

バイオバンク横断検索システムの概要

1. システム概要

本システムは、以下の3つの要素から構成されます。

- ① 利用者認証
- ② バイオバンク横断検索入力及び検索結果表示等の利用者とのインタフェースを担うとともに、検索内容を各バイオバンクに問合せてその結果を受ける「バイオバンク横断検索機能」
- ③ 複数の基幹バイオバンクに実装され、「バイオバンク横断検索機能」からの検索条件を受けて、生体試料・情報を検索し、その検索結果を返却する、「バイオバンクノード」

(1) 横断検索の流れ

- ① 利用者は、「バイオバンク横断検索システム (<https://www.biobank-network.jp/cross-search>)」での認証を受けてシステムに入ります。
- ② 利用者は、「バイオバンク横断検索機能」の「クライアント側ユーザインタフェース」を介して、検索条件を設定して所望の試料の検索を行います。
- ③ 「バイオバンク横断検索機能」は、利用者から受け付けた検索条件をインプットとして、各「バイオバンクノード」が公開している「横断検索 API」に Web 経由で問合せを行います。
- ④ 各「バイオバンクノード」では、「横断検索 API」を介して受け取った検索リクエストに応答して、自施設で管理している試料に関する情報の検索を行い、その検索結果を「横断検索 API」を介して Web 経由で返却します。
- ⑤ 「バイオバンク横断検索機能」は、各「横断検索 API」から返却された検索結果を集約して、利用者に表示します。

(2) 基幹バイオバンクでの対応

基幹バイオバンクでは、以下の機能を有する「バイオバンクノード」を準備します。

- ① 自機関ならびに、「バイオバンクノード」を有さないバイオバンクの生体試料・情報を管理するデータベースを用意し、横断検索システム全体で取り決められた「共通項目」による検索を可能とする検索機能を実装します。
- ② 「バイオバンク横断検索機能」からの Web 経由の検索リクエストを受け、検索結果を返却する「横断検索 API」を実装します。

2. バイオバンク間の共通項目

各バイオバンクに保管されている生体試料・情報を横断的に検索するためには、試料に付随する情報の定義を共通化する必要があります。どこまでの情報を共通化した定義として取扱うことができるか、という点は、最初にコンセンサスを得るべき課題であり、これを「共通項目」と称し、参加するバイオバンク間で調整を行いました。

この共通項目は、国際的には例えば MIABIS 2.0 Core (Minimum Information About BIobank data Sharing)^{※1} として、欧州のバイオバンクコンソーシアム (BBMRI-ERIC ; Biobanking and BioMolecular Resources Research Infrastructure – European Research Infrastructure Consortium)^{※2} で定義されており、これらを参考にしてバイオバンク横断検索システムでの共通項目を定義しました。

共通項目は、横断検索時の検索項目として使用可能とするとともに、検索結果の試料の情報として表示するもので、以下の項目を共通項目としています。

共通項目一覧

分類	項目	
バイオバンク	バイオバンク ID	横断検索システムで使用する各バイオバンク固有の ID
協力者	協力者 ID	試料提供者の各バイオバンク内での ID (外部公表可能な ID)
	性別	ISO5218/MIABIS2.0
	既往症・併存症コード	ICD-10
	既往症・併存症	テキスト
試料	病名コード	ICD-10
	病名	テキスト
	疾患特異的臨床情報	※3
	試料種類	MIABIS2.0
	試料採取時年齢	
	試料品質管理情報	※4
	同意情報	※5
解析情報	解析情報種類	
	ベンダー	
	プラットフォーム	

※1 <http://www.bbmri-eric.eu/news-events/miabis-2-0-published/>

※2 <http://www.bbmri-eric.eu/BBMRI-ERIC/about-us/>

- ※3 疾患特異的臨床情報については、口腔がん、頭頸部癌（口唇・口腔）、頭頸部癌（上咽頭）、頭頸部癌（中咽頭）、頭頸部癌（下咽頭・咽頭・気管・副咽頭間隙）、頭頸部癌（鼻腔・副鼻腔・頭蓋底）、食道癌、胃癌、大腸癌、肝癌、胆道癌、膵癌、肺癌、乳癌、甲状腺癌、悪性骨腫瘍、悪性軟部腫瘍、子宮頸癌、子宮体癌、卵巣癌、腎癌、副腎腫瘍、腎盂・尿管・膀胱癌、前立腺癌、精巣腫瘍について、標準を選定して共通項目とすること。
- ※4 試料品質管理情報については、保管期間、SOP からの逸脱、保管温度、保管条件（容器の種類）、液性試料については、採取から処理開始までの状態（時間）、採取容器の種類、採取から処理開始までの状態（温度）、組織試料については冷阻血時間（摘出から固定開始までの時間）、固定時間、固定方法、核酸試料については凍結融解回数、由来試料、核酸抽出の方法（抽出方法）、核酸抽出の方法（キット）、保管媒体について、国際的な標準を選定して、共通項目とすること。
- ※5 データ利用条件に基づく国際的な標準を選定して、共通項目とすること。

3. 横断検索 Web API (アプリケーション・プログラミング・インターフェース)

(1) 概要

- ・ バイオバンク横断検索機能からの検索リクエストに回答して、バイオバンクで管理している協力者情報の検索及び検索結果を返却する API を実装します。
- ・ バイオバンク横断検索機能は利用者から受け付けた検索条件をインプットとして基幹バイオバンクが公開している API に問合せを行い、API から返却された検索結果を集約して利用者に表示します。

(2) 横断検索 API の設計方針

バイオバンク横断検索機能からの検索リクエストに合致する情報が大量になる場合、返却するデータが増大する事が想定されるので、レスポンス内容を圧縮するとともに、バイオバンクノード側で 1 回の API 実行における協力者数を調整してデータ通信量の削減できるようにすること。

① 検索 API

バイオバンク横断検索機能から送信された検索条件と合致する協力者情報を検索し、検索条件と合致した協力者情報の件数と、検索条件と合致した協力者情報を特定するための協力者 ID を返却する API。

② データ取得 API

バイオバンク横断検索機能から送信された協力者 ID と合致する協力者情報を返却する API。(協力者情報の協力者 ID は複数指定することも可能)

(3) セキュリティ

- ・ バイオバンク横断検索機能と各バイオバンクの API 間は HTTPS または HTTP を用いて通信を行います。
- ・ 各バイオバンクはバイオバンク横断検索機能からの問合せのみ受け付けるため、API のアクセス元 IP アドレスの制限を行います。
- ・ 各バイオバンクの API を利用するには、各バイオバンクで発行した API キーをリクエストで指定するものとします。API キーは定期的に変更する運用とし、API キーの移行期間を考慮して最大 2 個の API キーを同時に許可できるものとします。
- ・ API キーは HTTP ヘッダに格納して通信を行います。

(4) API の仕様について

① 共通事項

- ・ API へのリクエスト、及び API からのレスポンスは JSON フォーマットで送受信します。
- ・ API の実行結果はレスポンスボディで返却すると共に、実行結果のステータスコードを HTTP ステータスコードで返却します。API から返却する HTTP ステータスコードは RFC 2616 等に準拠する形式で定義します。

② 検索 API

- ・ バイオバンク横断検索機能からのリクエストにより検索条件を取得し、該当するデータの検索を行います。
- ・ 検案件数の上限値を横断検索 API 毎に保持します。
- ・ 検索結果は、協力者 ID の昇順でソートし、検案件数の上限値の範囲内の協力者のデータのみ返します。
- ・ 検索リクエストから取得した API キーを用いた認証処理に失敗した場合は検索不可とします。

③ データ取得用 API

- ・ バイオバンク横断検索機能からのリクエストよりデータの協力者 ID の一覧を取得して、該当するデータの詳細情報を返却します。

(5) API の実行について

バイオバンク横断検索機能から各バイオバンクの API を実行する際は、以下の点を考慮します。

- ・ 各バイオバンクの API を実行する際は、一定時間内に応答がない場合に接続を切断する仕組みを実装します。
- ・ 同時期に検索できるリクエスト数は東北大側の横断検索結果表示 API で制限を設けます。

4. 横断検索 Web UI (検索画面ユーザインタフェース)

バイオバンク横断検索システムでは、利用者が直感的かつ効果的にバイオバンクの生体試料・情報の横断検索を行えるようにします。現バイオバンク横断検索システムは <https://www.biobank-network.jp/cross-search> をご参照ください。

5. 利用者認証

バイオバンク横断検索システムでは、各バイオバンクに保管してある生体試料・情報の詳細が検索でき、個票の内容を把握することも可能となるため、利用者の認証について考慮が必要です。一般的に、ユーザ認証には色々な方法や考え方があり、実運用に向けてはその検討が必要です。

6. その他

バイオバンク横断検索システムは、Web を使用して複数の異なるシステム間で検索リクエストを行い、応答を返すシステムであり、予想外の待ち時間等が発生する可能性があるため、実行時間や検索件数等に制限を設けて、利用者の利便性を確保することが必要です。そのため、下表に示す項目に制限を設けています。

項目	説明	設定値	単位
実行時間（リクエストタイムアウト）	各横断検索の完了を待ち受けるタイムアウト時間	180	秒
同時実行数（上限）	利用者からの検索リクエストを最大何件同時に処理するか	20	スレッド
最大検索待ち件数	同時実行数を超過した場合に、何件まで処理待ちを許可するか	100	スレッド
シャットダウン時の実行完了待ち時間	指定された時間、実行完了まで待つてから停止する	3	分
各バイオバンク検索の同時実行数（上限）	バイオバンク毎の検索スレッド数（上限）	80	スレッド
各バイオバンク検索の最大検索待ち件数	各バイオバンク検索の同時実行数を超過した場合に、何件まで処理待ちを許可するか 基本的に、各バイオバンクの検索待ちは起きない想定	100	件
各バイオバンク検索のシャットダウン時の実行完了待ち時間	指定された時間、実行完了まで待つてから停止する	3	分
個票の最大返却件数		1,000	件
検索履歴の最大返却件数		30	件