

DNA塩基配列関連メタデータ検索システムDDBJ Searchの刷新

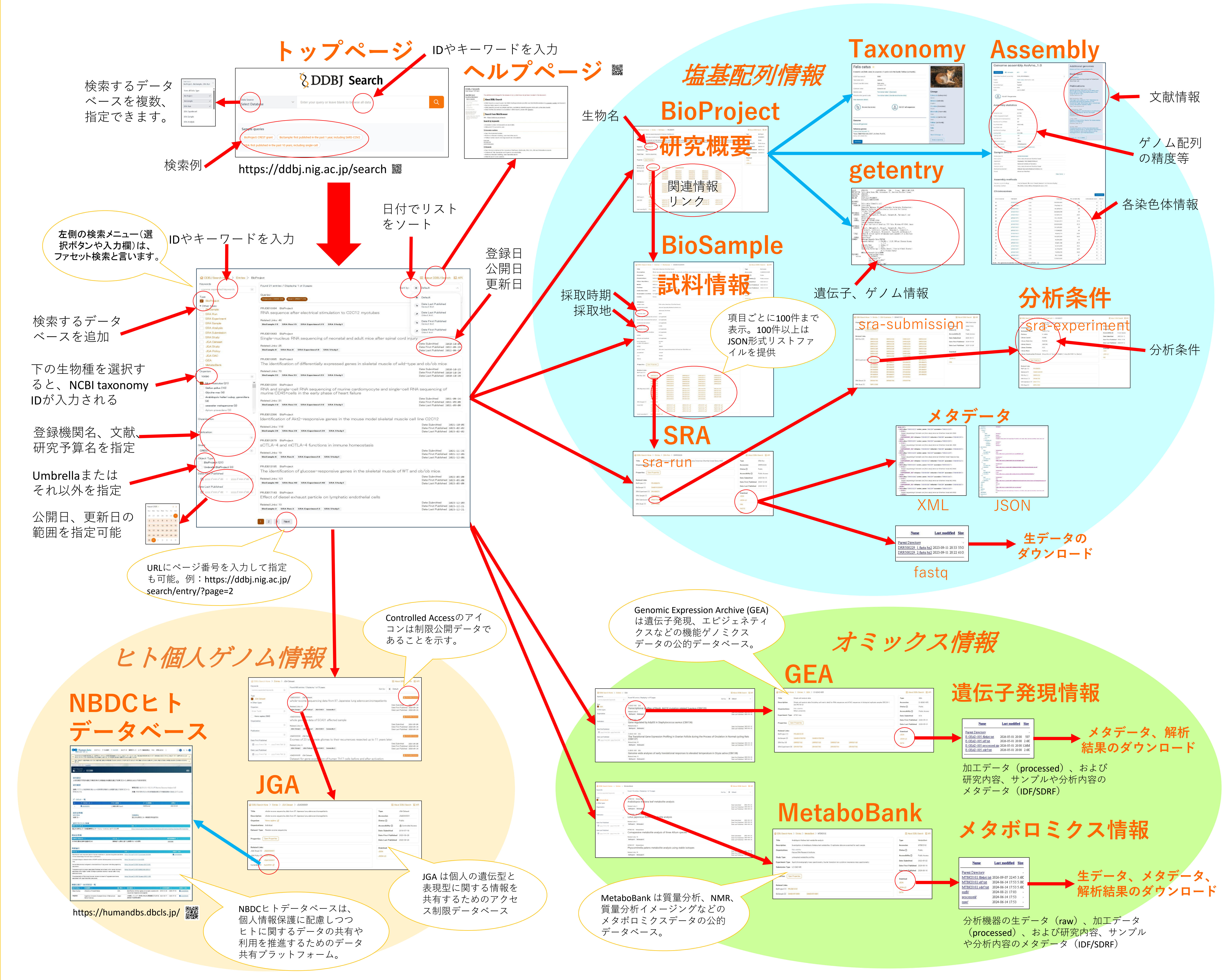
○藤澤貴智、荒 武、児玉悠一(遺伝研)、末竹裕貴 (Sator, Inc.)、小野田智 (Yohak LLC.)、大石直哉 (dogrun Inc.)、有田正規(遺伝研)

要約

DDBJが参画する国際塩基配列データベース連携（INSDC）は、生命医科学研究をより深くサポートするために、配列データに関連する様々なメタデータを種別ごとに維持・管理するインフラや標準づくりに注力してきた。それらメタデータは、大規模な塩基配列データを分類・選別してゲノム医療やバイオものづくりに役立てるのに必要なだけでなく、来歴や地理情報をたどって環境モニタリングや生物資源の保全に活かすなど、様々な二次データベースの基盤となっている。

DDBJの提供する検索システム（Search）は、これまでINSDCが管理するBioProject、BioSampleなどのメタデータ、次世代シーケンス配列（SRA）の付随情報を対象としてきた。しかし近年はマルチオミックスの研究が急増し、ヒト個人ゲノム情報データベース（JGA）の関連情報、DNAチップやRNA-Seqなどによる遺伝子発現量（GEA）、メタボロミクス情報（MetaboBank）を含めた統合検索が必要とされている。そこで、検索システム全体を刷新し、データ自体も整備して、より高度なサービスを提供できるSearchを公開した。具体的には、バックエンド処理の効率化によるデータ更新速度の改善、APIの改良による検索機能の拡充、画面構成の整理と再構築による表示機能の改善を実現したので紹介する。

DDBJ Searchによる塩基配列及びオミックス情報の探索



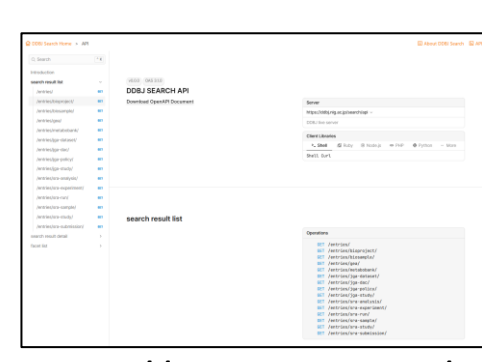
APIサービス

DDBJ Searchの検索と表示機能は、APIの機能を利用して構築されています。API自体は以下のGitHubで公開しています。API機能の詳細につきましては、以下の文書等をご参照ください。



システム構成や仕様の詳細等の説明があります。

<https://github.com/ddbj/ddbj-search-api>



APIの各機能について、パラメータを変えてテスト実行ができます。

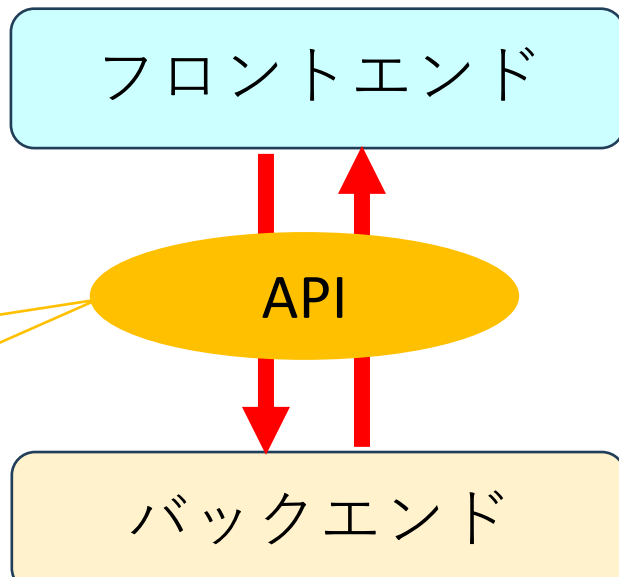
<https://ddbj.nig.ac.jp/search/api-doc/>



APIの各機能について、パラメータや利用事例の説明があります。

<https://ddbj.nig.ac.jp/search/api/docs>

APIを利用すると、自分のスクリプトやプログラムで検索ができます。



今後の展望

1. メタデータと検索性の拡充

表示可能なメタデータ項目を増やし、検索可能なメタデータ項目を拡充します。INSDC関連データのアセンブリデータのID検索の追加を行います。

2. NBDCヒトデータベースの申請・利用の強化

ヒト個人ゲノム配列等の登録に必要な、NBDCヒトデータベースへの申請の効率化、利用拡大のための連携強化を行います。

3. 生物資源の利活用と生物多様性への貢献

生物資源データベースNBRPの利用促進や、生物多様性に関わる情報の拡充を行うために、関連データベースとの連携強化を進めていきます。



Human data



NBRP GBIF

DDBJのサービスや活動への質問等は、以下よりお問い合わせください。

<https://www.ddbj.nig.ac.jp/contact-ddbj.html>



大学共同利用機関法人 情報・システム研究機構 国立遺伝学研究所 バイオデータ研究拠点
〒411-8540 静岡県三島市谷田1111 お問い合わせ: <https://www.ddbj.nig.ac.jp/contact.html>



Creative Commons 表示4.0国際ライセンス