

検索例：豚糞を利用した土壌改良資材に関する特許の検索

```
=> FILE CAPLUS

=> S ((SOIL OR EDAPHIC) (10A) (IMPROV? OR AMEND?)(L)(PIG OR PIGLET OR HOG))/BI,CLM
L1        2093 ((SOIL OR EDAPHIC) (10A) (IMPROV? OR AMEND?)(L)(PIG OR PIGLET OR HOG))/BI,CLM

=> D 20 ALL CLM

L1  ANSWER 20 OF 2093  CAPLUS  COPYRIGHT 2022 ACS on STN
AN   2022:1509508  CAPLUS Full-text
DN   179:151475
ED   Entered STN:  16 Jun 2022
TI   Soil improvement material containing methane fermentation digestive juice
 :
DT   Patent
LA   Japanese
CLMN 8
 :
PI
     PATENT NO.          KIND  DATE       APPLICATION NO.          DATE
     ----------------    ----  --------   ---------------------    --------
     JP 2022089338       A     20220616   JP 2020-201668           20201204
PRAI JP 2020-201668
 :
```

/BI,CLM で基本索引と請求項を検索

ALL

請求項の数

請求項の数を確認できる（ALL，MAX 表示形式に含まれる）

日本の特許

© 2022 化学情報協会

特許請求項の検索 - ALL CLM 表示 9

```
 :
AB   The invention relates to a soil improvement material which can effectively
     utilize methane fermn. digestive juice and is easy to handle, and has an
     effect of suppressing soil pests.
 :
IT   Aggregates
        (porous; soil improvement material contg. methane fermn. digestive
        juice)
 :
```

ALL

請求項

```
CLM  What is claimed is:
     1. A soil improvement material that has an inhibitory effect on soil
     disease bacteria, and is a soil improvement material obtained by
     immobilizing methane fermentation digestive juice with a carrier
     material made of a porous material.

        :（途中省略）

     4. The soil improvement material according to any one of claims 1 to
     3, wherein the methane fermentation digestive juice is made from
     sewage sludge, human waste, cow dung, pig dung, and food waste.
        :
```

CLM

書誌情報
抄録
索引
引用情報
請求項

請求項 4

ベーシック特許の請求項が**英語**で収録

請求項のみに記載されたキーワードでヒット

【請求項３】
前記メタン発酵消化液の細菌叢における主要細菌が、クロストリジウム（Clostridium）属細菌及びクロストリジウム綱（Clostridia）に属する細菌である、請求項１又は２に記載の土壌改良資材。
【請求項４】
前記メタン発酵消化液が、下水汚泥、し尿、牛糞、豚糞、食品廃棄物を原料とする、請求項１～３のいずれか１項に記載の、土壌改良資材。

日本語

© 2022 化学情報協会

Prior Art Analysis 機能の搭載 10

CAS が開発した AI ベースの類似特許検索エンジンを使用して、自動的に先行技術文献を検索

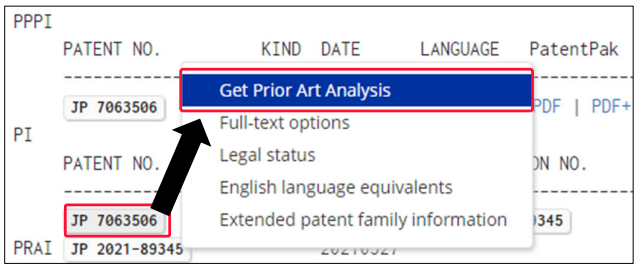
- 他の検索方法では見つからなかった先行技術文献が得られる可能性がある
- 簡単かつ迅速に検索できる
- 特許および非特許の文献が得られる

ワンクリックで
先行技術文献が得られる

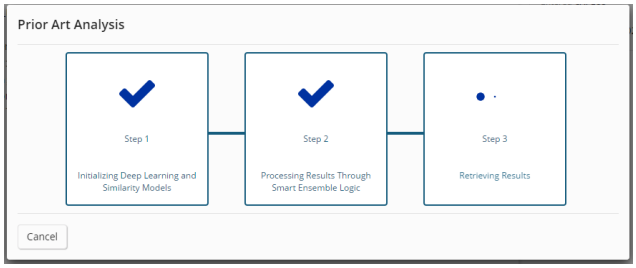
© 2022 化学情報協会

Prior Art Analysis 機能の利用方法 11

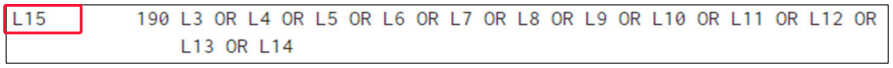
① 特許番号をクリックして
Get Prior Art Analysis を選択



② AI による検索が開始



③ 得られた複数のレコードが HCAplus ファイルで一つにまとめられ、回答集合が得られる



© 2022 化学情報協会

Prior Art Analysis - 検索例

12

新型コロナウイルスの検出方法に関する特許 US11149320 の先行技術文献を調査する

```
=> FILE CAPLUS

=> S US11149320/PN.B
L1            1 US11149320/PN.B

=> D BIB

L1   ANSWER 1 OF 1  CAPLUS  COPYRIGHT 2022 ACS on STN
AN   2021:2255055  CAPLUS Full-text
DN   176:397366
TI   LAMP- and PCR-based assays for the detection of SARS-CoV-2
IN   Brambati, Chiara; Bocchetta, Simone; Minnucci, Giulia
 :
PI
     PATENT NO.                                  TION NO.          DATE
     ------------                                ---------------   --------
     US 11149320                                 -17078249         20201023
     US 2021032448
     US 10815539                                 -16837364         20200401
     WO 2021198325        A1   20211007    WO 2021-EP58424         20210331
     US 20210340636       A1   20211104    US 2021-17374333        20210713
```

Get Prior Art Analysis
Full-text options
Legal status
English language equivalents
Extended patent family information

特許 US11149320
「LAMP 法および PCR 法に基づいた SARS-CoV-2 の検出方法」において **Get Prior Art Analysis** を実行

© 2022 化学情報協会

Prior Art Analysis - 実施されるコマンド

13

```
=> fil HCAPLUS  ←自動的に HCAplus ファイルに入る
                                 得られた先行技術文献のレコード番号(AN)による検索が実行
=> QUE (2020:1564571 OR 2020:1585215 OR 2020:2286915 OR 2020:1972097 OR 2020:2334651 OR
2020:2540492 OR 2020:1810140 OR 2020:2286901 OR 2020:1319887 OR 2020:1370729 OR 2020:2202449
OR 2020:2046015 OR 202                                :904311 OR 2010:293140 OR
2020:2487665)/AN

L2   QUE (2020:1564571 OR 2020:1585215 OR 2020:2286915 OR 2020:1972097 OR
      ....  OR 2020:936850 OR 2020:904311 OR 2010:293140 OR 2020:2487665)/AN
 :
=> QUE (2013:828379 OR 2005:1055750 OR 2018:1542754 OR 2013:835535 OR 2015:675177 OR
      ....  OR 2020:2157046 OR 2014:617571)/AN

L7   QUE (2013:828379 OR 2005:                                3:835535 OR 2015:675177 OR
      ....  OR 2020:2157046 OR 2014:617571)/AN
 :
=> QUE (1999:194298 OR 2017:1964291 OR 1997:735877 OR 2005:1026840 OR 2008:805956 OR
1997:491567 OR 2001:748059 OR 2002:522048 OR 2003:472652 OR 2009:1536083 OR 2008:1363987)/AN

L13  QUE (1999:194298 OR 2017:1964291 OR 1997:735877 OR 2005:1026840 OR
     .... OR 2009:1536083 OR 2008:1363987)/AN

=> S L2 OR L3 OR L4 OR L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11 OR L12 OR L13
L14        192 L2 OR L3 OR L4 OR L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11 OR L12 OR L13

=> S L14 AND P/DT
L15         93 L14 AND P/DT

=> S L14 NOT P/DT
L16         99 L14 NOT P/DT
```

最も関連度が高い非特許文献

最も関連度が高い特許文献

自動で実行される

レコード番号は非特許文献、特許文献の順にまとまっている
➢ それぞれ関連度の高い順に整理される

最大 200 件の文献が得られる

192 件の文献が得られた
(特許 93 件、非特許 99 件)

© 2022 化学情報協会

Prior Art Analysis - ヒットしたレコード 14

ヒットした関連度が最も高い非特許文献

```
AN   2020:1564571  HCAPLUS Full-text
DN   173:382188
ED   Entered STN:  12 Aug 2020
TI   Comparison of the analytical sensitivity of seven commonly used commercial
     SARS-CoV-2 automated molecular assays
AU   Mostafa, Heba H.; Hardick, Justin; Morehead, Elizabeth; Miller, Jo-Anne;
     Gaydos, Charlotte A.; Manabe, Yukari C.
CS   Division of Medical Microbiology, School of Medicine, Department of
     Pathology, Johns Hopkins University, Baltimore, MD, USA
SO   Journal of Clinical Virology (2020), 130, 104578
     CODEN: JCVIFB; ISSN: 1386-6532
 :
AB   The SARS-CoV-2 pandemic has challenged mol. microbiol. labs. to quickly
     implement and validate diagnostic assays and to expand testing capacity in
     a short timeframe.  Multiple mol. diagnostic methods received FDA
     emergency use authorization (EUA) and were promptly validated for use
     nationwide.
 :
IT   Microbial gene
     RL: ANT (Analyte); DGN (Diagnostic use); ANST (Analytical study); BIOL
     (Biological study); USES (Uses)
        (E; comparison of the anal. sensitivity of seven commonly used com.
        SARS-CoV-2 automated mol. assays)
 :
IT   COVID-19
     Homo sapiens
     Human
 :
```

SARS-CoV-2 の検出方法の比較についての非特許文献が得られた

© 2022 化学情報協会

Prior Art Analysis - ヒットしたレコード 15

ヒットした関連度が最も高い特許文献

```
AN   2020:2157046  HCAPLUS Full-text
DN   173:810242
ED   Entered STN:  28 Oct 2020
TI   Oligonucleotide primers and probes and assays for the detection of
     SARS-CoV-2
IN   Brambati, Chiara; Bocchetta, Simone; Minnucci, Giulia
PA   DiaSorin S.p.A., Italy
 :
PI
     PATENT NO.          KIND  DATE        APPLICATION NO.          DATE
     ----------------    ----  --------    ---------------------    --------
     US 10815539          B1   20201027    US 2020-16837364         20200401
     US 11149320          B1   20211019    US 2020-17078249         20201023
     US 20210324484       A1   20211021
     WO 2021198325        A1   20211007    WO 2021-EP58424          20210331
     WO 2021198326        A1   20211007    WO 2021-EP58426          20210331
     US 20210340636       A1   20211104    US 2021-17374333         20210713
 :
AB   The present invention is directed to methods for assaying for the presence
     of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in a
     sample, …
 :
IT   Microbial gene
     RL: ANT (Analyte); DGN (Diagnostic use); ANST (Analytical study); BIOL
     (Biological study); USES (Uses)
        (ORF1ab; oligonucleotide primers and probes and assays for detection
        of SARS-CoV-2)
```

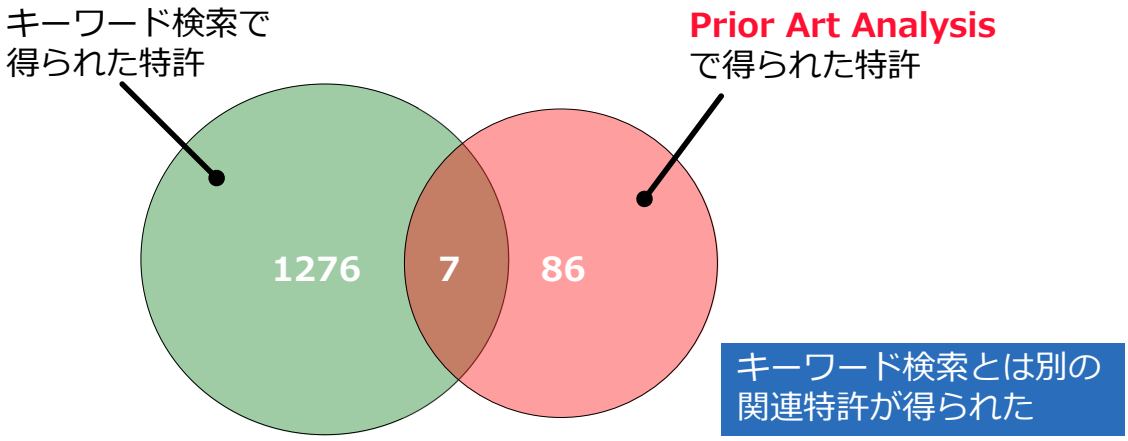
SARS-CoV-2 検出方法に関する特許文献が得られた

© 2022 化学情報協会

Prior Art Analysis – キーワード検索との比較 16

```
=> S (SARS COV 2 OR NOVEL CORONA VIRUS OR COVID 19) (10A) (DETECT? OR ANALY?) AND
(AD<20201023 OR PRD<20201023)
L17       1283 (SARS COV 2 OR NOVEL CORONA VIRUS OR COVID 19) (10A) (DETECT?
               OR ANALY?) AND (AD<20201023 OR PRD<20201023)

=> S L15 AND L17
L18          7 L15 AND L17

=> S L15 NOT L17
L19         86 L15 NOT L17
```

キーワード検索で SARS-CoV-2 の検出についての特許調査



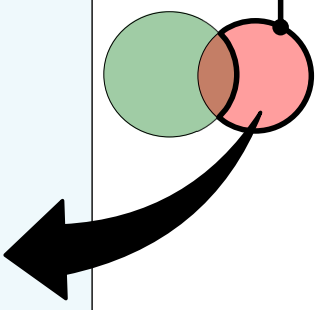
キーワード検索とは別の関連特許が得られた

© 2022 化学情報協会

Prior Art Analysis – ヒットしたレコード 17

```
=> D L19 4 ALL

L19  ANSWER 4 OF 86  HCAPLUS  COPYRIGHT 2022 ACS on STN
AN   2017:1964291  HCAPLUS Full-text
DN   168:72736
ED   Entered STN:  14 Dec 2017
TI   Enzyme-linked oligonucleotide sorbent assay method for detecting
     oligonucleotides in a test sample
IN   Walker, Ameae M.; Lorenson, Mary Y.
PA   The Regents of the University of California, USA
 :
PI
     PATENT NO.          KIND  DATE        APPLICATION NO.          DATE
     ---------------     ----  --------    ---------------------    --------
     WO 2017214145        A1   20171214    WO 2017-US36153          20170606
     US 20190136302       A1   20190509    US 2018-16306807         20181203
 :
AB   Certain embodiments of the invention provide a method (i.e., Enzyme Linked
     Oligonucleotide Sorbent Assay (ELOSA)) for the detection and/or
     quantification of a test oligonucleotide (e.g., a small oligonucleotide)
     in a test sample, such as a biol. fluid, comprising:
 :
IT   Bioassay
        (ELOSA (Enzyme-linked oligonucleotide sorbent); enzyme-linked
        oligonucleotide sorbent assay method for detecting oligonucleotides in
        a test sample)
 :
```

Prior Art Analysis のみで得られた特許

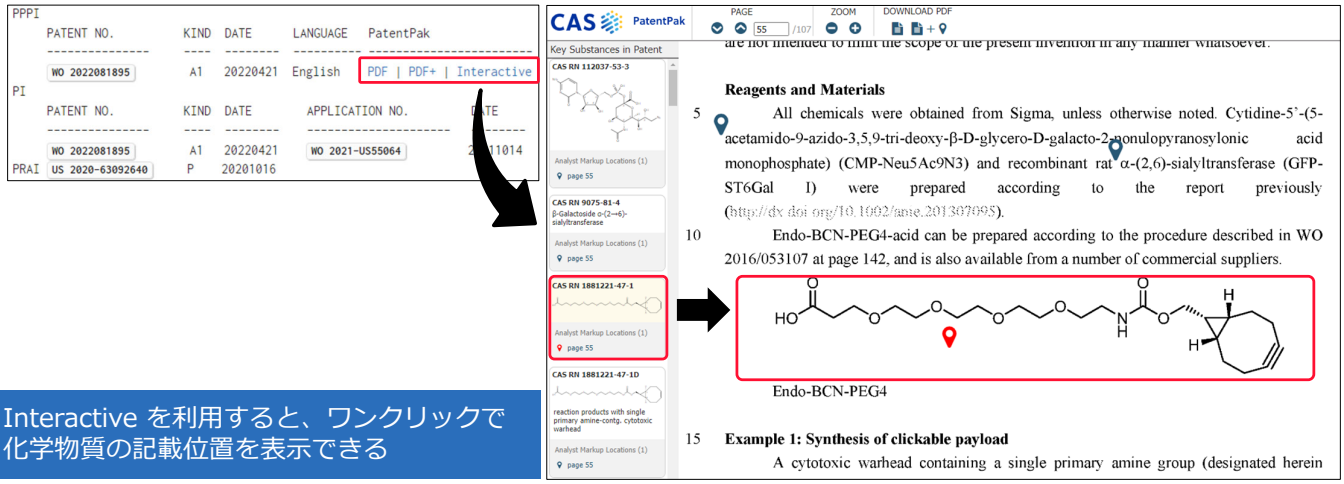


核酸（オリゴヌクレオチド）の検出に関する特許文献が得られた

© 2022 化学情報協会

CAS PatentPak 18

- 特許明細書中における化学物質の記載位置を瞬時に確認できる CAS 独自の機能
- 明細書 PDF をワンクリックで簡単にダウンロードできる

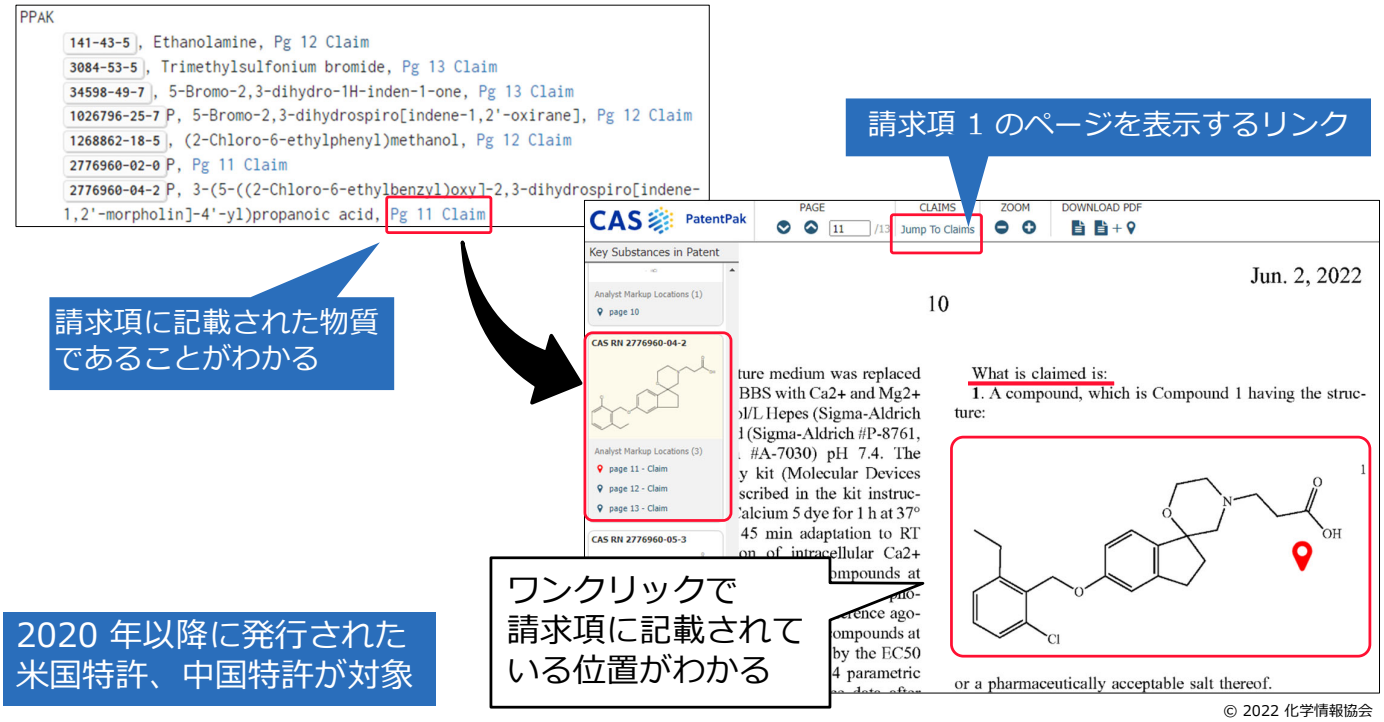


Interactive を利用すると、ワンクリックで化学物質の記載位置を表示できる

© 2022 化学情報協会

CAS PatentPak - Claim タグの付与開始 19

物質が請求項に記載されている場合に
PPAK フィールドに Claim タグが付くようになった



請求項に記載された物質であることがわかる

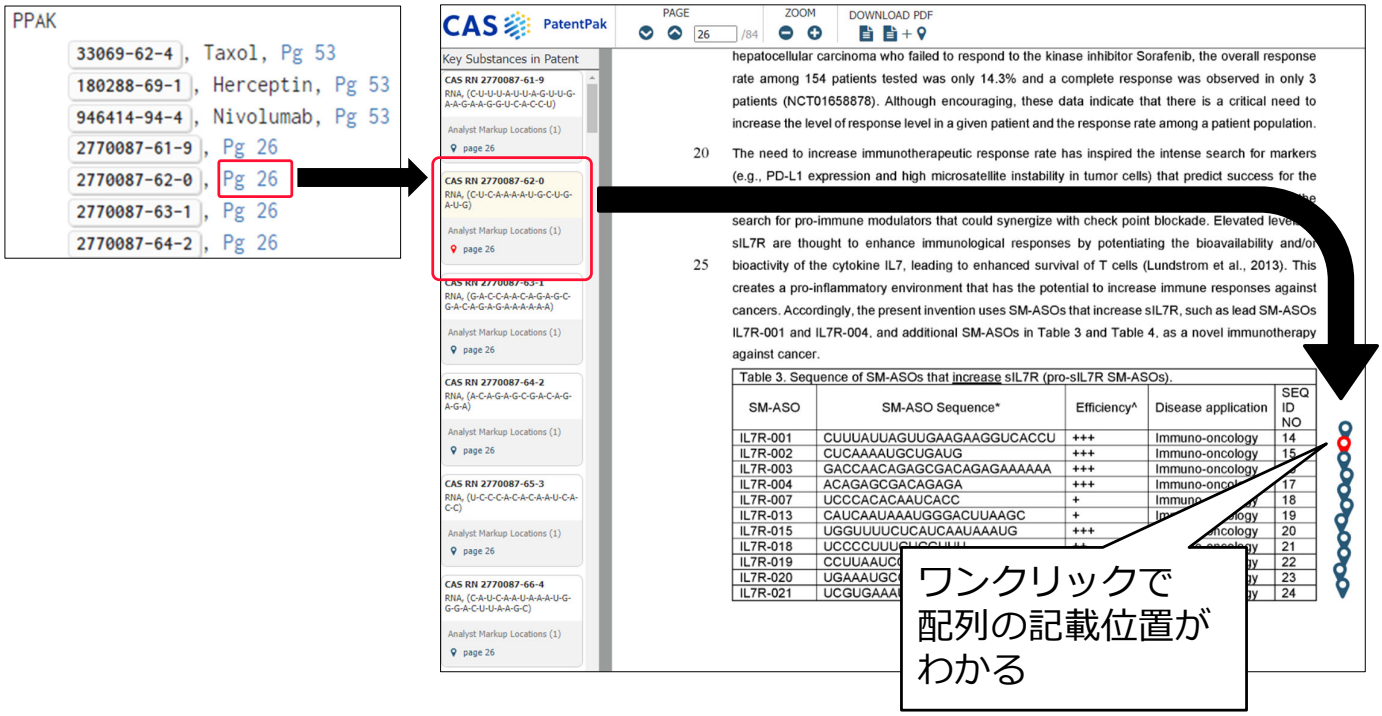
ワンクリックで請求項に記載されている位置がわかる

2020 年以降に発行された米国特許、中国特許が対象

© 2022 化学情報協会

CAS PatentPak - 配列記載位置の付与開始 20

配列の記載位置の収録開始



© 2022 化学情報協会

CAS PatentPak のリンクの有効期限延長 21

リンクの有効期限が 365 日に延長

レポート (Excel)

	A	B	C	D	E	F	G
1	Answer Line	Database	PatentPak PDF	PatentPak PDF+	PatentPak Interactive	Title	Accession Number
2	L3 8 OF 4761	CAPLUS	PatentPak PDF	PatentPak PDF+	PatentPak Interactive	Soluble interleukin-7 receptor (IL7R) modulating therapy to treat autoimmune disorder and cancer	2022:119741 8 CAPLUS
3							
4	L3 10 OF 4761	CAPLUS	PatentPak PDF	PatentPak PDF+	PatentPak Interactive	Preparation of prodrug glycoconjugates having antibody cell-binding agent	2022:103957 3 CAPLUS

リンクの有効期限が
90 日から 365 日 へ延長

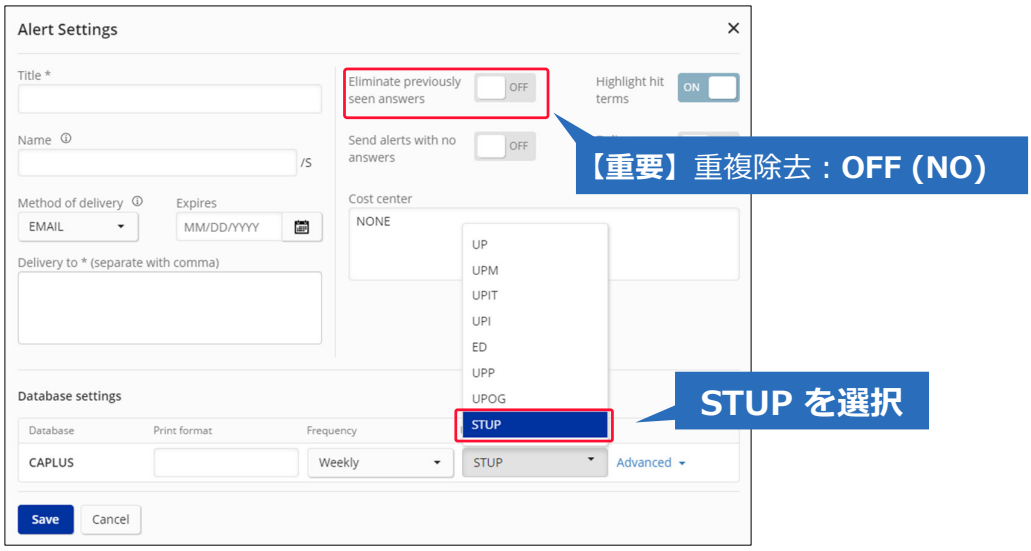
セッション記録

```
L3   ANSWER 10 OF 4761  CAPLUS  COPYRIGHT 2022 ACS on STN
PatentPak PDF | PatentPak PDF+ | PatentPak Interactive
PPAK
     65-47-4, CTP, Pg 56
     98-59-9, 4-Toluenesulfonyl chloride, Pg 74
     104-94-9, 4-Methoxyaniline, Pg 71
     107-86-8, 3-Methylbut-2-enal, Pg 71
     108-24-7, Acetic anhydride, Pg 76
     108-30-5, Succinic anhydride, Pg 66
     814-68-6, Acryloyl chloride, Pg 74
     821-09-0, 4-Penten-1-ol, Pg 74
```

© 2022 化学情報協会

特許ステータスのアラート設定 22

特許ステータス更新コード STUP がアラート更新コードとして設定可能になった



CAplus/CA ファイルで特許ステータス更新のウォッチングが可能に！

© 2022 化学情報協会

MARPAT ファイルの強化

MARPAT ファイル 24

マルクーシュ構造が検索できる特許データベース

(2022 年 8 月)

製作者	CAS
レコード構成	文献単位 (発明単位)
収録期間	1961 年～
収録件数	130 万件以上のマルクーシュ構造
特長	• CA ファイルに収録する特許のうち、マルクーシュ構造の記載がある特許を収録 • 有機化合物、有機金属化合物、重合度が 10 以下のオリゴマーが収録対象

© 2022 化学情報協会

特許ステータス収録開始 25

書誌情報に特許ステータスが収録されるようになった

```
AN   163:544811  MARPAT Full-text
TI   Asparagine endopeptidase inhibitors for managing cancer and related
     compositions
IN   Ye, Keqiang; Fu, Haian; Du, Yuhong
PA   Emory University, USA
 :
PI
     PATENT NO.          KIND  DATE        APPLICATION NO.          DATE
     ----------------    ----  --------    ---------------------    --------
     WO 2015157376        A1   20151015    WO 2015-US24850          20150408
     EP 3129375           A1   20170215    EP 2015-776775           20150408
     US 20170166569       A1   20170615    US 2016-15303481         20161011
     US 10759803          B2   20200901
PRAI US 2014-61978225          20140411
     WO 2015-US24850           20150408
PSPI
     PATENT NO.          KIND  STATUS          STATUS DATE
     ----------------    ----  -------------   -----------
     WO 2015157376        A1   Dead            20201202
     EP 3129375           A1   Dead            20201203
     US 20170166569       A1   Alive           20201121
     US 10759803          B2   Alive           20201121
 :
```

CAplus/CA ファイルの書誌情報、抄録、索引が収録されている

特許ステータス表示 PSPI フィールド

- 特許ステータス (PSPI) は CAplus, CA, MARPAT ファイルに収録されている。
  より詳細な法的状況データは INPADOCDB/INPAFAMDB ファイルで確認できる。

© 2022 化学情報協会

特許ステータス（PSPI）分類

特許ステータス	1文字コード	内容
ALIVE	**A**	出願中，特許権存続期間中など、有効な場合に付与
DEAD	**D**	期間満了、放棄、取り下げなど、有効な可能性がない場合に付与
TRANSITIONAL	**T**	最新の法的状況イベントで特許失効になる可能性があることが示された場合に付与（取り下げ、年金未払いなど）
INDETERMINATE	**I**	欧州登録特許固有のステータス。欧州登録特許は各国の国内特許権となるため、期間満了日 (計算値) までの間付与

ALIVE に戻る可能性もある

© 2022 化学情報協会

CHEMLIST ファイルの強化

CHEMLIST ファイルの概要 28

化学物質の規制に関するデータベース

(2022 年 8 月)

製作者	CAS
レコード構成	化学物質単位
収録期間	1980 年～
収録件数	41 万件以上
特長	• 世界の主要国の既存化学物質リスト情報を収録 • その他各国により規制されている化学物質情報を収録

© 2022 化学情報協会

UK REACH の収録開始 29

- RSTR フィールドにイギリスの UK REACH より下記リストが追加された
  - ・Substances of Very High Concern (SVHC)
  - ・Substances Subject to Authorisation (Annex XIV)

```
=> FILE CHEMLIST

=> S UK REACH/RSTR
L1          293 UK REACH/RSTR
                ((UK(W)REACH)/RSTR)

=> D L1 1 IDE

L1   ANSWER 1 of 293  CHEMLIST  COPYRIGHT 2022 ACS on STN
AN    440519  CHEMLIST
CN    Oligomers of chromic acid and dichromic acid
FS    Restricted Chemical Lists: RSTR
CBI   Public
RSTR  UK REACH: Candidate List of Substances of Very High Concern for
      Authorisation
      Listed Name(s): Oligomers of chromic acid and dichromic acid
      Reason for Inclusion: Carcinogenic (Article 57a).
FA    RSTR    Restricted Chemical Lists
```

=> S UK REACH/RSTR で検索

制限されている化学物質リスト (RSTR フィールド) に表示

© 2022 化学情報協会

CAS STNext ユーザーミーティング

WPI ファイルのリロード

WPI ファイルのリロード

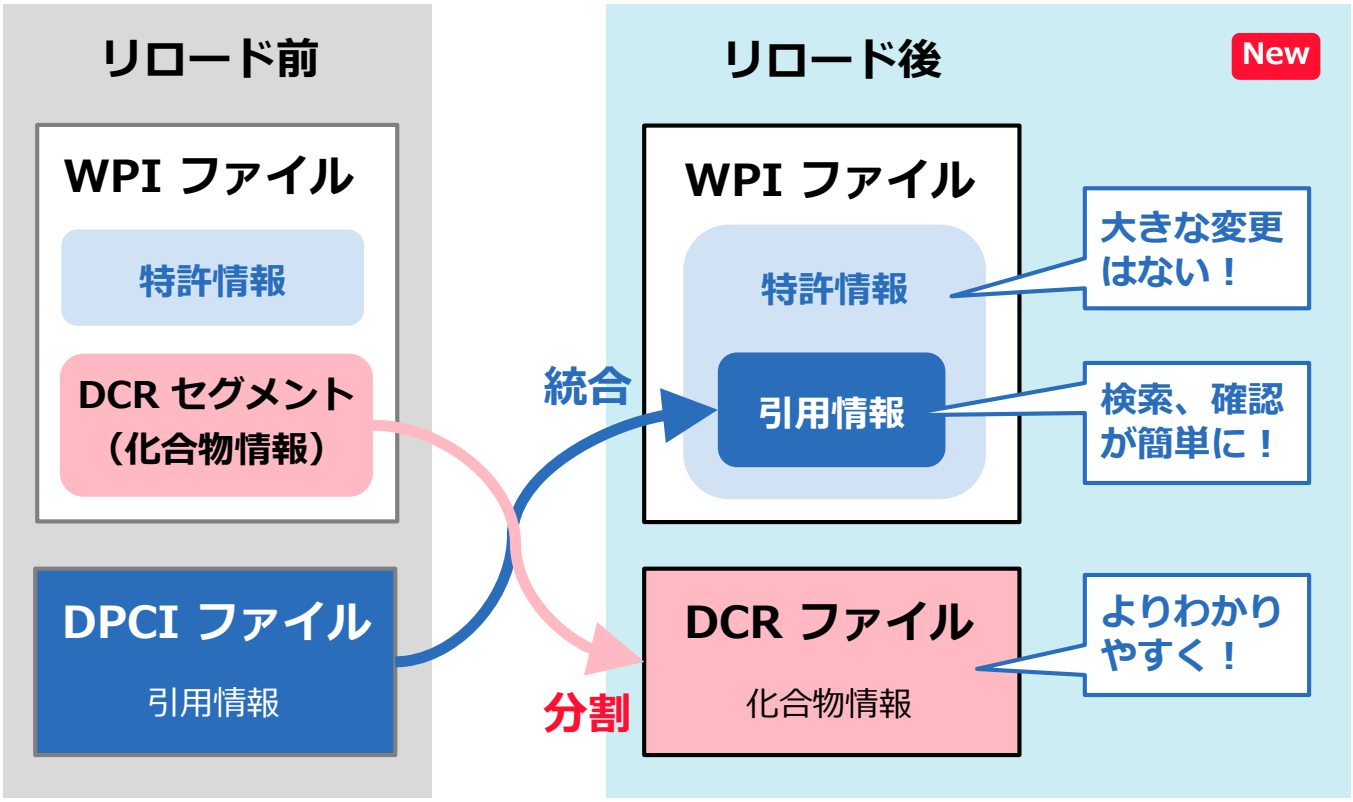
化学情報協会　情報事業部

目次

1. リロードの概要
2. WPI ファイル
3. DCR ファイル

© 2022 化学情報協会

リロードの概要 2



© 2022 化学情報協会

WPI ファイル

WPI ファイルの概要 4

世界中の特許を収録しているデータベース

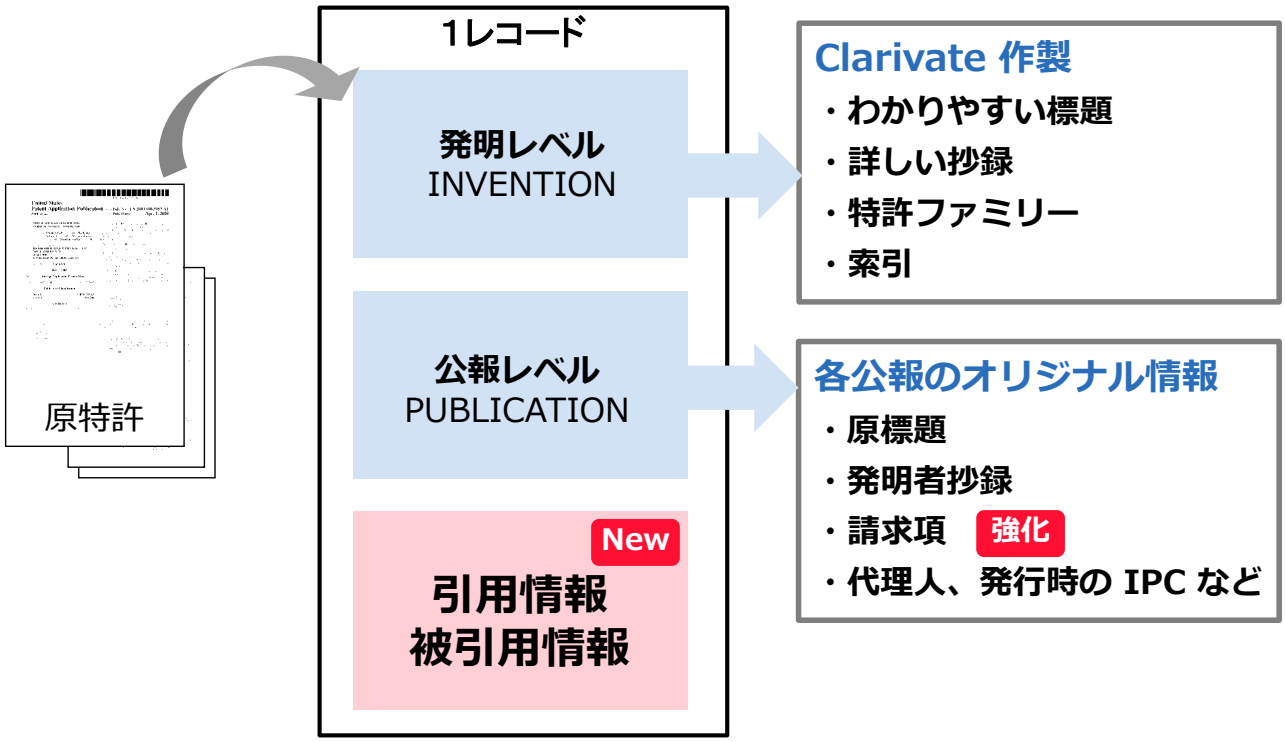
(2022 年 8 月)

製作者	Clarivate
収録国	59 特許発行機関、 2 技術公開誌
収録内容	書誌情報、抄録、索引、特許分類、独自の分類コード
収録分野	**全技術分野**
レコード構成	**発明単位**
収録期間	1963 年~ (国、分野によって異なる)
更新頻度	週 2 回
特長	• Clarivate 作製の詳細な特許標題、抄録、索引を利用できる • 特許公報由来の原題、発明者抄録、クレームなども一部収録

© 2022 化学情報協会

WPI ファイルのレコード構成 5

1 レコード＝1 発明＝ 1 特許ファミリー



© 2022 化学情報協会

WPI ファイルのリロードのハイライト 6

特許情報はこれまでと同様に検索できる

強化点

- 引用情報の収録開始（旧 DPCI ファイルの統合）
- 特許請求項の収録対象の拡大
- 中間一致・後方一致検索可能なフィールドの拡大
- ケミカルフラグメンテーションコードに関する強化

© 2022 化学情報協会

引用情報の収録開始 7

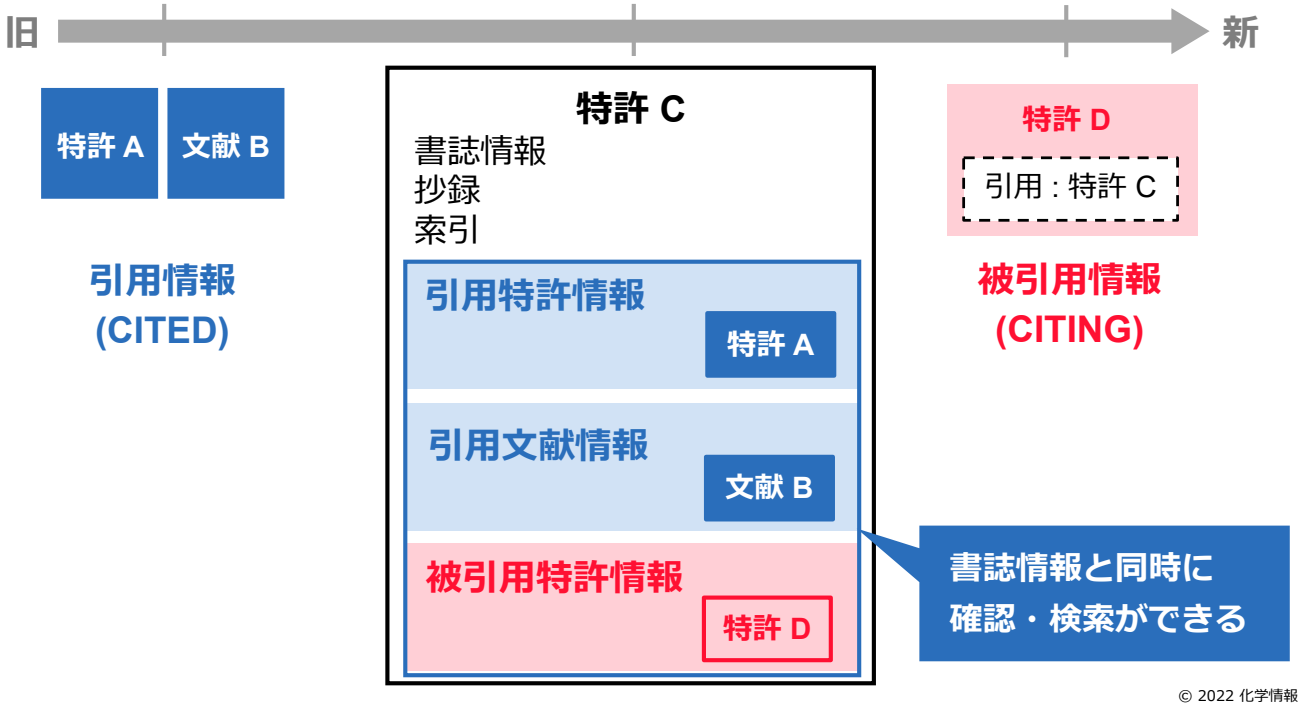
旧 DPCI ファイルの引用情報が WPI ファイルに統合

- 書誌情報と引用・被引用情報を同時に表示可能に
- 特許情報と組み合わせた引用・被引用情報の検索が WPI ファイルで完結

© 2022 化学情報協会

引用情報の収録（レコード構成） 8

引用特許情報、引用文献情報、被引用特許情報がレコードに収録



© 2022 化学情報協会

レコード例（BIB CITN 表示形式） 9

BIB

```
AN   2014-E51864 [201428]   WPINDEX Full-text
TI   New indole carboxamide derivative, useful for treating tuberculosis in a human having sputum
     smear-positive, sputum smear-negative, or extrapulmonary tuberculosis, caused by drug resistant
     Mycobacterium tuberculosis complex organisms
DC   B02
IN   JIRICEK J; KONDREDDI R; KONDREDDI R R; KONDREDDI RAVINDER R; PAUL W; SMITH P W; SMITH PAUL W; SMITH R
PA   (NOVS-C) NOVARTIS AG; (JIRI-I) JIRICEK J; (KOND-I) KONDREDDI R R; (SMIT-I) SMITH P W
CYC  130
PI   WO 2014037900    A1 20140313 (201428) * EN 93[0]
     CA 2881351       A1 20140313 (201514)   EN
     AU 2013311235    A1 20150226 (201518)   EN
 :
```

引用・被引用数は、特許の価値を評価する指標として広く利用される

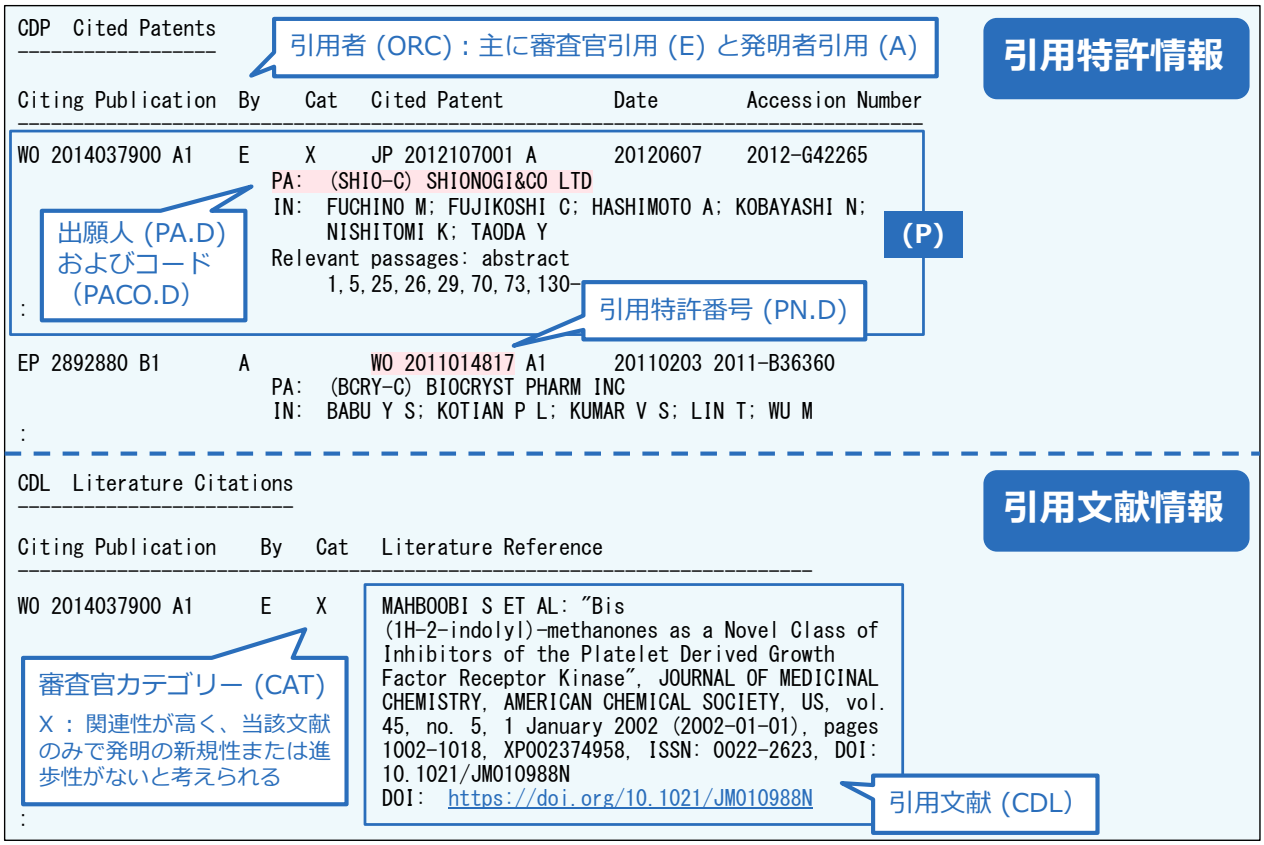
CITN

```
Citation Counters
-----------------
PNC.D     23      Cited Patents Count                      ← 引用特許数 (/PNC.D)
PCC.D     4       Cited Patents Country Count              ← 引用特許の特許発行国数 (/PCC.D)
CRC       125     Cited Literature Reference Count         ← 引用文献数 (/CRC)
REC       148     Cited Reference Count (total)            ← 引用数 (合計) (/REC)
ANC.D     15      Cited DWPI Accession Number Count        ← 引用特許の WPI レコード数 (/ANC.D)
PNC.G     15      Citing Patents Count                     ← 被引用特許数 (/PNC.G)
PCC.G     5       Citing Patents Country Count             ← 被引用特許の特許発行国数 (/PCC.G)
ANC.G     13      Citing DWPI Accession Number Count       ← 被引用特許の WPI レコード数 (/ANC.G)

EXF Examiner 's Field of Search                            ← 審査官調査フィールド
-------------------------------

Citing Publication    EXF  Examiner's Field of Search
-----------------------------------------------------
JP 6223452 B2         JPC  C07D; A61K; A61P
```

特許審査官が引用特許を検索した際に使用した特許分類を収録

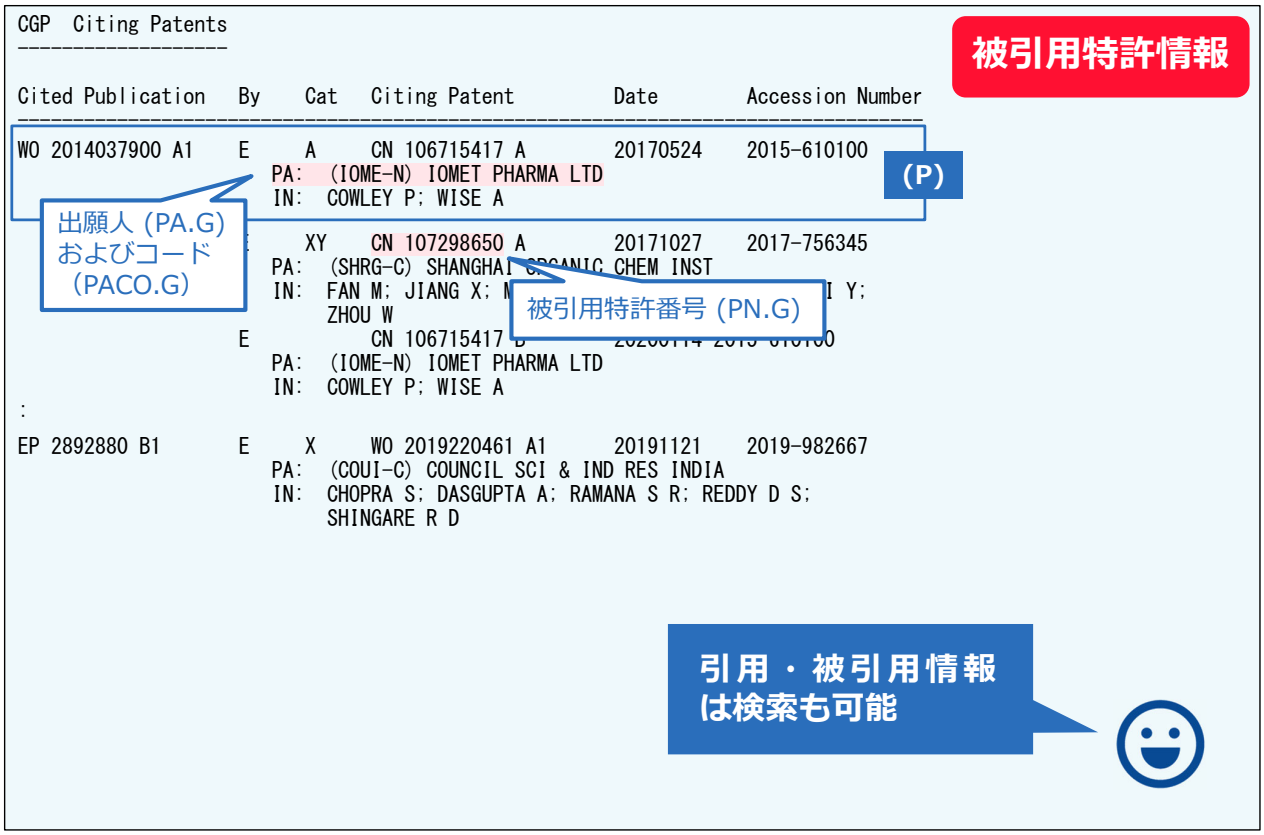
© 2022 化学情報協会

レコード例（つづき） 10



© 2022 化学情報協会

レコード例（つづき） 11



© 2022 化学情報協会

引用情報の検索フィールド 12

専用の検索フィールドを使うことで、簡単に検索可能

検索フィールド	本願	引用特許	被引用特許
レコード番号	/AN	/AN.D	/AN.G
出願人コード	/PACO	/PACO.D	/PACO.G
発行国	/PC	/PC.D	/PC.G
特許番号	/PN	/PN.D	/PN.G
発行年	/PY	/PY.D	/PY.G
審査官カテゴリー	-	/CAT	
		/CAT.D	/CAT.G
引用者	-	/ORC	
引用・被引用特許のレコード数	-	/ANC.D	/ANC.G

引用特許（CITE**D**）の検索には **.D** を、
被引用特許（CITIN**G**）の検索には **.G** をつけるだけ！

© 2022 化学情報協会

引用情報の表示形式 13

引用・被引用情報を両方表示したい場合、**CITN** 表示形式（省略形式： **CITNA** 表示形式）を利用

	表示形式 (括弧内は重複を除いた省略形式*)	
引用数	CITC	CITN (CITNA)
審査官の検索分野	EXF (EXFA)	
引用特許	CDP (CDPA)	
引用文献	CDL (CDLA)	
被引用特許	CGP (CGPA)	

* レコードに含まれるファミリ内のすべての特許ごとに引用・被引用情報が収録されるため、重複している情報もある。省略形式を利用すると重複を除いてコンパクトに表示することができる

© 2022 化学情報協会

CRISPR/Cas システムによる多重ゲノム編集について広く特許調査する

- キーワード検索と引用・被引用情報の検索を併用
  - キーワード検索では公報レベルを含めて網羅的に検索
    /BI,BIEX
  - 引用・被引用特許情報中の WPI レコード番号を利用して、引用・被引用特許を検索

CRISPR/Cas システム：
DNA の二本鎖切断を原理とする高精度・簡単・安価な遺伝子改変技術

© 2022 化学情報協会


```
=> FILE WPINDEX                          ← WPI ファイルに入る
=> S ((CRISPR (S) CAS##) AND (MULTI? (2A) GEN? (2A) EDIT?))/BI,BIEX

L1        102 ((CRISPR (S) CAS##) AND (MULTI? (2A) GEN? (2A) EDIT?))/BI,BIEX

=> D CITNA 100-102                       ← 引用情報を確認（CITNA 表示形式：コンパクトに引用、被引用情報を表示）

L1   ANSWER 100 OF 102  WPINDEX COPYRIGHT 2022   CLARIVATE on STN
AN   2014-T38051 [201469]   WPINDEX

Citation Counters
-----------------
PNC.D     63     Cited Patents Count
:
PNC.G     148    Citing Patents Count
:

Cited Patents
-------------
Cited Publication     By          Accession Number
-------------------------------------------------
US 5571690 A          E           1996-505399
US 8252535 B2         E           2008-N67499
US 8697359 B1         E           2014-G67519
:

Citing Patents
--------------
Citing Publication    By          Accession Number
-------------------------------------------------
CN 112195164 A        E           2021-08582H
CN 112195164 B        E           2021-08582H
:
```

キーワード検索

引用・被引用特許番号、引用者および WPI レコード番号がコンパクトに表示

引用特許のレコード番号
/AN.D

被引用特許のレコード番号
/AN.G

関連技術分野も広く調査するために、引用・被引用特許を検索

© 2022 化学情報協会

検索例 1


```
=> TRA L1 AN.D AN.G /AN
L2          TRANSFER L1 1- AN.D AN.G :     1521 TERMS
L3        1414 L2/AN

=> S L3 OR L1              ← L1 と L3 をまとめる
L4        1478 L3 OR L1

=> D MAX MEMB 1-2          ← MAX MEMB 表示形式で回答を確認

L4   ANSWER 1 OF 1478  WPINDEX COPYRIGHT 2022   CLARIVATE on STN
AN   2022-77868U [2022056]   WPINDEX Full-text
CR   2021-52487U; 2021-A1774W; 2021-D02876; 2022-351510
TI   Type I CRISPR system for identifying cells having genome that is rationally edited comprises
     first vector comprising nucleic acid sequence encoding first half of split protein reporter
     system comprising cascade-T7 polymerase fusion protein
DC   B04; C06; D16
IN   FEDEROWICZ S; GARST A; MIR A; SEAMON K
:

L4   ANSWER 2 OF 1478  WPINDEX COPYRIGHT 2022   CLARIVATE on STN
AN   2022-713894 [2022052]   WPINDEX Full-text
TI   Genetically altering target cells by combining within each of wells of first well plate, liposome
     comprising crRNA portion of guide RNA, liposome comprising tracrRNA portion, liposome comprising Cas
     enzyme, and target cell
:
AB   US 20220162601 A1       UPAB 20220701
:
 ADVANTAGE - The methods using CRISPR-Cas systems provide a platform for high efficiency genome editing of
     cells. The use of a multi-well plate and CRISPR-Cas system components produces high throughput,
     multiplexed genome-wide cell editing.
:
```

L1 のレコードから引用・被引用情報中の WPI レコード番号（AN.G, AN.D）を抽出し、レコード番号 /AN で検索

引用・被引用検索でヒット

キーワード検索でヒット

© 2022 化学情報協会

検索例 1


```
Member (0001)
PI   US 20220162601  A1 20220526 (2022052)* EN
TIEN HIGH THROUGHPUT GENE EDITING SYSTEM AND METHOD
:
CLMEN [CLAIM 1] 1. A method of genetically altering a plurality of target cells
      comprising (a) combining within each of a plurality of wells of a first
      well plate (1) a liposome comprising a crRNA portion of a guide RNA,
       :
        :
      [CLAIM 29] 29. The method of claim 1 wherein the well concentration of
      Cas enzyme is between 60 nM and 80 nM.
      [CLAIM 30] 30. The method of claim 1 wherein the liposome comprising a
      crRNA portion of a guide RNA, the liposome comprising a tracrRNA portion
      of the guide RNA, the liposome comprising a Cas enzyme or a nucleic acid
      sequence encoding the Cas enzyme, and the target cell are provided to the
      well by droplet transfer from a source container to the well using sound
      waves.
      [CLAIM 31] 31. The method of claim 1 wherein the Cas enzyme is a Cas
      enzyme of a Type II CRISPR system.
```

(S)
(S)
(S)
(S)
(P)

特許請求項の演算子の範囲 New
- (P) 演算子：同一公報の同一言語内
- (S) 演算子：同一クレーム（パラグラフ）内

© 2022 化学情報協会

特許請求項の収録対象国の拡大 18

下記 21 特許発行機関で全クレームが収録されるように

全クレーム収録				
アルゼンチン (AR)	ブラジル (BR)	カナダ (CA)	スイス (CH)	中国 (CN)
ヨーロッパ特許庁 (EP)	スペイン (ES)	フランス (FR)	イギリス (GB)	インドネシア (ID)
インド (IN)	日本 (JP)	韓国 (KR)	マレーシア (MY)	ロシア (RU)
タイ (TH)	トルコ (TR)	台湾 (TW)	アメリカ (US)	ベトナム (VN)
PCT 出願 (WO)				

リロード前、多くはメインクレームのみ収録。(中国・韓国はリロード前も全クレーム収録)

メインクレーム収録		
ドイツ (DE)	メキシコ (MX)	シンガポール (SG)

多くの機関は 2015 年以降の特許レコードで収録。詳細は下記ページを参照
https://www.stn-international.com/en/training-center/documentation/dwpi-summary-table-first-level-patent-data

© 2022 化学情報協会

中間一致・後方一致検索可能なフィールドの拡大 19

下記フィールドで中間一致・後方一致検索が可能

入力例）=> S ?SENSOR?/TIEN

「BIOSENSOR」や「SENSORY」もヒット

基本索引	/BI, /BIEX
標題	/TI, **/TIEN, /TIDE, /TIFR, /TIES, /TIOL**
クレーム	/CLM, **/CLMEN, /CLMDE, /CLMFR, /CLMES, /CLMOL**
抄録	/AB, /TECH, /ABEX※, /ABDE, /ABEN, /ABES, /ABFR, /ABOL, **/ABDT※**

※ WPIX ファイルでのみ表示可能。 ABEX：拡張抄録、ABDT：ドキュメンテーション抄録

赤字が追加

公報レベルの検索がより便利に！

© 2022 化学情報協会

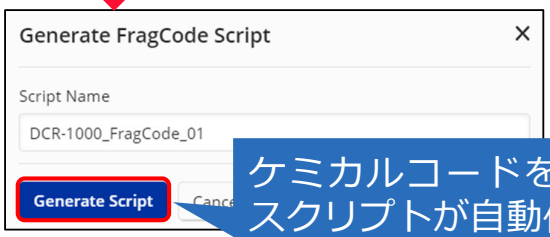
ケミカルフラグメンテーションコードに関する強化　20

CAS STNext に構造質問式からケミカルコードの検索式を自動作成する機能が搭載！

ケミカルフラグメンテーションコードとは、構造を断片的にコード化したもので、WPI ファイルの索引に利用される

```
M3 *01*  D000 D601 M280 M320 M412 M511 M520 M530 M540 M750 N134 Q431
         M905 M904 M910
         DCN-R01003-X DCN-R01003-K
         DCR-1000-X DCR-1000-K
```

WPI ファイルの索引

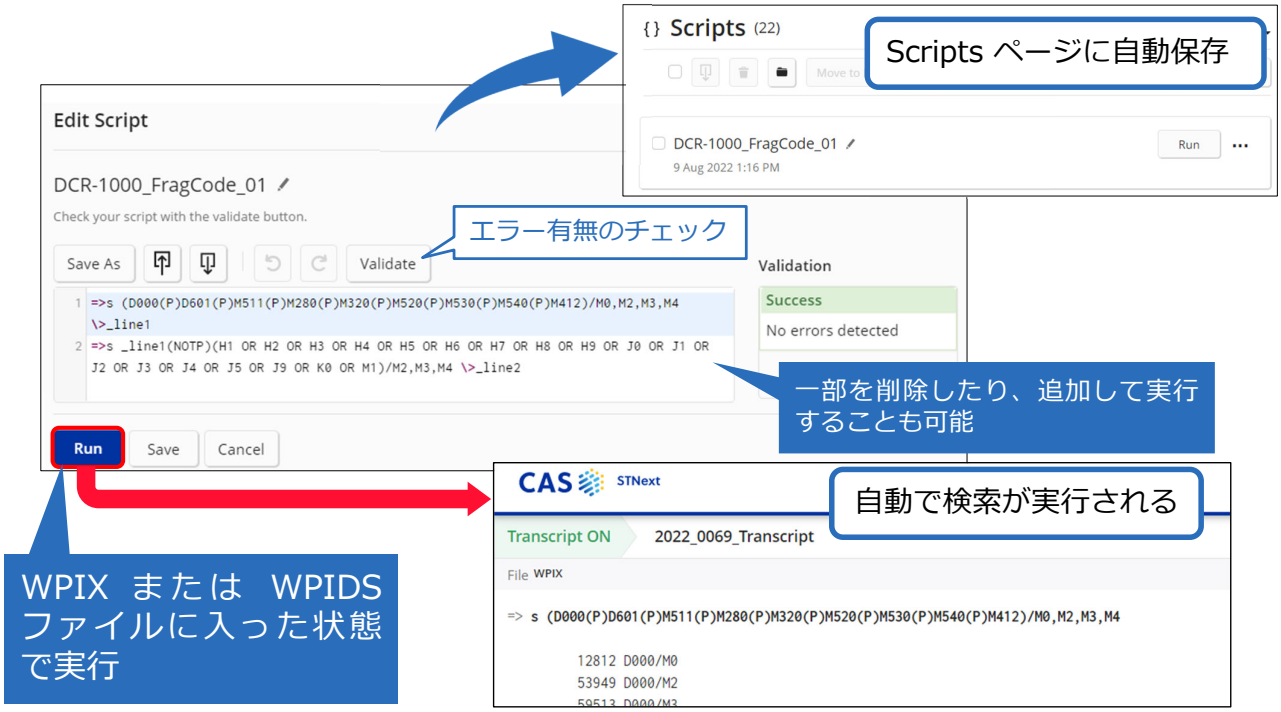


ケミカルコードを含むスクリプトが自動作成

© 2022 化学情報協会

ケミカルフラグメンテーションコードに関する強化　21

作成されたスクリプトの編集画面が自動で立ち上がる



* => S L# (NOTP)(DCR OR MCN)/MALL で DCR, MCN の索引がないレコードに限定できる

© 2022 化学情報協会

その他の変更点

- テキスト中の数値検索機能の強化

  - ppm、 磁界強度、屈折率などが追加され、59 種類に

- 更新コードの追加

新規更新コード	定義
UPD	引用情報の更新
UPG	被引用情報の更新
UPIN	発明者情報 (発明レベル) の更新
UPPA	出願人情報 (発明レベル) の更新

- 特許番号入力形式の変更

	入力形式	検索例
日本登録特許 (JPB, JPB1, JPB2)	番号の後ろに B を追加	=> S JP6972444B/PN
実用新案	番号の後ろに U を追加	=> S CN214259850U/PN

- ロールの検索フィールド /RL の追加（後述）

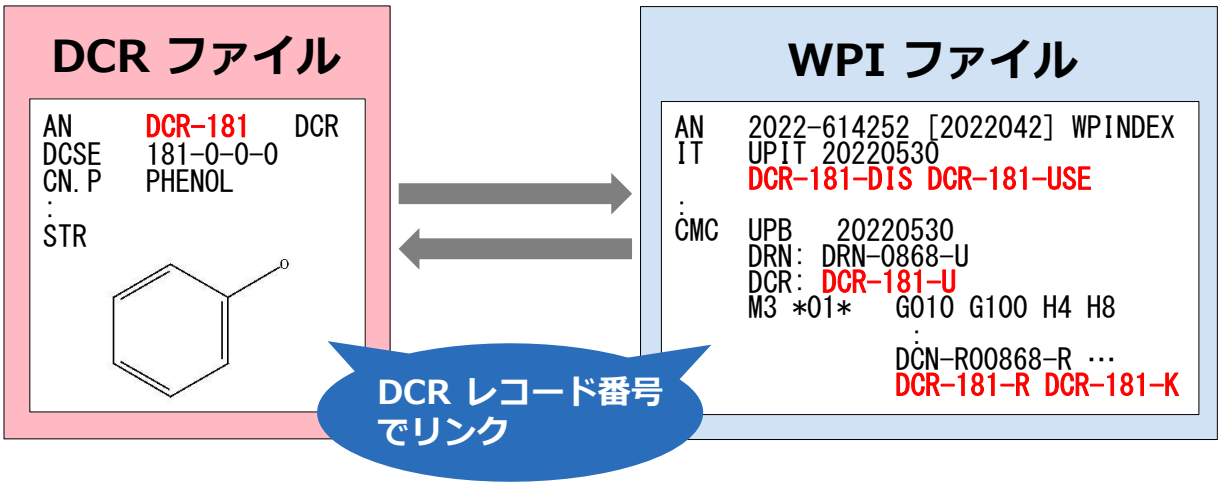
© 2022 化学情報協会

DCR ファイル

DCR ファイルのリリース 24

化学物質情報が WPI ファイルから分離され、独立したデータベースとしてリリース！

- WPI レコードに索引された重要な特定化学物質が収録
  (ダウエントセクション B (医薬)、C (農薬)、E (一般化学))



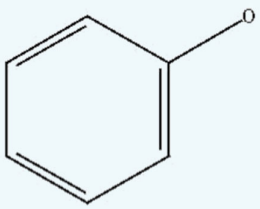
© 2022 化学情報協会

DCR ファイルのレコード例 25

DCR レコード番号
DCR 番号
優先化学物質名
統制化学物質名
同義名

```
AN      DCR-181    DCR
DCSE    181-0-0-0
CN.P    PHENOL
CN.S    Phenol
SY      2,5-CYCLOHEXADIENE-1-ONE; CARBOL; CARBOLIC ACID; CARBOLIC-ACID;
        CEPASTAT; CHLORASEPTIC; CHLORASEPTTC; DOBELLS-SOLUTION; GREEN THROAT
        SPRAY; JABON-FENICADO; JASON-FENICADO; MENTOSEPTAN; MIKRO-BAC; NORAVERD;
        PAOSCLE; PHENASEPTIC; PHENOL; PHENOL-IN-OIL; PHOH; RED THROAT SPRAY;
        TRIAMINIC SORE THROAT SPRAY
```

収録内容や収録状況に変更点はない



分子式
標準化分子式
分子量
ダウエント化合物番号
ダウエント登録番号
入力日・更新日

```
MF      C6 H6 O
SMF     TYPE *1; C6 H6 O *1; TOTAL *1
MW      94.1124
SDCN    R00868
SDRN    0868
ED      Entered STN: 26 Apr 1999
        Last updated on STN: 30 Jan 2017
        Update DWPI Cross Ref.: 4 Aug 2022
```

主な検索フィールド	
/AN	DCR レコード番号
/CN	化学物質名称 (CN.P, CN.S, SY)
/CNS	化学物質名称セグメント
/MF	分子式
/ELC	元素数
/ELS	元素記号

© 2022 化学情報協会

DCR ファイルのハイライト 26

化学物質から特許情報を検索する方法が大きく変更！

DCR ファイルの L 番号を WPI ファイルで検索

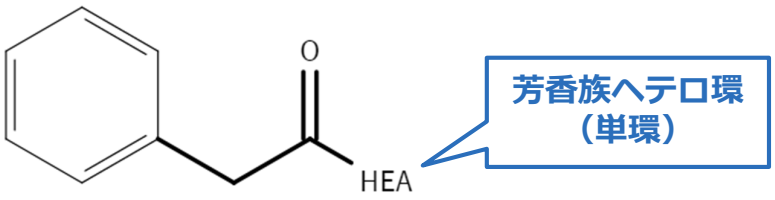
=> TRA L## DCR /AN

物質検索から文献を得る検索の流れが、すべて L 番号による
クロスオーバー検索に統一！

© 2022 化学情報協会

検索例2 27

下記の構造を持つ物質を合成している特許の検索



- **DCR ファイル**で部分構造検索
- WPI ファイルに**クロスオーバー検索**
- ロール（特許中の化学物質の役割を表すコード）を利用
  - **/RL フィールド**（WPI ファイル） New
    - ・合成のロール　　：**PRD** または **P**
    - ・新規物質のロール：**NEW** または **N**

© 2022 化学情報協会

検索例2

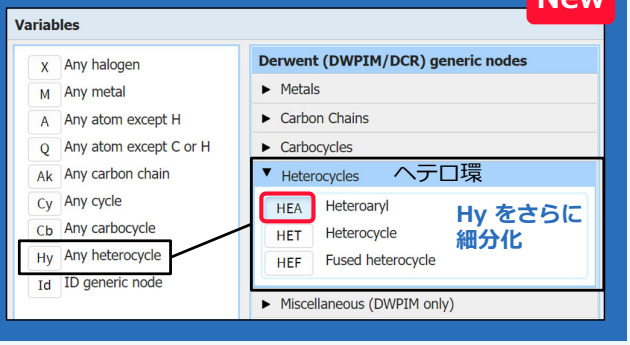

```
=> FILE DCR

=>
Uploading structure file: 2022_0020_Structure
```

DCR ファイルで構造検索

```
Node Attributes
Ring Nodes : 5  6  7  8  9  10
Chain Nodes : 1  2  4
Bond Attributes
Ring Bonds : 5-6  6-7  7-8  8-9  9-10  10-5
Chain Bonds : 1-2  1-5  2-3  2-4
Exact Bonds : 1-2  1-5
Normalized Bonds : 5-6  6-7  7-8  8-9  9-10  10-5
Exact/Normalized Bonds : 2-3  2-4
Markush Attributes
Match Level (ATOM) : 1  3  5  6  7  8  9  10
Match Level (CLASS) : 2  4
Element Count Level (LIMITED) : 1  2  3  4  5  6  7

L1      STRUCTURE UPLOADED
```

DCR/DWPIM 用**スーパーアトム**を利用できるように

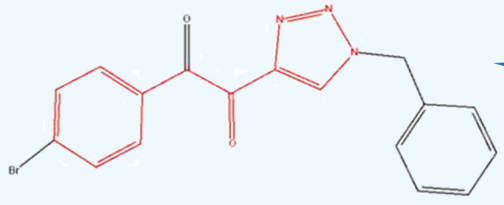
New



© 2022 化学情報協会

検索例2


```
=> S L1                          ← 部分構造検索（SSS）のサンプル検索（SAM）：SSS と SAM は省略可
 :
L2           45 SEA SSS SAM L1

=> D SAM 1-                      ← SAMPLE 表示形式でヒットした構造を確認

L2      ANSWER 1 OF 45 DCR  COPYRIGHT 2022 CLARIVATE on STN.
MF      C17 H12 Br N3 O2
```



ヒット部分がハイライト New

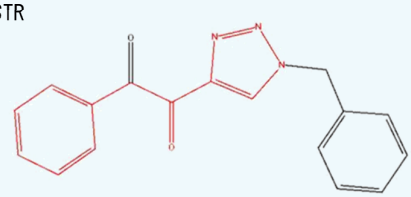
```
 :
=> S L1 FUL                      ← 部分構造検索のフルファイル検索（FUL）

FULL SEARCH INITIATED 15:30:48 FILE 'DCR'
FULL SCREEN SEARCH COMPLETED -  4512125 TO ITERATE
 :
L3           412 SEA SSS FUL L1
```

強化

検索範囲	回答件数の制限
サンプル検索	50
フルファイル検索	1,500,000

© 2022 化学情報協会

```
=> FILE WPINDEX                    ← WPI ファイルに入る

=> S L3 (T) (PRD OR P OR NEW OR N)/RL
L4        128 L3 (T) (PRD OR P OR NEW OR N)/RL

=> D BIB HITCODE HITSTR 1-
:
L4   ANSWER 2 OF 128  WPINDEX COPYRIGHT 2022   CLARIVATE
AN   2022-15956J [2022030]   WPINDEX Full-text
:
PI   CN 113912556    A  20220111 (2022030)* ZH
ADT  CN 113912556 A CN 2021-11375686 20211119
PRAI CN 2021-11375686      20211119
IT   UPIT 20220419
     DCR-5725479-EX DCR-5725479-PRD; DCR-5725480-EX DCR-5725480-PRD;
     DCR-5725481-EX DCR-5725481-PRD; DCR-148-CL DCR-148-RCT; DCR-129372-CL
:
AN.S DCR-5725479
CN.P 1-(1-benzyl-1H-1,2,3-triazol-4-yl)-2-phenylethane-1,2-dione
MF   C17 H13 N3 O2
STR
```



:

WPI ファイルへは L 番号でクロスオーバー！
WPI ファイルへクロスオーバー可能な DCR ファイルの回答件数は 20 万件

ロール検索をできる /RL が追加 **New**
化学物質とロールを組み合わせる場合は **(T) 演算子**を利用

HITSTR 表示形式を利用すると、WPI ファイル中でヒットした構造を確認できる

HITCODE

HITSTR

© 2022 化学情報協会

参考：WPI ファイルのロール

■ DCR レコード番号用ロール

ロール	定義	説明
CL	CLAIM	クレームに記載
EX	EXAMPLE	実施例に記載
DIS	DISCLOSURE	発明の詳細な説明に記載
NEW	NEW	新規の物質、方法、装置
PRD	PRODUCED	合成方法、製造方法
USE	USE	用途
DET	DETECTED	検出条件、検出された物質

ロール	定義	適用範囲
RCT	REACTANT	出発物質、中間体
RGT	REAGENT	試薬
CMP	COMPONENT	混合物の構成成分
PUR	PURIFIED	精製
REM	REMOVED	除去された物質、不純物
TES	TESTED	テストされた物質、機器
ST	SALT	塩

■ ダウエント化合物番号（DCN）用ロール

ロール	説明
A	分析、検出
C	触媒
D	検出剤
E	賦形剤
K	既知化合物
M	混合物の構成要素
N	新規の物質
P	既知物質の合成方法、製造方法

ロール	説明
Q	出発物質によって規定された製品
R	除去剤、精製剤
S	出発物質、中間体
T	治療活性物質、プロドラッグ
U	用途
V	試薬
X	除去された物質
Z	その他

SMARTracker が利用可能に 32

DCR ファイルから WPI ファイルへのクロスオーバー検索についてアラート（**SMARTracker**）が設定可能に

SMARTracker

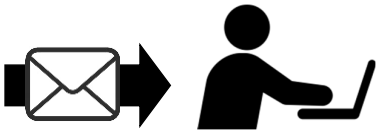
関心ある物質の最新文献情報を得る際に設定するアラート。物質データベースから文献データベースへのクロスオーバーまでの一連の検索について設定できる

=> **FILE DCR WPINDEX**

=> **SDI XFILE**

DCR と WPI の両ファイルに入る

定期的に最新情報がメール等で配信！



© 2022 化学情報協会

設定例（検索例２） 33

```
=> FILE DCR WPINDEX          ← DCR ファイルと WPI ファイル（マルチファイル）に入る
=> SDI XFILE

ENTER QUERY L# FOR SDI REQUEST OR (END):L4
ENTER UPDATE FIELD CODE (UP), ED, UPP, UPPA, UPIN, UPIT, UPCP, UPIC, UPFT, UPTI, UPAB, UPMC, UPB, UPA, UPAG,
UPAT, UPAA, UPCL, UPG, UPD OR ?:UPIT
ENTER SDI REQUEST NAME, (AA052/S), OR END:PARA/S
ENTER COST CENTER (NONE) OR NONE:.
ENTER TYPE OF SEARCH (SSS), CSS, FAMILY, OR EXACT:SSS
ENTER TITLE (NONE):PARA DCR
ENTER METHOD OF DELIVERY (EMAIL), ONLINE OR RSS:.
ENTER EMAIL ID (7152T):OOOO@jaici.or.jp
RECEIVE DELIVERY NOTIFICATION? Y/(N):N
ELIMINATE PREVIOUSLY SEEN ANSWERS WITH EACH SDI RUN?  Y/(N):Y
ENTER PRINT FORMAT (STD) OR ?:IBIB AB HITCODE HITSTR
HIGHLIGHT HIT TERMS? (Y)/N:.
ARCHIVE ANSWERS? Y/(N):.
REDISTRIBUTE ANSWERS? Y/(N):.
ENTER MAXIMUM NUMBER OF HITS TO BE DELIVERED PER RUN (100):.
SORT SDI ANSWER SET (N)/Y?:.
SEND SDI WITH NO ANSWERS? (Y)/N:.
ENTER SDI RUN FREQUENCY - WEEKLY, (EVERYUPDATE), MONTHLY, OR ?:WEEKLY
ENTER SDI EXPIRATION DATE 'YYYYMMDD' OR (NONE):.
QUERY L4 HAS BEEN SAVED AS SDI REQUEST 'PARA/S'
```

DCR ファイルからのクロスオーバー検索を含む L 番号

UPIT：索引情報の更新

構造検索のタイプを指定

表示形式を指定

実行頻度は毎週、更新毎、毎月から選択

© 2022 化学情報協会

その他の変更点 34

- 新規検索フィールドの追加

検索フィールド	内容
BI	基本索引　(CN, CMT, MF, SCT から切り出した語を含む)
FA	フィールドの存在

- 検索フィールドの変更

検索フィールド	内容
/NC	成分数（リロード前の /NFRAG は削除）
/NC.TOT	係数を含めた成分数（リロード前の /NC から変更）
/MF	分子式（スペースの有無にかかわらず検索できるように変更）

© 2022 化学情報協会

まとめ 35

WPI ファイルがリロードされ、

- 引用情報（旧DPCI ファイルの情報）が含まれるように
- 21 特許発行機関で全クレーム情報が収録されるように
- DCR ファイルが独立し、REGISTRY/CAplus ファイルと同様の仕組みで検索が可能に

© 2022 化学情報協会

CAS STNext ユーザーミーティング

配列ファイルのリロード

配列ファイルのリロード

化学情報協会　情報事業部

目次

1. 配列ファイル
2. リロードに伴う強化点
3. 検索例

© 2022 化学情報協会

配列ファイル

CAS STNext の配列関連のファイル 3

GENESEQ (旧 DGENE), USGENE, PATGENE (旧 PCTGEN) ファイルがリロードされた

ファイル名	内容	配列検索機能		
		完全配列/部分配列	BLAST ホモロジー	GETSIM ホモロジー
GENESEQ	世界の特許から抽出した配列を収録	○	○	○
USGENE	米国特許に記載された配列を収録			
PATGENE	PCT 出願に記載された配列を収録			
REGISTRY	世界の雑誌や特許から抽出した配列を収録	○	○	-
GenBank	米国国立衛生研究所作成の核酸配列データベース	-	-	-

リロード

機能	内容	配列検索機能
Biosequences Search	下記から配列を収録 • REGISTRY ファイル • 73 ヵ国の特許発行機関の特許 • NCBI 由来の配列	BLAST 配列検索 CDR 配列検索 Motif 配列検索

© 2022 化学情報協会

GENESEQ, USGENE, PATGENE ファイルの概要 4

(2022 年 8 月)

	GENESEQ (旧 DGENE)	USGENE	PATGENE (旧 PCTGEN)
製作者	Clarivate	SequenceBase Corp	FIZ Karlsruhe
収録源	WPI ファイルのベーシック特許	・ INSCD (核酸) ・ NCBNI/EMBL-EBI (タンパク質) ・ USPTO PSIPS 配列情報 ・ 米国特許の Sequence Listing	PCT 出願
レコード構成	配列単位	配列単位	配列単位
収録期間	1981 年～	1981 年～	2001 年 8 月～
収録件数　(合計)	5,900 万件以上	1 億件以上	2,500 万件以上
核酸	4,100 万件以上	7,300 万件以上	1,800 万件以上
タンパク質	1,800 万件以上	2,800 万件以上	720 万件以上
更新頻度	毎週	毎週	毎週
特長	GENESEQ 独自の抄録を作成しているので配列に関する説明情報が豊富	速報性に優れている (通常 3 日以内)	速報性に優れている (通常 1 日)

© 2022 化学情報協会

RUN コマンドを使用した配列検索 5

配列検索は RUN コマンドで実行する

BLAST ホモロジー検索　=> RUN BLAST　コード/検索フィールド パラメータ

GETSIM ホモロジー検索　=> RUN GETSIM コード/検索フィールド パラメータ

完全配列・部分配列検索　=> RUN GETSEQ コード/検索フィールド パラメータ

- コードは以下の順で入力する
  - 核酸：5′ 末端から 3′ 末端の順
  - タンパク質：N 末端 (NH_2) から C 末端 (COOH)
- パラメータは各ファイルの HELP OPTIONS および HELP GSEQ で確認できる

© 2022 化学情報協会

リロードに伴う強化点

ホモロジー検索のプログラムの強化 7

BLAST ホモロジー検索

- 最新版の BLAST プログラム Version 2.12.0 になり検索タイプが追加された

GETSIM ホモロジー検索

- 最新版の GETSIM プログラム Version 36.3.8h になり検索速度が大幅に改善された

© 2022 化学情報協会

BLAST ホモロジー検索の検索タイプの追加 8

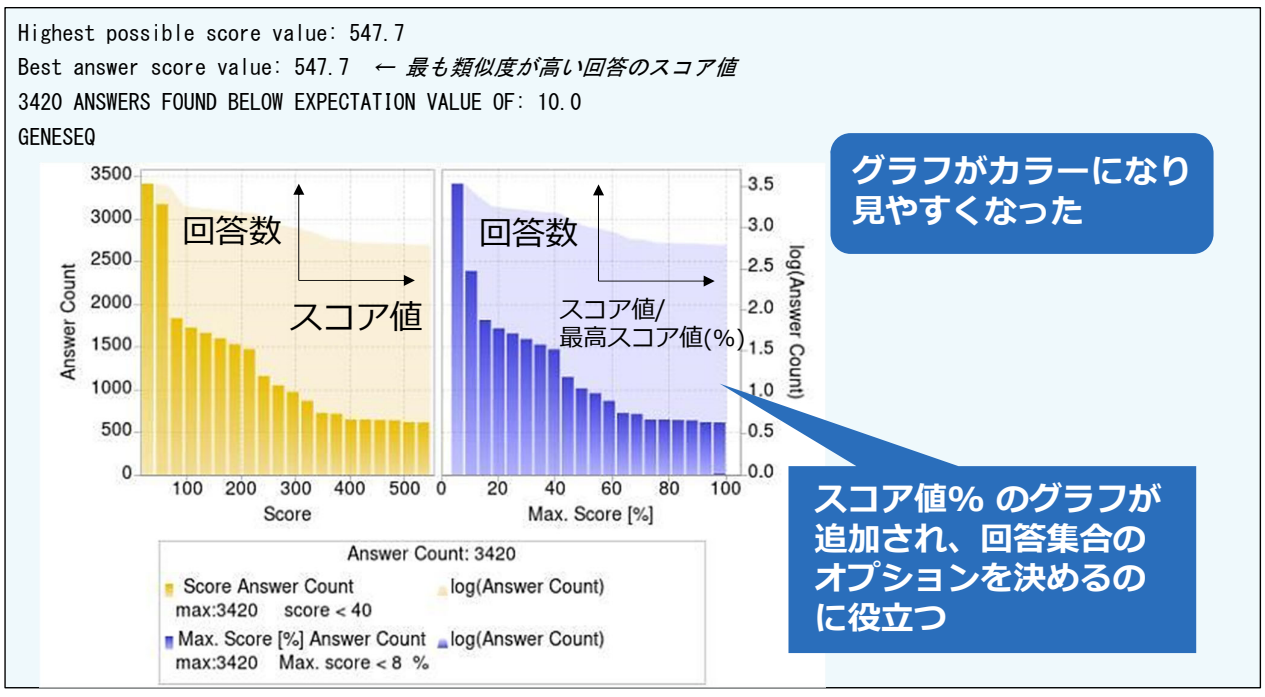
BLAST ホモロジー検索は検索タイプが 7 つに増加

検索フィールド	BLAST	GETSIM	内容	質問式	回答
/SQN	○	○	塩基配列の質問式に類似した塩基配列を検索	塩基配列	塩基配列
/TSQN	○	○	データベース中の塩基配列をアミノ酸に翻訳した配列の中からアミノ酸配列の質問式に類似した配列を検索	アミノ酸配列	塩基配列
/SQP	○	○	アミノ酸配列の質問式に類似したアミノ酸配列を検索	アミノ酸配列	アミノ酸配列
New /SQM	○	-	非常に類似した（種内などの）配列用に最適化された BLASTn (megaBLAST)	塩基配列	塩基配列
/SQDM	○	-	一部の塩基を無視し（多少のミスマッチを許容し），より離れた（種間などの）配列を検索するために最適化された BLASTn (discontiguous megaBLAST)	塩基配列	塩基配列
/TSQP	○	-	塩基配列の質問式をアミノ酸配列に翻訳してこれに類似したアミノ酸配列を検索 (BLASTx)	塩基配列	アミノ酸配列
/TSQNX	○	-	塩基配列の質問式をアミノ酸配列に翻訳してこれに類似したアミノ酸配列に翻訳された塩基配列を検索 (tBLASTx)	塩基配列	塩基配列

© 2022 化学情報協会

ホモロジー検索のワークフローの変更 9

ワークフローがより便利になり、様々な観点で回答集合を作成可能に！



© 2022 化学情報協会

ホモロジー検索のワークフローの変更 (続き)


```
Query time:          681
① EITHER "ALL" TO KEEP ALL ANSWERS
② OR ENTER THE MINIMUM SCORE VALUE YOU WISH TO KEEP
③ OR ENTER THE MINIMUM PERCENT OF SCORE FOLLOWED BY "% SCORE"
④ OR ENTER THE MINIMUM PERCENT OF IDENTITY FOLLOWED BY "% IDENT"
⑤ OR COMBINE MINIMUM PERCENT OF SCORE AND IDENTITY AS "X% SCORE Y% IDENT"
⑥ OR ENTER "END". "END" MUST BE ENTERED TO COMPLETE THE RUN COMMAND.

ENTER (ALL) OR ?:ALL

L2     RUN STATEMENT CREATED
L2         3420 GCTCCCAGAATGCCAGAGGCT
                TCCTACACCGGCGGCCCCTGC
                CTGTCCCTTCCCAGAAAACCT
                TTCTTGCATTCTGGGACAGCC
                CCTCAACAAGATGTTTTGCCA
                GGGTTGATTCCACACCCCCGC
                TAC/SQN -F F

① ENTER EITHER "ALL" TO KEEP ALL ANSWE
② OR ENTER THE MINIMUM SCORE VALUE YOU WISH TO KEEP
③ OR ENTER THE MINIMUM PERCENT OF SCORE FOLLOWED BY "% SCORE"
④ OR ENTER THE MINIMUM PERCENT OF IDENTITY FOLLOWED BY "% IDENT
⑤ OR COMBINE MINIMUM PERCENT OF SCORE AND IDENTITY AS "X% SCORE
⑥ OR ENTER "END". "END" MUST BE ENTERED TO COMPLETE THE RUN COM
ENTER (ALL) OR ?:
```

回答集合のオプションを選択する

	内容	入力例
①	回答全件 (ALL と入力)	ALL
②	入手したいスコア値の最低値 (数字を入力)	300
③	入手したいスコア値パーセント (スコア値/最高スコア値) の最低値	85% または 85% SCORE
④	入手したい同一性パーセント (一致コード数/Alignment コード数) の最低値	100% IDENT
⑤	入手したいスコア値パーセントの最低値と同一性パーセントの最低値	85% SCORE 100% IDENT
⑥	終了時 (END を入力)	END

回答集合のオプションを変えて、何度でも回答集合を作成できる

© 2022 化学情報協会

回答件数の上限の緩和

ホモロジー検索において回答上限数が増加

	リロード前	リロード後
BLAST ホモロジー検索 GETSIM ホモロジー検索	10,000 件	15,000件 (デフォルト) ● パラメータ設定で最大 100,000 件まで変更可能
完全配列検索 部分配列検索	250,000 件 (L 番号は 25,000 件ずつ作成されるので最後に複数の L 番号をまとめる)	250,000 件 (1 つの L 番号にまとまる)

- ホモロジー検索で回答数の上限を変更するには、-MAXSEQ パラメータを利用する。

  => RUN BLAST L1/SQN **–MAXSEQ 100000**

この設定で最大 10 万件に!

© 2022 化学情報協会

配列質問式の長さの制限値緩和 12

- ホモロジー検索

検索フィールド	BLAST		GETSIM	
	最小	最大	最小	最大
/SQN	9	30,000	9	30,000
/TSQN	9	20,000	3	10,000
/SQP	6	30,000	4	30,000
/SQM	28	30,000	-	
/SQDM	36	30,000	-	
/TSQP	18	30,000	-	
/TSQNX	18	5,000	-	

リロードにより、最大値が 1 万コードから 3 万コードに増加

- 完全配列検索、部分配列検索

検索フィールド	最小	最大
/SQEN	2	10,000
/SQSN	8	10,000
/SQEP	2	10,000
/SQEFP	2	10,000
/SQSP	3	10,000
/SQSFP	5	10,000

リロードにより、最大値が 2000 コードから 1 万コードに増加

© 2022 化学情報協会

相補鎖検索の強化 13

核酸検索で相補鎖が自動的に含まれるようになった

- 相補鎖を含めるかどうかはパラメータで変更可能

=>RUN GETSIM コード/SQN ← 入力したコードと相補鎖を検索

相補鎖のコードを別途検索する必要がなくなった！

=>RUN GETSIM コード/SQN **–S SIN** ← 入力したコードのみ検索

-S オプションで指定

オプション	定義	BLAST 検索タイプ	GETSIM 検索タイプ	完全配列検索、部分配列検索
SIN	入力した配列コードのみ検索	/SQN, /SQM, /SQDM, /TSQNX	/SQN	/SQEN, /SQSN
COM	入力した配列コードの相補鎖のみを検索			
BOTH（デフォルト）	入力した配列コードとその相補鎖の両方を検索			

© 2022 化学情報協会

核酸表、アミノ酸表の追加 14

核酸またはアミノ酸の種類、数、比率 (%) を示す表が ALL および SQIDE 表示形式に追加

- 核酸

```
SEQ
  1 gcaggtatta ccctgggtat ggatgaataa
NA   ← 核酸表
        Code     Count  Percent
           A         9     30.0
           C         4     13.3
           G         9     30.0
           T         8     26.7   (S)
           U         0      0.0
       Other         0      0.0
```

検索フィールド	内容
NA	核酸の種類
NA.CNT	核酸の数
NA.PER	核酸の比率

(S) 演算子で同一コードの情報に限定
例) T が 20-30% 含まれる
=>S T/NA(S)20-30/NA.PER

- アミノ酸

```
SEQ
  1 hsqgtftsdk seyldserar dfvawleagg
AA   ← アミノ酸表
        Code     Count  Percent
           A         3     10.0
           R         2      6.7   (S)
           N         0      0.0
           D         3     10.0
           B         0      0.0
           :
           V         1      3.3
      Others         0      0.0
```

検索フィールド	内容
AA	アミノ酸の種類
AA.CNT	アミノ酸の数
AA.PER	アミノ酸の比率

© 2022 化学情報協会

配列キーの追加 15

配列を SHA-2 アルゴリズムによりコード化した配列キー (SEQK) の収録を開始

```
=> FILE GENESEQ          ← GENESEQ ファイルに入る

=> RUN GETSEQ GACATGCGGAAGCACGTGGCCAT/SQEN      ← 完全配列検索
GENESEQ
Query time:          323
L1     RUN STATEMENT CREATED
L1           3 GACATGCGGAAGCACGTGGCCAT/SQEN

=> D SQIDE               ← SQIDE 表示形式で表示

L1      ANSWER 1 OF 3 GENESEQ COPYRIGHT 2022 CLARIVATE on
AN      BHT33478    GENESEQ  ED 20211030 UP 20211030
            :
SEQK    812c12a2e0fe13dad3e72bc50e3f764cee046b0efbdb4cff8456dd3e2ee29d2f
SEQ
                1 gacatgcgga agcacgtggc cat

=> FILE PATGENE          ← PATGENE ファイルに入る

=> S 812c12a2e0fe13dad3e72bc50e3f764cee046b0efbdb4cff8456dd3e2ee29d2f/SEQK
L2          2 812C12A2E0FE13DAD3E72BC50E3F764CEE046B0EFBDB4CFF8456DD3E2EE29D2F/SEQK
```

配列キー

- **GENESEQ,USGENE,PATGENE ファイルにおいて同一配列には同一の配列キーを収録している**
- **どの生物に由来する配列であっても、同じ配列であれば同一の文字列で表現される**

完全配列検索と同じ結果が得られる

© 2022 化学情報協会

その他の新規検索フィールド 16

検索フィールド	内容	入力例
GENESEQ, USGENE, PATGENE		
/APO	出願番号、オリジナル	S WOBE000003/APO
/DED	データ入力日	S 20190307/DED
/DUPD*	データ更新日	S 20190307/DUPD
/PNO	特許番号、オリジナル	S WO200206834/PNO
/PRDF	最先の優先権主張日	S 20150608/PRDF
/PRNO	優先権出版番号、オリジナル	S DE04447388/PRNO
USGENE		
/INA	発明者住所	S LONDON/INA
/PAA	特許出願人住所	S NEW YORK/PAA
/RLPC	関連出願の特許発行国	S WO/RLPC
/RLPD	関連出願の特許発行日	S 20140116/RLPD
/RLPN	関連出願の特許番号	S WO2014001422/RLPN
/RLPY	関連出願の特許発行年	S 2015/RLPY
/RLT	関連出願のタイプ	S EARLIER APPLIATION/RLT

* GENESEQ, PATGENE のみ利用可能

© 2022 化学情報協会

検索例

調査テーマ 18

関節リウマチ治療薬であるアダリムマブの重鎖と類似している配列を調べる

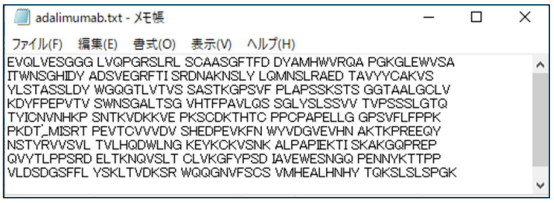
```
EVQLVESGGG LVQPGRSLRL SCAASGFTFD DYAMHWVRQA PGKGLEWVSA
ITWNSGHIDY ADSVEGRFTI SRDNAKNSLY LQMNSLRAED TAVYYCAKVS
YLSTASSLDY WGQGTLVTVS SASTKGPSVF PLAPSSKSTS GGTAALGCLV
KDYFPEPVTV SWNSGALTSG VHTFPAVLQS SGLYSLSSVV TVPSSSLGTQ
TYICNVNHKP SNTKVDKKVE PKSCDKTHTC PPCPAPELLG GPSVFLFPPK
PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY
NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP
QVYTLPPSRD ELTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP
VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPGK
```

- 検索条件
  - GENESEQ ファイルで BLAST ホモロジー検索を行う
    - 直接入力できるコードは 278 コードまでのため、長い配列質問式はアップロード機能を利用する
  - 抗体のキーワードで限定
  - クレームされている配列に限定

© 2022 化学情報協会

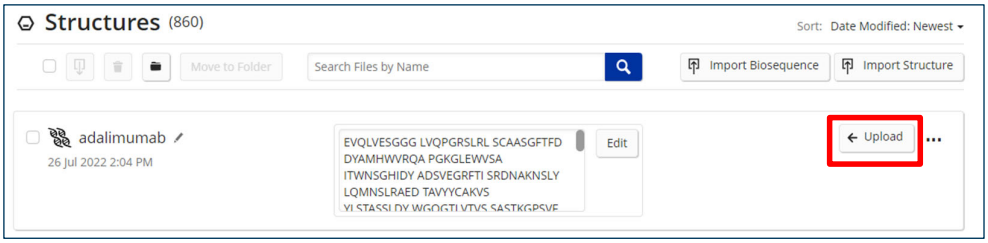
配列質問式のアップロード 19

1. 配列質問式のテキストファイルを作成



2. GENESEQ ファイルに入り、My Files の Structures ページで Import Biosequence をクリックし、上記 1 のファイルをインポート

3. Upload をクリックする



© 2022 化学情報協会

BLAST ホモロジー検索


```
=> FILE GENESEQ

=>
Uploading sequence file: adalimumab
UPLOAD R BLAST

UPLOAD SUCCESSFULLY COMPLETED
L1    GENERATED                    ← アップロードした配列質問式の L 番号が表示される

=> D LQUE                          ← LQUE 表示形式でアップロードした配列を確認できる

L1      ANSWER 1 GENESEQ  COPYRIGHT 2022 CLARIVATE on STN.
LQUE    EVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSAITWNSGHIDYADSVEGRFTISR
        DNAKNSLYLQMNSLRAEDTAVYYCAKVSYLSTASSLDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTA
        ALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVD
        KKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV
        HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRD
        ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVM
        HEALHNHYTQKSLSLSPGK

=> RUN BLAST L1/SQP -F F           ← フィルターを外して BLAST ホモロジー検索を実行する

Algorithm: BLAST - BLAST
      :
```

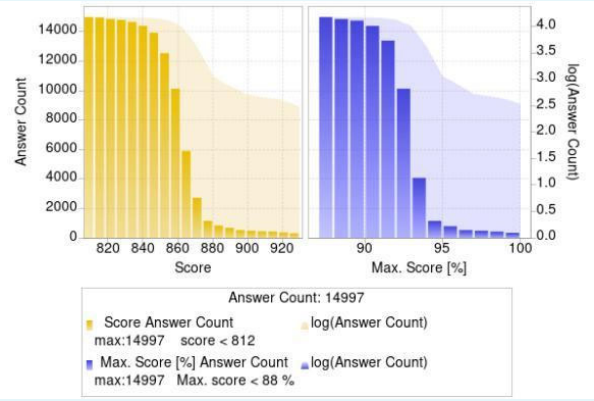
低複雑度領域フィルタ

低複雑度領域のマスクフィルタリングが行われ、生物学的に無意味なアライメントは取り除かれる。特許性調査の場合は -F F でフィルタをはずした方がよい

© 2022 化学情報協会

検索例


```
Highest possible score value: 929.9
Best answer score value: 929.9       ← 最も類似度が高い回答のスコア値

14997 ANSWERS FOUND BELOW EXPECTATION VALUE OF: 10.0        ← 回答件数
GENESEQ
```




```
Query time:          4108
ENTER EITHER "ALL" TO KEEP ALL ANSWERS
OR ENTER THE MINIMUM SCORE VALUE YOU WISH TO KEEP
OR ENTER THE MINIMUM PERCENT OF SCORE FOLLOWED BY "% SCORE"
OR ENTER THE MINIMUM PERCENT OF IDENTITY FOLLOWED BY "% IDENT"
OR COMBINE MINIMUM PERCENT OF SCORE AND IDENTITY AS "X% SCORE Y% IDENT"
OR ENTER "END". "END" MUST BE EN
ENTER (ALL) OR ?:ALL
```

全件を得るには ALL を入力

© 2022 化学情報協会

検索例

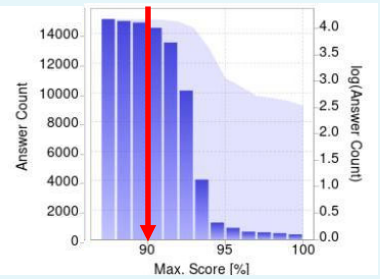

```
L2    RUN STATEMENT CREATED
L2       14997 EVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSA
                    :
                VLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK/SQP -F F

ENTER EITHER "ALL" TO KEEP ALL ANSWERS
OR ENTER THE MINIMUM SCORE VALUE YOU WISH TO KEEP
OR ENTER THE MINIMUM PERCENT OF SCORE FOLLOWED BY "% SCORE"
            :
ENTER (ALL) OR ?:90%

L3    RUN STATEMENT CREAT
L3       14611 EVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSA
                VLD                                QKSLSLSPGK/SQP -F F

ENTER EITHER "ALL" TO KEEP ALL ANSWERS
OR ENTER THE MINIMUM SCORE VALUE YOU WISH TO KEEP
OR ENTER THE MINIMUM PERCENT OF SCORE FOLLOWED BY "% SCORE"
 :
ENTER (ALL) OR ?:END
```

任意

回答集合のオプションは複数回指定できる
例) スコア値が 90% 以上の回答のみを得る

新たに別の L 番号が作成される

終了するには END を入力



© 2022 化学情報協会

検索例 – SORT を利用した回答の並び替え


```
=> S L2 AND ANTIBOD?                ← 抗体のキーワードで限定
L4       14863 L2 AND ANTIBOD?

=> S L4 AND CLAIM?/PSL              ← クレームされている配列に限定
L5        5574 L4 AND CLAIM?/PSL

=> SORT L5 1- SCORE D IDENT D
L6        5574 SORT L5 1- SCORE D IDENT D

=> FSORT L6
 :
L7         SEL L6 1- PN APPS :    4803 TERMS

'L7' DELETED
L7        5574 FSO L6

           668 Multi-record Families   Answers 1-4892      ← 関連特許ファミリー
                 Family 1                Answers 1-4
                 Family 2                Answers 5-15
                    :
                 Family 668              Answers 4889-4892
           682 Individual Records      Answers 4893-5574   ← 関連特許ファミリーがないレコード
             0 Non-patent Records
```

SORT コマンドを利用してスコア値 (SCORE) の降順 (D) かつ同一性 (IDENT) の降順 (D) で回答を並び替える

特許情報の重複表示を回避したい場合は、FSORT コマンドを使用して関連特許ファミリーごとにまとめる

FSORT コマンド
回答から PN (特許番号), AP (出願番号), PRN (優先権出願番号) を抽出して関連特許ファミリーごとにまとめる

© 2022 化学情報協会

検索例 - 回答の表示 (BIB AB)


```
=> D L7 4889 BIB AB ALIGN            ← BIB AB ALIGN 表示形式で表示

L7     ANSWER 4889 OF 5574 GENESEQ COPYRIGHT 2022 CLARIVATE on STN. FAMILY 668
AN     AAM52156   GENESEQ  ED 20211030 UP 20211030
       DED 20020205 Full-text
TI     Novel compound used to treat cancer has target cell-specific portion comprising
       humanized monoclonal antibody having specificity for polymorphic epithelial mucin,
       and cytotoxic portion having endonucleolytic activity.
IN     Young RJ
PA     ANTISOMA RES LTD (ANTI-N)
LA     English
DT     Patent
PI     WO 2001074905   A1        20011011
PIT    WOA1 INTERNATIONAL APPLICATION PUBLISHED WITH INTERNATIONAL SEARCH REPORT
AI     WO 2001-GB1324            20010326
PRAI   GB 2000-8049              20000403
       US 2000-237159P           20001002
FS     PROTEIN; PS
OS     2001-662969 [76]                                    ← WPI のレコード番号
MTY    protein                                             ← 分子タイプ
PSL    Claim 20; Fig 7; 176pp                              ← 配列情報の記載位置
DESC   Humanised HMFG-1 heavy chain/DNase I fusion protein 1.   ← 配列の説明
AB     The invention relates to a compound which comprises a target cell-
       specific portion, comprising an humanised monoclonal antibody, having
       specificity for polymorphic epithelial mucin (PEM) or its antigen
            :
```

配列独自の抄録

© 2022 化学情報協会

検索例 - 回答の表示 (ALIGN)


```
ALIGN
  ALIGNMENT FROM L-NUMBER L2
  Query Length: 451; Sequence Length: 731;                ← 配列質問式の長さ; 回答配列の配列長
  Score: 814.7 bits (2103), 87.6% of highest possible score 929.9;   ← スコア値
  Expect value: 2.341e-234;                               ← 期待値
  Identities: 396 / 451 (87.8%); Positives: 419 / 451 (92.9%);   ← 同一性パーセント
  Query Identity: 87.8%; Query Coverage: 100.0%;
  Subject Identity: 54.2%; Subject Coverage: 61.7%;
  Alignment Length: 451;
  Q:   1 EVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSAITWNSGHIDY 60
         +||||+||  + +|| |+++|| |||+||  | + |||||||||||||  |   | +  |
  S:  20 QVQLVQSGAEVKKPGASVKVSCKASGYTFSAYWIEWVRQAPGKGLEWVGEILPGSNNSRY 79

  Q:  61 ADSVEGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCAKVSYLSTASSLDYWGQGTLVTVS 120
          +  +|| |++|| + |+ |++++|||+|||||||||+ ||    +   ||||||||||
  S:  80 NEKFKGRVTVTRDTSTNTAYMELSSLRSEDTAVYYCAR-SY--DFAWFAYWGQGTLVTVS 136

  Q: 121 SASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  S: 137 SASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS 196
                                    :
  Q: 421 WQQGNVFSCSVMHEALHNHYTQKSLSLSPGK                          451
         ||||||||||||||||||||||||||||||
  S: 437 WQQGNVFSCSVMHEALHNHYTQKSLSLSPGK                          467
```

ヒットした回答配列の配列長は731 コードであり、ヒットは20～467 コードの位置で一致している

アライメント表示
Q (配列質問式：Query)
S (回答の配列：Subject)
| は一致しているコード
+ はファミリーで一致しているコード

© 2022 化学情報協会

まとめ

GENESEQ, USGENE, PATGENE ファイルが
リロードされ、

- ホモロジー検索では何度もオプションを変えて
  複数の回答集合を作成できる
- ホモロジー検索の回答件数の上限がリロード前に
  比べ最大 10 倍になった
- 核酸検索ではデフォルトで相補鎖が含まれる

© 2022 化学情報協会

CAS STNext ユーザーミーティング

CAS STNext の強化

CAS STNext の強化

化学情報協会　情報事業部

目次

1. 検索記録の強化
2. 構造作図に関する強化
3. CAS SciFinder^{n} との連携
4. その他の強化

© 2022 化学情報協会

検索記録の強化

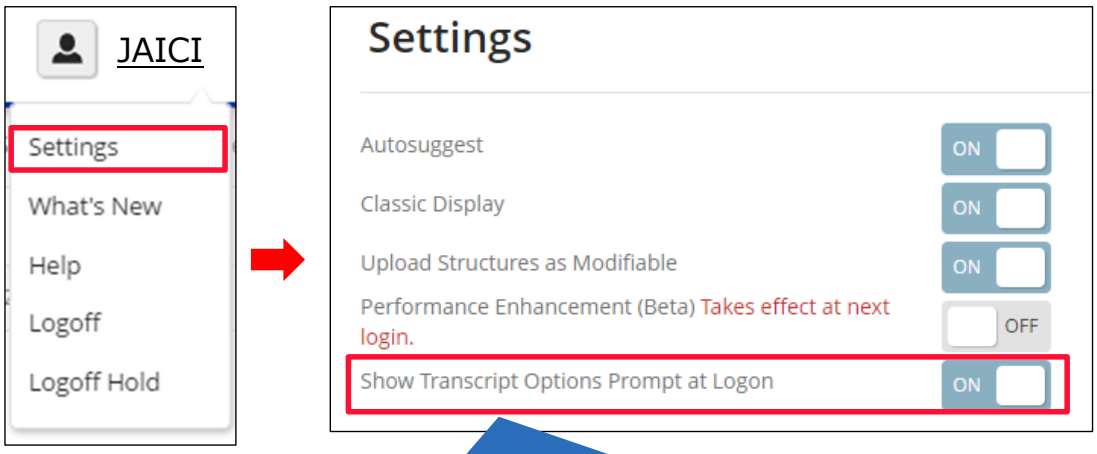
検索記録の強化内容 3

- セッション開始時に Transcript Options を表示可能に
- 検索記録のヒットタームハイライト色を選択可能に
- レポート作成時にハイライト色などを選択可能に
- Excel 形式テーブルのダウンロードオプション追加
- レポートを XML, BizInt 形式でダウンロード可能に

© 2022 化学情報協会

セッション開始時に Transcript Options を表示可能に 4

Settings で表示の ON/OFF を選択できる

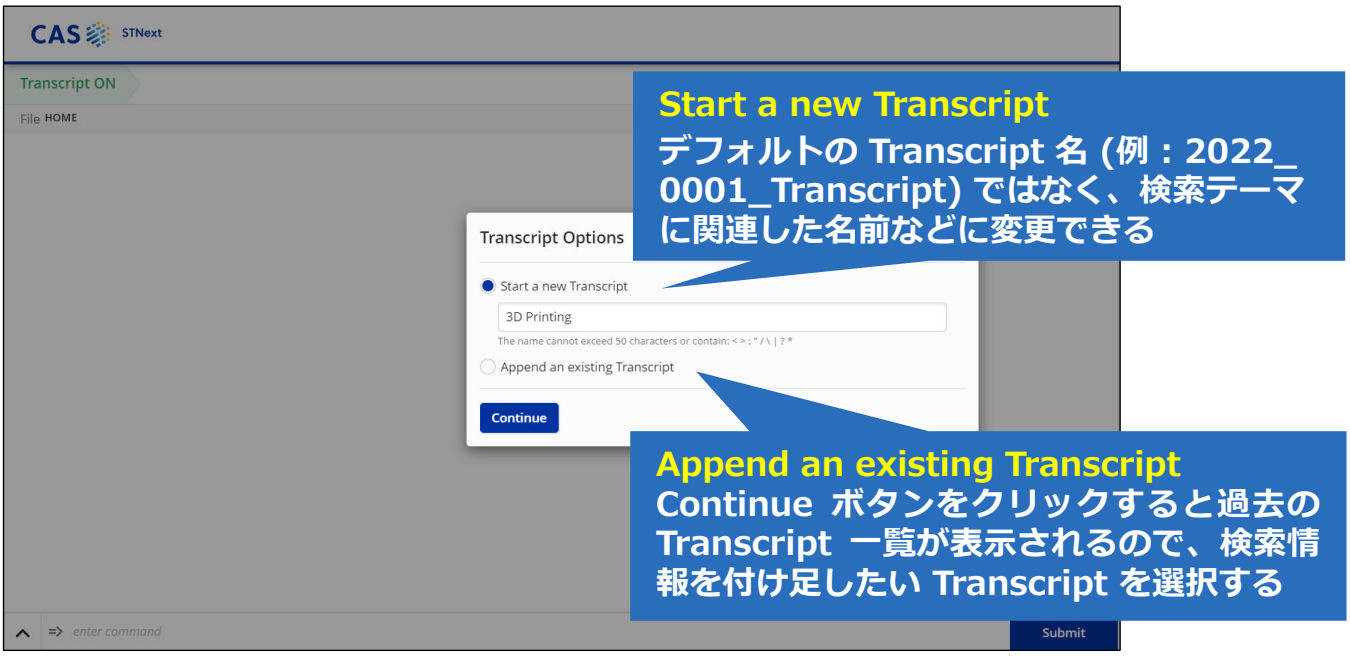


セッション開始時に Transcript (検索記録) Options 画面を表示する

© 2022 化学情報協会

セッション開始時に Transcript Options を表示可能に 5

分かりやすい Transcript 名に変更してから検索開始



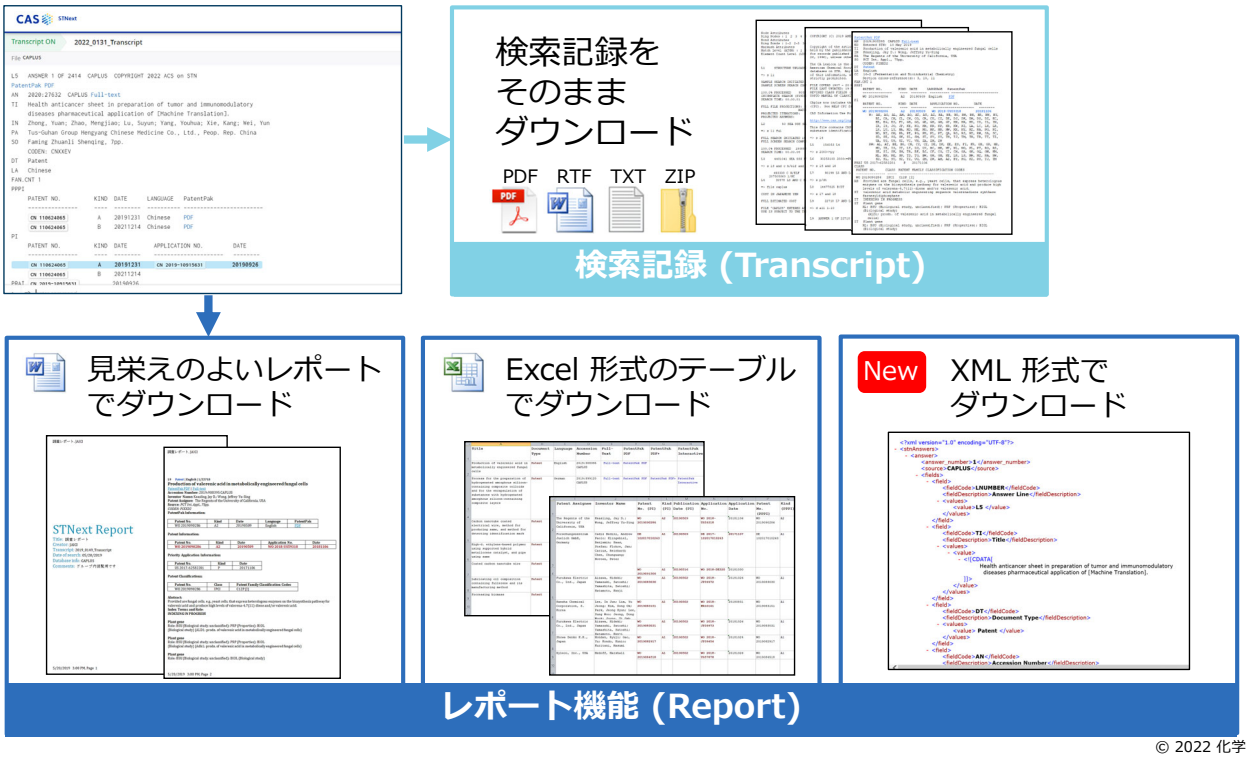
Start a new Transcript
デフォルトの Transcript 名 (例：2022_0001_Transcript) ではなく、検索テーマに関連した名前などに変更できる

Append an existing Transcript
Continue ボタンをクリックすると過去の Transcript 一覧が表示されるので、検索情報を付け足したい Transcript を選択する

© 2022 化学情報協会

検索記録のダウンロード 6

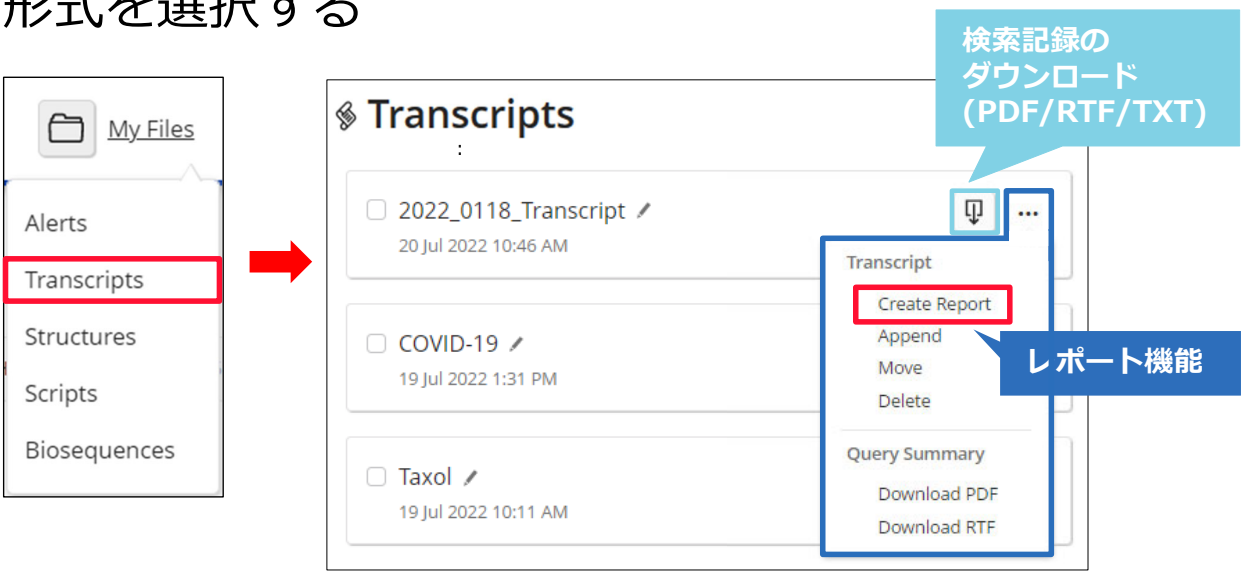
出力した検索記録は様々な形式でダウンロードできる



© 2022 化学情報協会

検索記録のダウンロード方法 7

My Files の Transcripts からダウンロードしたい形式を選択する

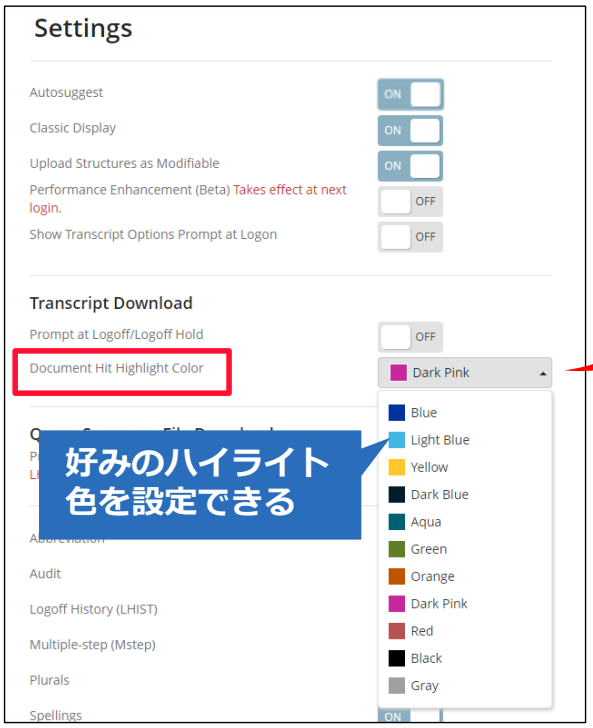


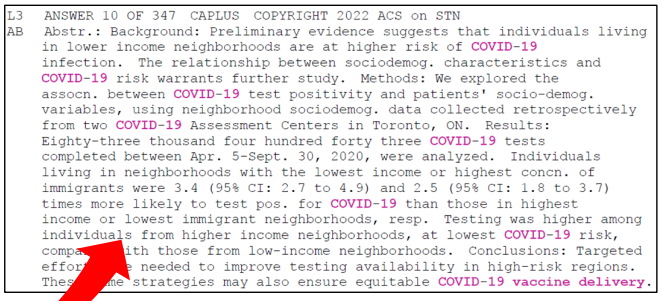
＊ レポート機能の詳細は「CAS STNext 利用ガイド」参照
https://www.jaici.or.jp/stn-ip-protection-suite/cas-stnext/documents/

© 2022 化学情報協会

検索記録のヒットタームハイライト色を選択可能に 8

ハイライト色は Settings で設定する



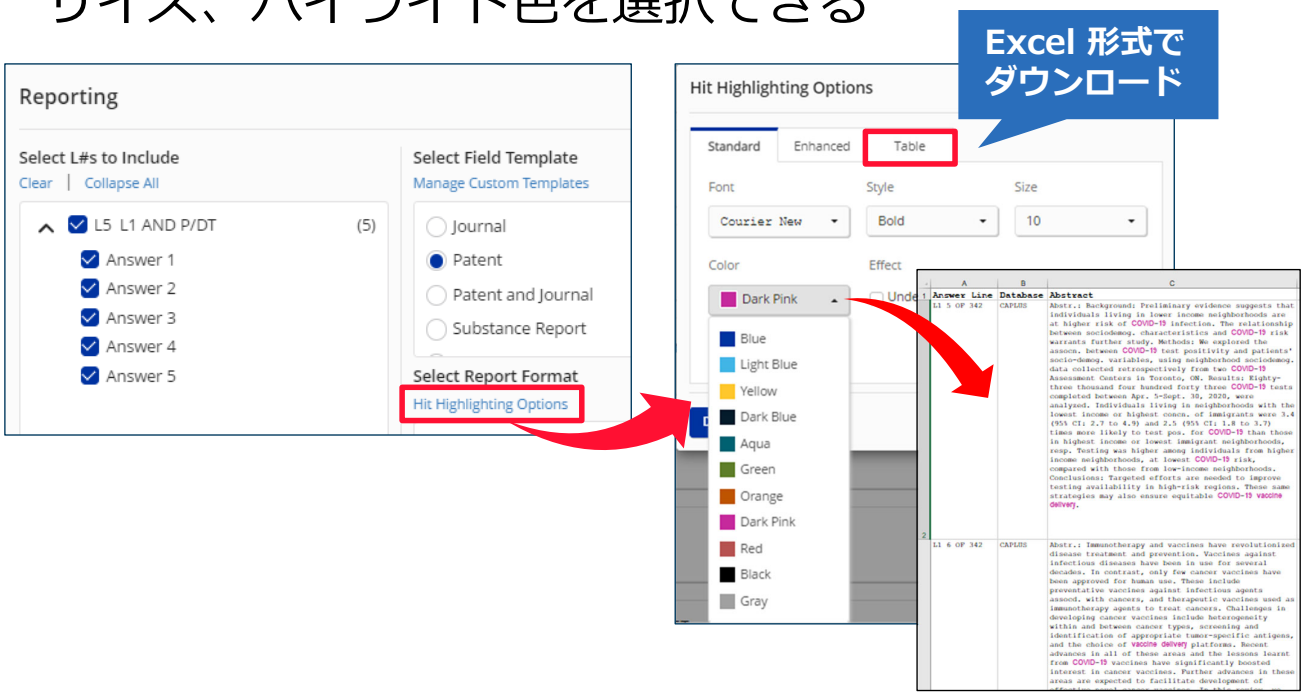


L3 ANSWER 10 OF 347 CAPLUS COPYRIGHT 2022 ACS on STN
AB Abstr.: Background: Preliminary evidence suggests that individuals living in lower income neighborhoods are at higher risk of COVID-19 infection. The relationship between sociodemog. characteristics and COVID-19 risk warrants further study. Methods: We explored the assocn. between COVID-19 test positivity and patients' socio-demog. variables, using neighborhood sociodemog. data collected retrospectively from two COVID-19 Assessment Centers in Toronto, ON. Results: Eighty-three thousand four hundred forty three COVID-19 tests completed between Apr. 5-Sept. 30, 2020, were analyzed. Individuals living in neighborhoods with the lowest income or highest concn. of immigrants were 3.4 (95% CI: 2.7 to 4.9) and 2.5 (95% CI: 1.8 to 3.7) times more likely to test pos. for COVID-19 than those in highest income or lowest immigrant neighborhoods, resp. Testing was higher among individuals from higher income neighborhoods, at lowest COVID-19 risk, compa[illegible]ith those from low-income neighborhoods. Conclusions: Targeted effor[illegible]e needed to improve testing availability in high-risk regions. Thes[illegible]me strategies may also ensure equitable COVID-19 vaccine delivery.

© 2022 化学情報協会

レポート作成時にハイライト色などを選択可能に 9

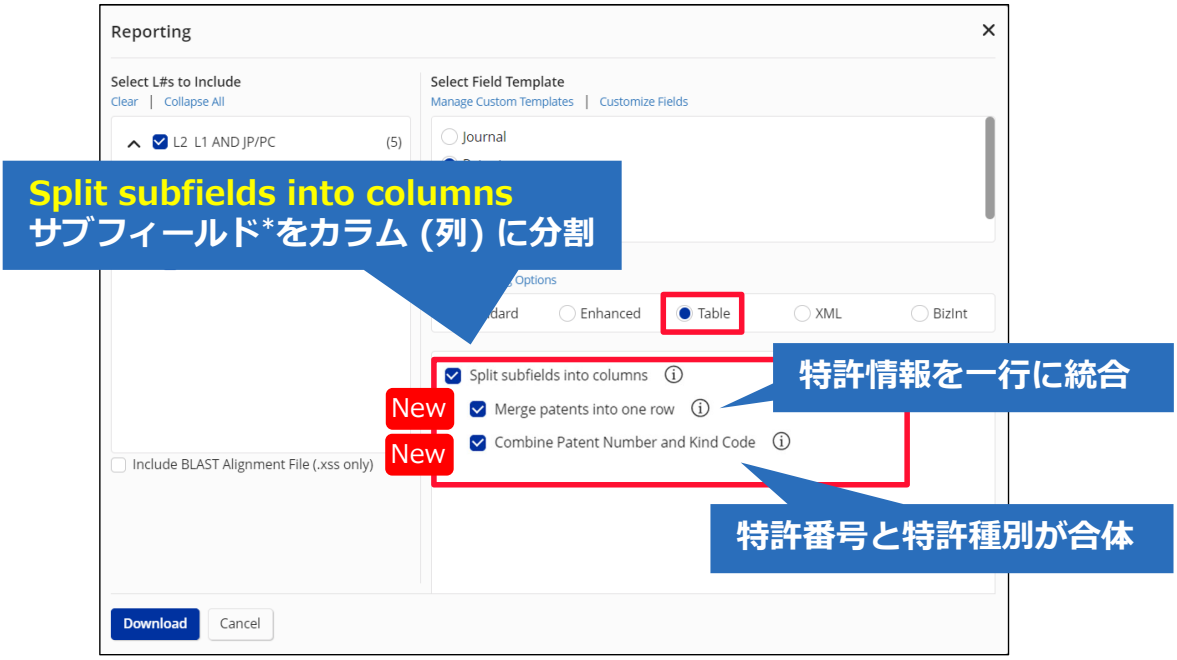
レポートのヒットタームのフォント、スタイル、サイズ、ハイライト色を選択できる



© 2022 化学情報協会

Excel 形式テーブルのダウンロードのオプション追加　10

Merge Patents into one row と Combine Patent Number and Kind Code が追加された

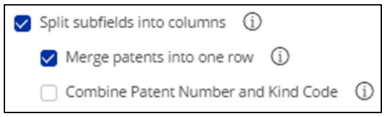


* PI フィールドの PN (特許番号), PK (特許種別), PD (特許発行日) など

© 2022 化学情報協会

Excel 形式テーブルのダウンロードのオプション追加　11

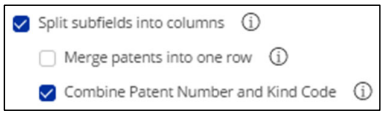
1. Merge Patents into one row のみチェックを入れる



Patent No. (PI)	Kind (PI)	Publication Date (PI)	Application No. (PI)	Application Date (PI)	Priority No.	Kind (PRAI)	Priority Date
CN 114767707	A	20220722	CN 2021-11437836	20211129	KR 2021-9584	A	20210122
KR 2022106912	A	20220801	KR 2021-119259	20210907	KR 2021-119259	A	20210907

特許情報を一行に統合

2. Combine Patent Number and Kind Code のみチェックを入れる



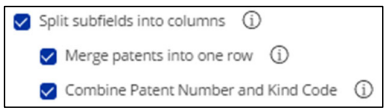
Patent No./Kind (PI)	Publication Date (PI)	Application No. (PI)	Application Date (PI)	Priority No.	Kind (PRAI)	Priority Date
CN 114767707 A	20220722	CN 2021-11437836	20211129	KR 2021-9584	A	20210122
KR 2022106912 A	20220801	KR 2021-119259	20210907	KR 2021-119259	A	20210907

特許番号と特許種別が合体

© 2022 化学情報協会

Excel 形式テーブルのダウンロードのオプション追加　12

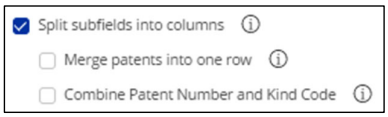
3. 両方にチェックを入れる



Patent No./Kind (PI)	Publication Date (PI)	Application No. (PI)	Application Date (PI)	Priority No.	Kind (PRAI)	Priority Date
CN 114767707 A KR 2022106912 A	20220722 20220801	CN 2021-11437836 KR 2021-119259	20211129 20210907	KR 2021-9584 KR 2021-119259	A A	20210122 20210907

特許情報を一行に統合
特許番号と特許種別が合体

4. 両方にチェックを入れない

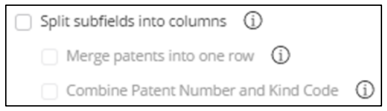


Patent No. (PI)	Kind (PI)	Publication Date (PI)	Application No. (PI)	Application Date (PI)	Priority No.	Kind (PRAI)	Priority Date
CN 114767707	A	20220722	CN 2021-11437836	20211129	KR 2021-9584	A	20210122
KR 2022106912	A	20220801	KR 2021-119259	20210907	KR 2021-119259	A	20210907

© 2022 化学情報協会

参考：特許情報と出願情報を一つのセルに入れられる　13

Split subfields into columns のチェックを入れずにテーブルを作成



Patent Information					Priority Application Information		
PATENT NO.	KIND	DATE	APPLICATION NO.	DATE	JP 2021-12663	A	20210129
US 20220242916	A1	20220804	US 2022-17578049	20220118			
JP 2022116482	A	20220810	JP 2021-12663	20210129			
PATENT NO.	KIND	DATE	APPLICATION NO.	DATE	KR 2021-9584	A	20210122
CN 114767707	A	20220722	CN 2021-11437836	20211129	KR 2021-119259	A	20210907
KR 2022106912	A	20220801	KR 2021-119259	20210907			

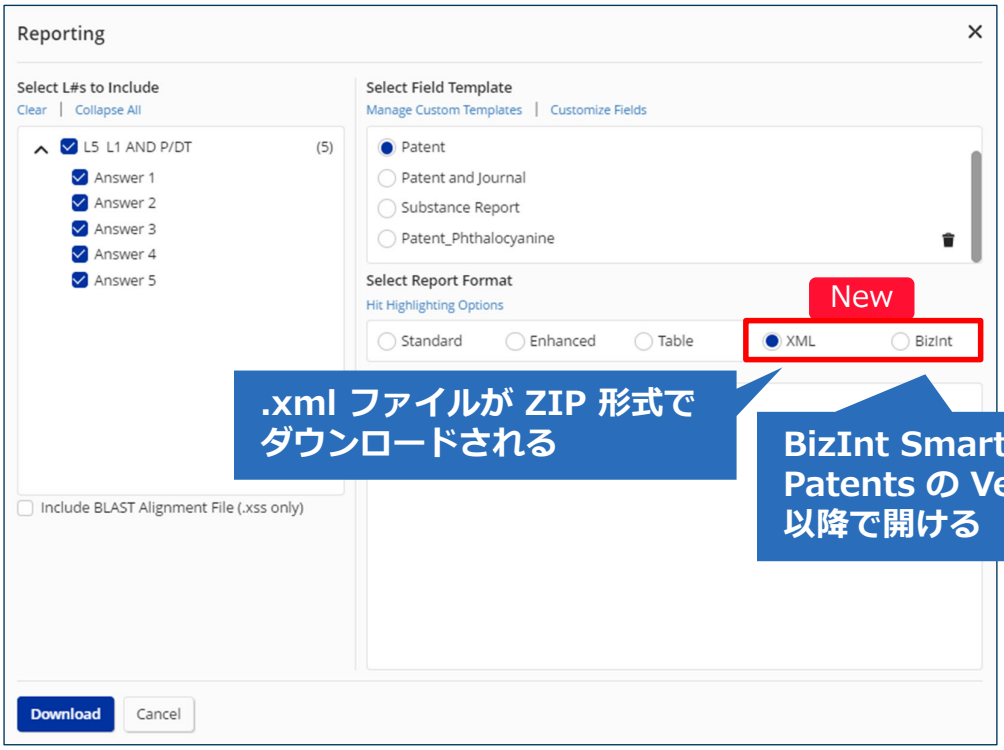
特許情報（特許番号、特許種別、特許発行日）と出願情報（出願番号、出願日）の表が一つのセルに入ったテーブルを作成できる

優先権情報の表も一つのセルに入る

© 2022 化学情報協会

レポートを XML, BizInt 形式でダウンロード可能に　14

2 つのダウンロード形式が追加された



.xml ファイルが ZIP 形式でダウンロードされる

BizInt Smart Charts for Patents の Version 5.6 以降で開ける

© 2022 化学情報協会

構造作図に関する強化

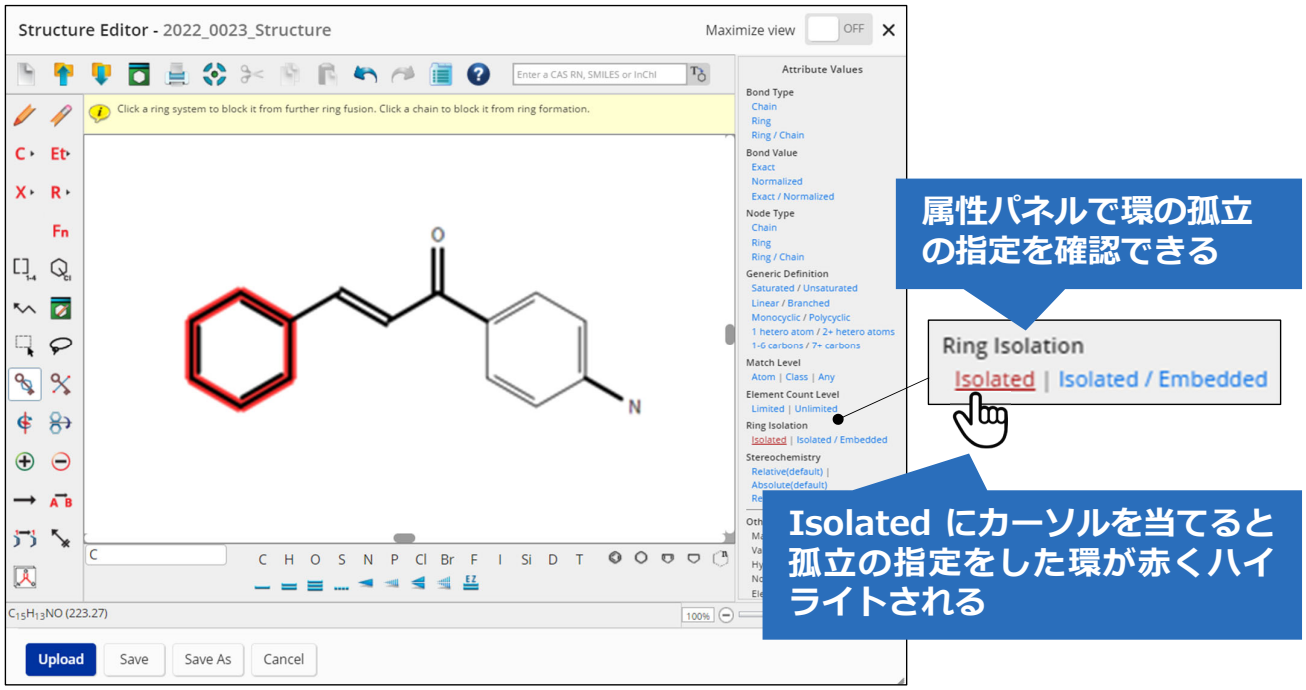
構造作図に関する強化内容 16

- 構造作図画面の属性パネルの強化
- 立体化学構造に関する設定の追加
- 構造質問式のインポートに関する強化
- 構造質問式のアップロードに関する強化

© 2022 化学情報協会

構造作図画面の属性パネルの強化 17

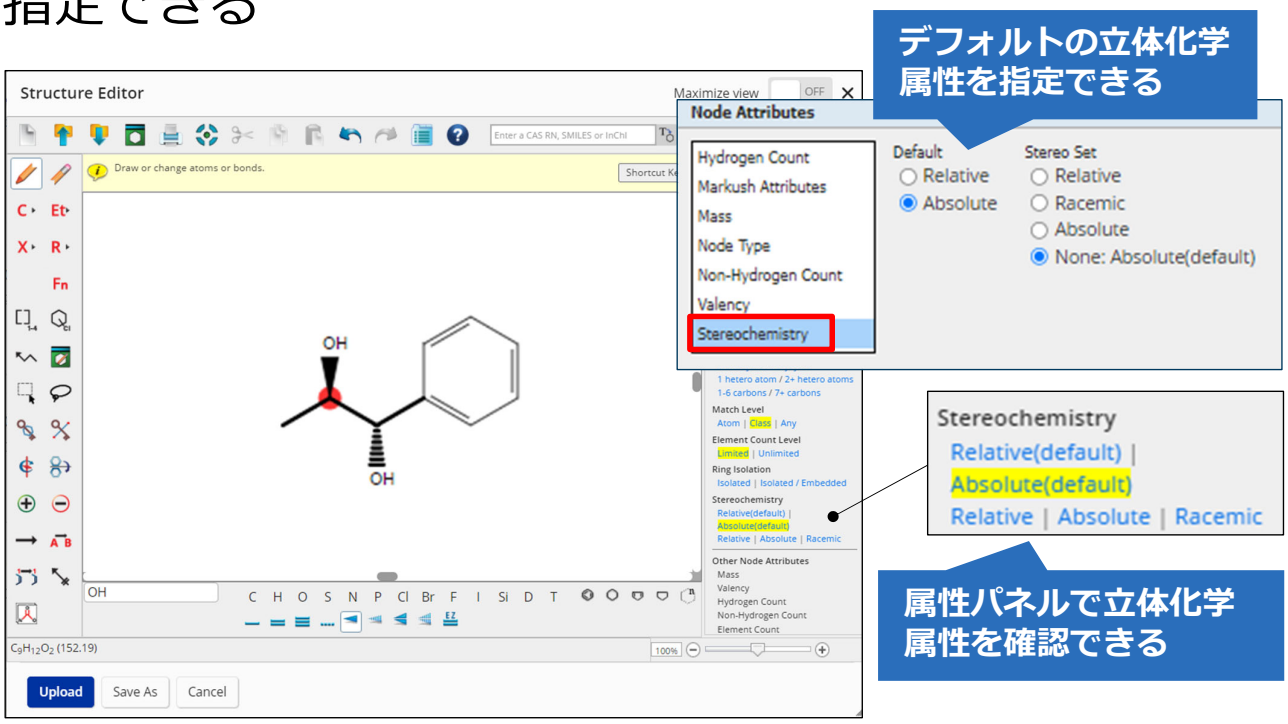
環の孤立の指定を確認できる



© 2022 化学情報協会

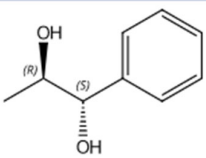
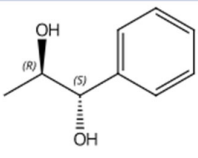
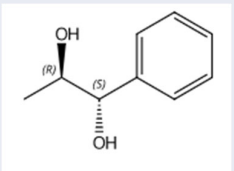
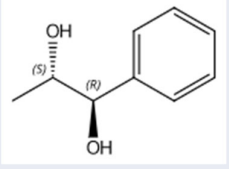
立体化学構造に関する設定の追加 18

不斉中心炭素の立体化学 (Stereochemistry) 属性を指定できる



© 2022 化学情報協会

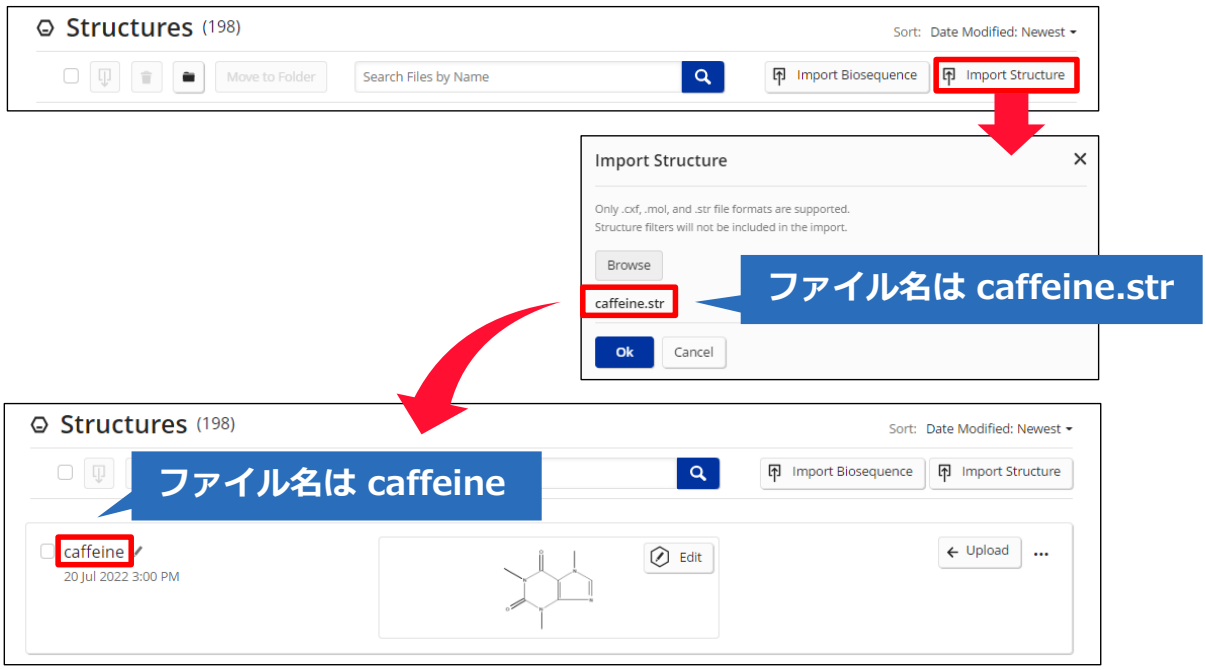
構造質問式の立体化学属性とヒットする回答 19

構造質問式		REGISTRY ファイルのレコード	
構造質問式	デフォルトの立体化学属性	 Absolute stereochemistry shown Rotation (+)	 Relative stereochemistry shown
	Absolute (絶対)	○	×
	Relative (相対)	○	○
	Absolute (絶対)	×	×
	Relative (相対)	○	○

© 2022 化学情報協会

構造質問式のインポートに関する強化 20

構造質問式をインポートした際、インポートしたファイルの名前が反映される

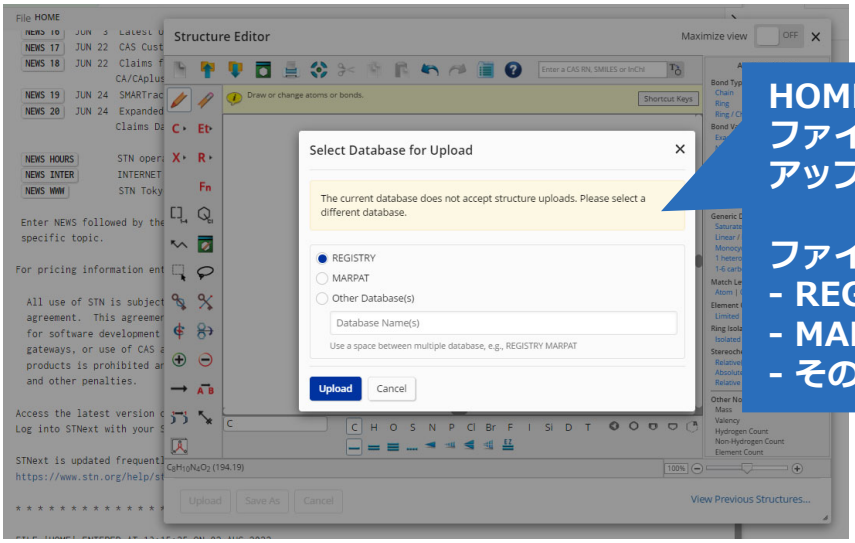


© 2022 化学情報協会

構造質問式のアップロードに関する強化 21

構造質問式のアップロードを HOME ファイルや STNGUIDE ファイルで行った場合でも、ファイルの選択ができる

例：HOME ファイルに入った状態で構造質問式をアップロード



© 2022 化学情報協会

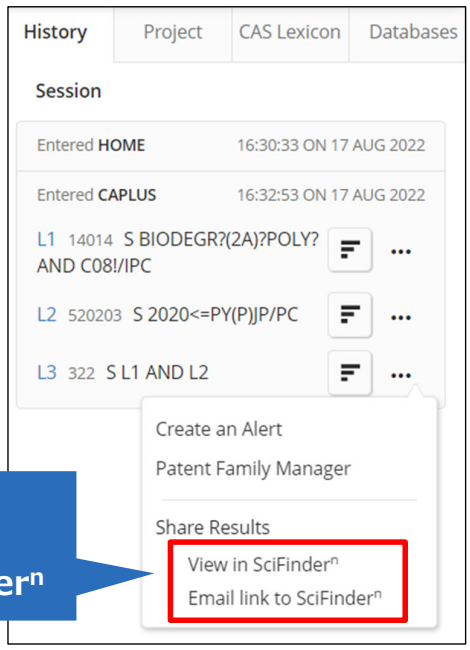
CAS SciFinder^n との連携

CAS SciFinder^n との連携 23

CAS STNext の検索結果を CAS SciFinder^n で再現できる

対象ファイル

- REGISTRY, ZREGISTRY, LREGISTRY
- CAplus, HCAplus, ZCAplus
- MEDLINE

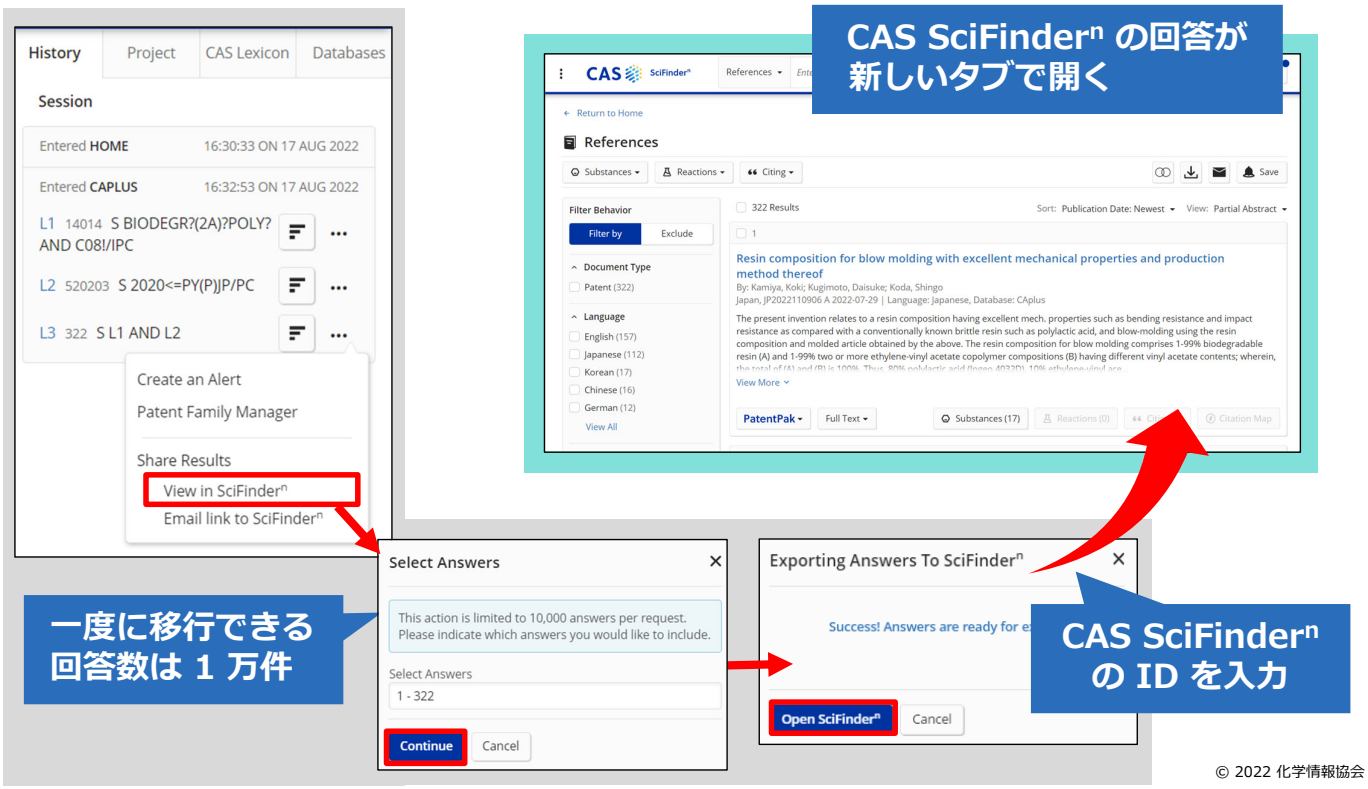


2 通りの方法がある
- View in SciFinder^n
- Email link to SciFinder^n

© 2022 化学情報協会

View in SciFinder^n をクリック 24

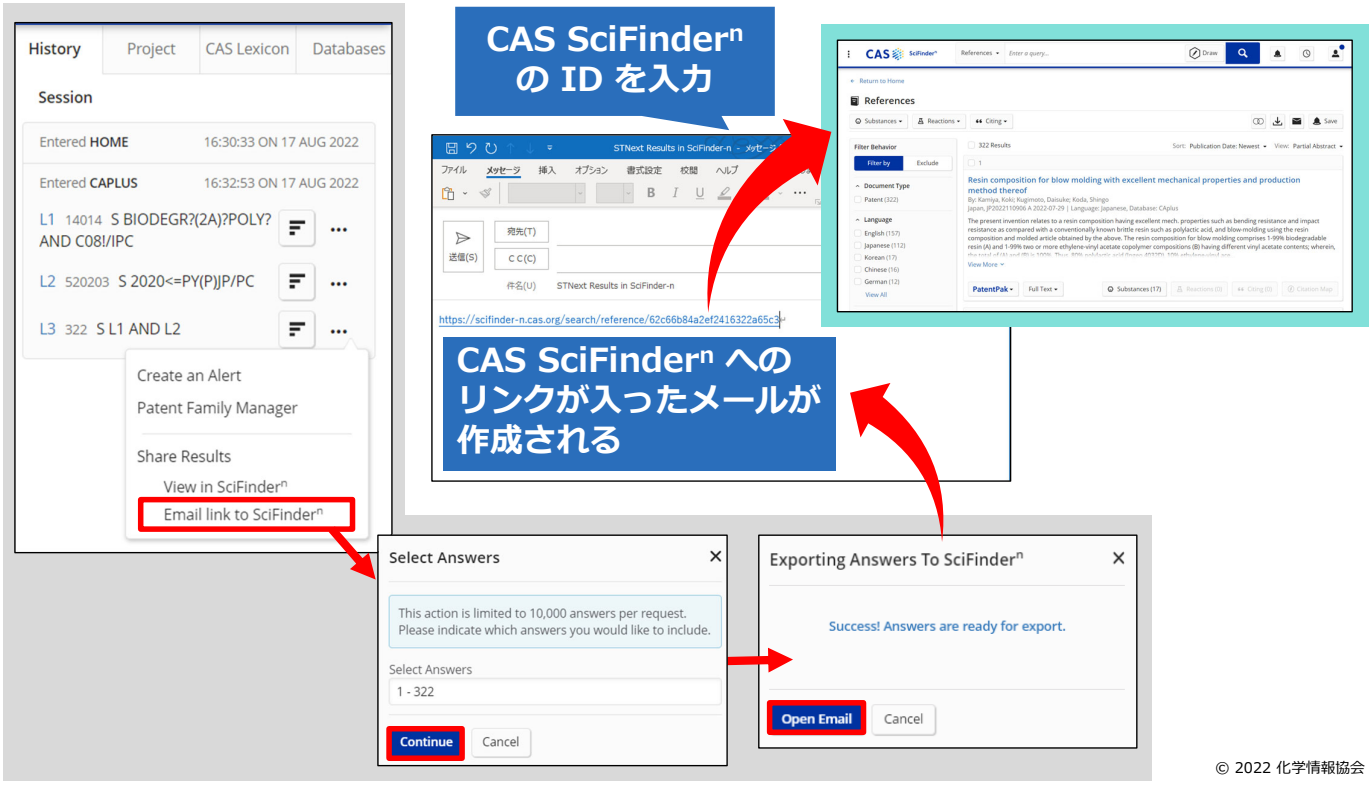
CAS SciFinder^n で回答を表示



© 2022 化学情報協会

Email link to SciFinder^n をクリック 25

CAS SciFinder^n へのリンクが入ったメールを作成

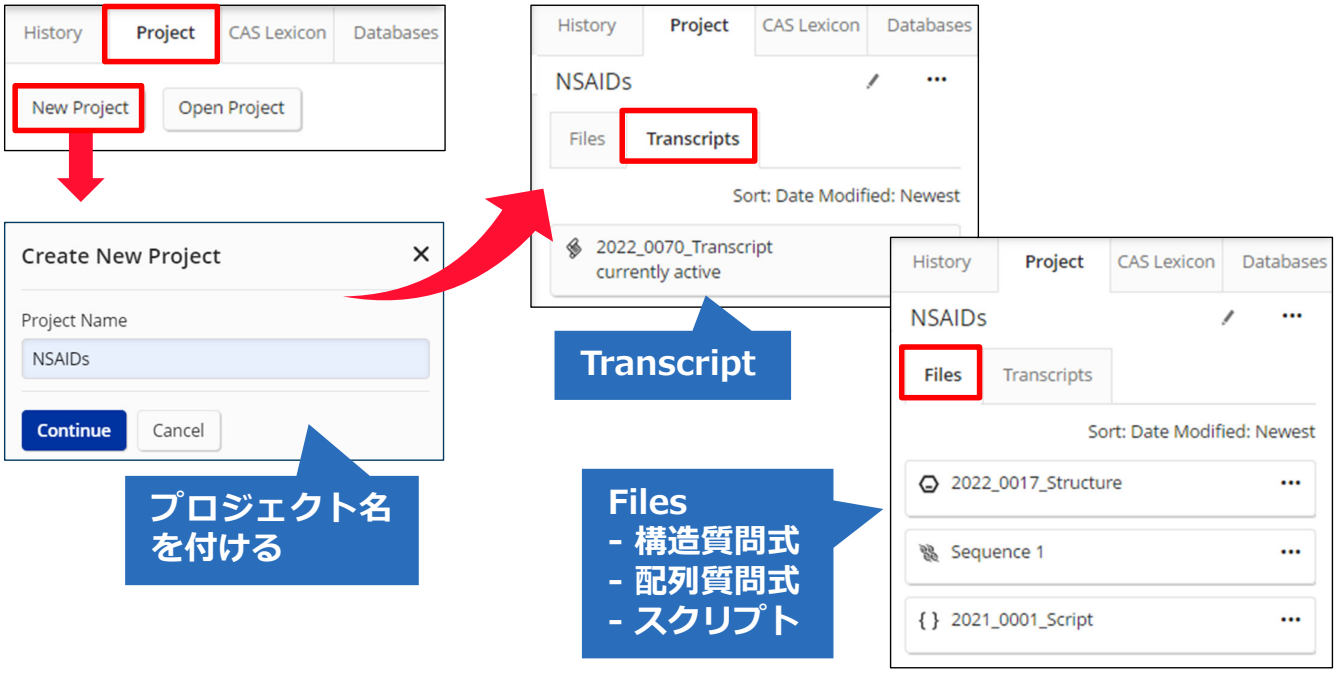


© 2022 化学情報協会

その他の強化

Project タブの追加 27

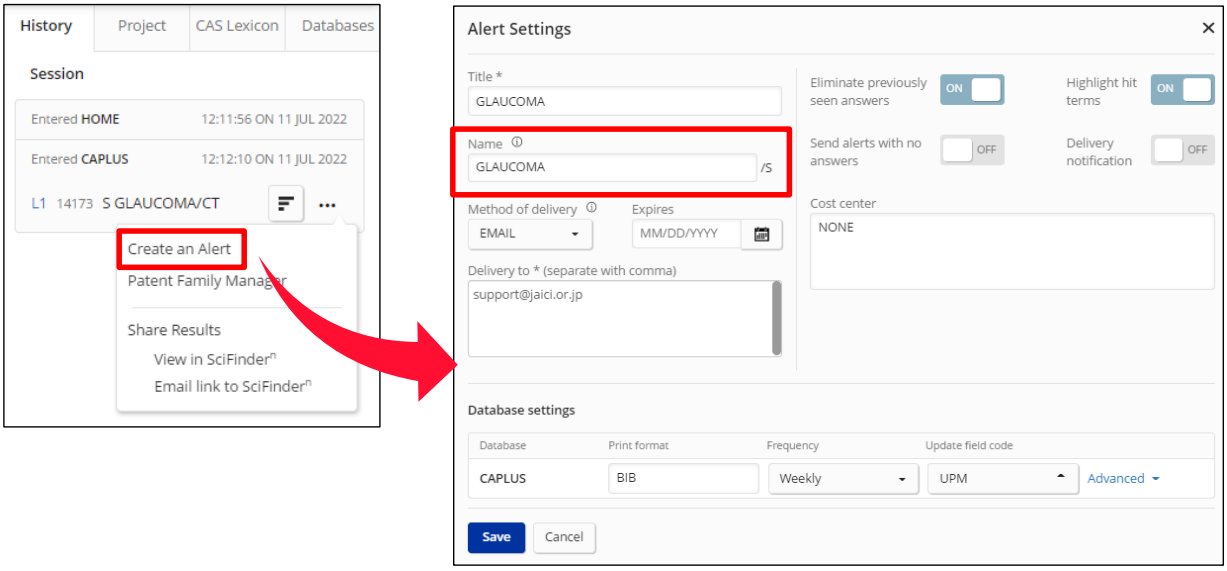
調査に関する情報を一つのプロジェクトにまとめて管理できる



© 2022 化学情報協会

アラートの強化 28

Alert Settings 画面でアラート登録名を設定できる

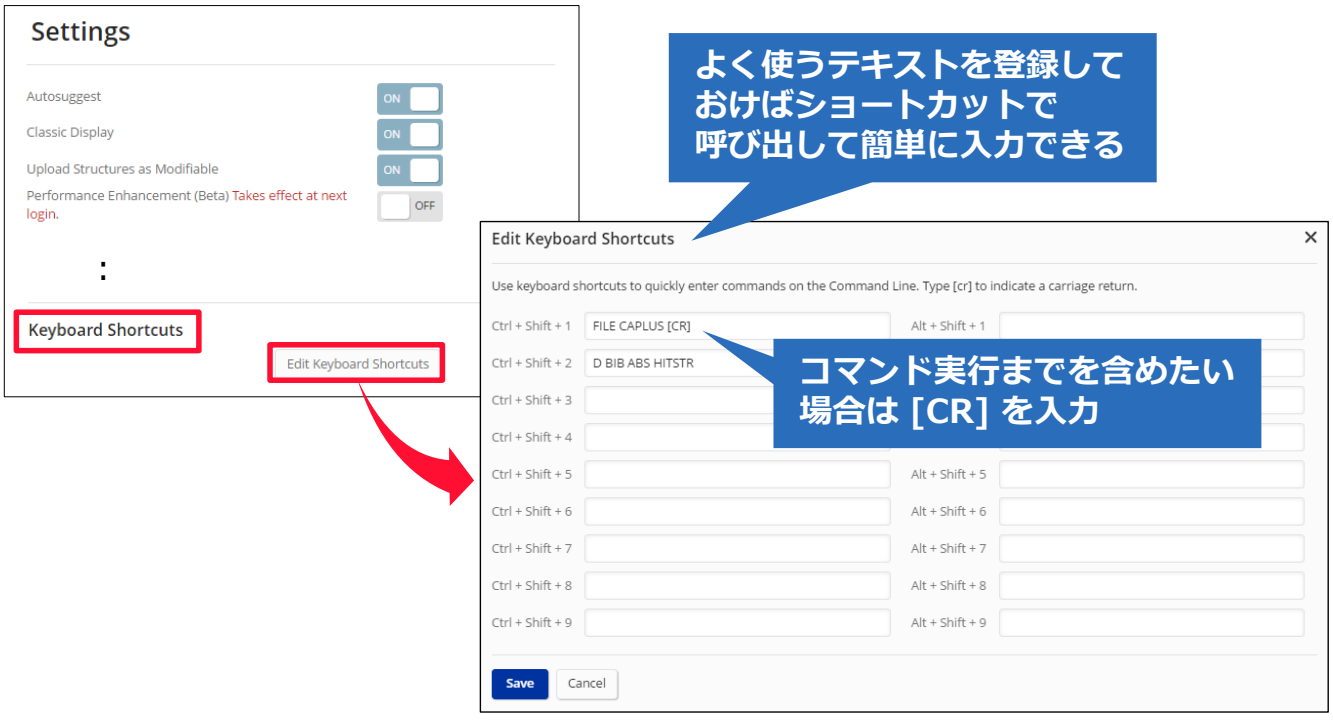


アラート登録名はアルファベットで始まる
12 文字以内の半角英数字

© 2022 化学情報協会

Keyboard Shortcuts 設定の追加 29

Settings に Keyboard Shortcuts が追加された



よく使うテキストを登録して
おけばショートカットで
呼び出して簡単に入力できる

コマンド実行までを含めたい
場合は [CR] を入力

© 2022 化学情報協会

まとめ 30

- セッション開始時に、デフォルトの Transcript 名から、わかりやすい名前に変更できる
- 検索記録やレポートでヒットタームハイライト色を選択できる
- 構造作図画面の属性パネルで環の孤立や立体化学属性を確認できる
- CAS STNext の検索結果を CAS SciFinder^{n} で再現できる

© 2022 化学情報協会

2021 年 6 月以降の強化一覧

■ データベースの強化・変更

ファイル名	強化点
ADISNEWS	・ Pharmacoeconomics and Outcomes News の発行終了
AUPATFULL	・ INPADOC 由来の情報を表示不可に
CANPATFULL	・ INPADOC 由来の情報を表示不可に
CAplus/CA	・ Prior Art Analysis（先行技術調査）機能の搭載 ・ 特許請求項の収録を拡大 ・ 更新コード STUP がアラートで利用可能に
CHEMLIST	・ UK REACH 関連情報の追加
DCR	・ リリース
CNFULL	・ INPADOC 由来の情報を表示不可に
DEFULL	・ INPADOC 由来の情報を表示不可に
EMBASE	・ EMTREE 語のオンラインシソーラス更新
EPFULL	・ アラートの実行頻度に毎月が追加 ・ INPADOC 由来の情報を表示不可に
FRFULL	・ INPADOC 由来の情報を表示不可に
GENESEQ （旧 DGENE）	・ リロード ・ 完全配列検索、部分配列検索に相補鎖オプション追加 ・ ホモロジー検索（BLAST, GETSIM）結果の回収オプションの追加
GRFULL	・ INPADOC 由来の情報を表示不可に
INFULL	・ INPADOC 由来の情報を表示不可に
INPADOCDB INPAFAMDB	・ 特許ステータスの追加 ・ 米国特許の特許調整期間が検索可能に ・ 法的状況イベントクラス（LSEC）の追加 ・ ウクライナ特許の収録強化
JPFULL	・ INPADOC 由来の情報を表示不可に
KRFULL	・ INPADOC 由来の情報を表示不可に
MARPAT	・ 特許ステータスが表示されるように
MEDLINE	・ リロード – 2022 年度版 MeSH シソーラスの搭載 – 索引付与の自動化
NTIS	・ リロード – /TI, /AB フィールドで後方一致検索，中間一致検索可能に – テキスト中の数値検索機能（Ver.3.3）の追加 – アラートを設定可能に – /PN, /AP フィールドの削除 – レコード番号の形式変更

■ データベースの強化（続き）

ファイル名	強化点
PATGENE (旧 PCTGEN)	・ リロード ・ ホモロジー検索の回答作成時のオプション変更 ・ 完全配列検索、部分配列検索に相補鎖オプション追加 ・ ホモロジー検索（BLAST, GETSIM）結果の回収オプションの追加
TOXCENTER	・ MEDLINE セグメントのリロード
RAPRA	・ 更新中止
RDISCLOSURE	・ リロード – /TX で後方一致検索，中間一致検索可能に – テキスト中の数値検索機能（Ver.3.3）の追加 – 国際特許分類（IPC），ヨーロッパ特許分類（EPC）関連情報を削除
RUFULL	・ リリース（ロシア特許全文データベース）
USGENE	・ リロード
WSCA	・ 更新中止
WPINDEX/WPIDS/WPIX	・ リロード ・ マニュアルコード改訂（2022 年版）

■ データベースのサービス終了

文献データベース	AEROSPACE	COMPUSCIENCE	DKF	DJSMDS/DJSMONLINE
	ENERGY	INSPHYS	METADEX	TRIBO
	WELDASEARCH			
特許データベース	DPCI	FRANCEPAT	JAPIO	KOREAPAT
	PATDD	PATDPA	PATDPAFULL	PATDPASPC
	RUSSIAPAT			
化学物質 データベース	ICSD			
化学物質規制 データベース	MSDS-OHS			
練習ファイル	LDPCI	LDRUG	LPATDPA	LWPI

2021 年 6 月以降の強化一覧

■ STNext の強化

	強化点
検索記録の強化	・ ログオン時に Transcript Options を表示可能に ・ 検索記録のヒットタームハイライト色を選択可能に ・ レポート作成時にハイライト色などを選択可能に ・ Excel 形式テーブルのダウンロードのオプション追加 ・ レポートを XML 形式でダウンロード可能に ・ レポートを BizInt 形式でダウンロード可能に ・ レポートに自動で含まれていたフィールドが選択可能に ・ HITSTR 表示形式の出力を含むレポートやテーブルの改善 （CAplus/CA ファイル） ・ HITSTR 表示形式の出力を含むレポートやテーブルの改善 （USPATFULL/USPATOLD/USPAT2 ファイル）
構造作図に関する強化	・ 構造作図画面の最大化が可能に ・ 環の孤立の設定の確認が簡単に ・ 立体化学構造のより詳細な指定が可能に ・ 炭素の表示を変更可能に ・ 結合の重なりの警告が表示されるように ・ インポートした構造質問式の名前がファイル名に反映されるように ・ 構造質問式をアップロードする際のデータベースに関する強化
CAS SciFinder^{n}との連携	・ CAS STNext の結果を CAS SciFinder^{n} で再現できるように
Biosequences Search の強化	・ NCBI 由来の配列の追加 ・ Biosequences Search の結果に非特許文献情報を追加 ・ Biosequences Search に Organism フィルターと Sequence Length フィルターを追加 ・ Motif 配列検索機能の強化
その他	・ Project タブの追加 ・ Project タブの Transcript から Query Summary をダウンロード可能に ・ Alert Settings 画面でアラート登録名を設定できるように ・ Alert Settings 画面で E-mail アドレスの入力漏れに対して警告が表示されるように ・ Keyboard Shortcuts 設定の追加 ・ スクリプト実行速度の改善 ・ 構造質問式からケミカルコードの検索式を作成可能に （ダウエント会員のみ）

2021 年 6 月以降の強化一覧

■ CAS PatentPak の強化

- PPAK フィールドに Claim タグを追加
- 物質情報付き明細書（Interactive Viewer）の強化
- 配列の記載位置情報の付与開始
- リンク有効期間が 365 日に

■ STN の従来型インターフェースのサービス終了

- STN のインターフェースは STNext に統一された。それに伴い下記のサービスは終了した。

  - STN 新プラットフォーム
  - STN Express
  - STN on the Web
  - STN Easy
  - STN AnaVist

JAICI
化学情報協会

情報事業部
〒113-0021 東京都文京区本駒込6-25-4 中居ビル
TEL: 0120-003-462 FAX: 03-5978-4090
URL: www.jaici.or.jp
E-mail: support@jaici.or.jp