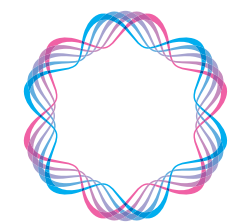


健康・医療研究開発データ 統合利活用プラットフォーム事業

AMED data utilization platform project



AMEDデータ利活用
プラットフォーム



国立研究開発法人 日本医療研究開発機構

ゲノム・データ基盤事業部 データ利活用推進課



〒100-0004 東京都千代田区大手町 1-7-1 読売新聞ビル 22F

TEL: 03-6870-2222

E-mail: platform@amed.go.jp

<https://www.amed.go.jp/program/list/14/04/001.html>



国立研究開発法人 日本医療研究開発機構

AMEDデータ利活用プラットフォーム

健康・医療研究開発データ統合利活用プラットフォーム事業は、AMED研究から生み出されたデータ利活用の促進を可能とする場を目指し、^①3大バイオバンクの全ゲノム解析データのメタデータの横断検索～データの処理までワンストップで実行可能なサービスを提供することで、ゲノムデータ利活用を推進する役割を果たします。



各種申請窓口一本化による手続簡素化
他システムとの同一IDによる簡便なログイン



3大バイオバンクの全ゲノム解析データ（約2万件）のメタデータの横断検索による研究計画立案の効率化



連携5拠点におけるシステムへワンストップで接続可能となることによるデータ処理までのスピード効率化



連携5拠点

データ処理

東北メディカル・メガバンク機構
東京大学 医科学研究所

国立遺伝学研究所
国立国際医療研究センター
国立がん研究センター

複数拠点に格納されているゲノムデータを横断して実施する統合的なデータ処理を含む

★
1 受付事務局確認

利用申請

利用者

- ・アカデミア研究者
- ・民間企業等

2

利用者認証

3

横断検索

4

予備的処理

★★
5 データ利用審査

データ利用申請

プラットフォームからワンストップで接続

他システムの利用

他システム例（調整中）

- ・バイオバンク横断検索システム
- ・全ゲノム解析等実行計画

臨床基本4属性情報および^②アレル頻度情報等を用いた研究計画の立案

3大バイオバンクの全ゲノム解析データのメタデータ
（^④JGA^⑤スキーマ+臨床基本4属性情報※）
※年齢階層（10歳単位）、性別、居住地または出生地、疾患名および疾病コード（^⑥ICD-10）

連携基盤

プラットフォーム連携基盤およびデータ利用に関する規則群

有識者会議

プラットフォーム及びデータ利用に関する規則、システム等、本プラットフォームの運営を支える重要な仕組みにおいて議論を行う

受付事務局

本プラットフォーム利用申請、全ゲノム解析データ及びデータ処理ノードの利用申請、各種問い合わせへの対応を行う

データ利用審査会

データ提供機関のデータ利用に際し、公平・公正・中立な立場でデータ利用審査を実施する

AMED事務局：ゲノム・データ基盤事業部

★
1

利用申請

受付事務局確認

- ・連携基盤 利用申請書
- ・データ取扱者の身分証明書の写し
- ・情報セキュリティガイドライン（データ利用機関向け）のチェックリスト、他

★★
5

データ利用申請

データ利用審査

- ・利用申請書（利活用個人データ及び連携拠点におけるシステム）
- ・研究計画書、研究計画書の機関内承認を示す文書、倫理審査委員会の審査承認を示す文書
- ・データ利用計画の遂行能力を持つことを示す文書（例：関連分野での具体的な研究実績、研究論文、等）
- ・「本人識別等の禁止行為に関するオンライン研修」の受講を終了したことを示す書類
- ・AMEDデータ利活用プラットフォームデータ利用許諾等に関する契約書（利用承認後）、他

AMEDデータ利活用プラットフォームの想定活用事例

1

製薬企業における創業プロジェクトの優先順位付け

概要

- AMEDデータ利活用プラットフォームの利用申請でアクセス可能な約2万件について、標的生体分子の変異情報と疾患との関連を中心に確認を進めた結果、標的酵素Zの変異と疾患Aとの関連が見出された。
- 従来のゲノム関連データベースにより見えなかった相関を捉えられたのは、疾患Aの罹患率が日本人において高いためと推察される。人種による変異頻度の差がある場合、地域特異性を考慮したゲノム解析が必要と考えられた。

2

日本人由来がん細胞株データからの生殖系列多型の除去

概要

- AMEDデータ利活用プラットフォームの利用申請でアクセス可能な約2万件について、日本人由来がん細胞株の全ゲノム配列データからGermline由来のSNPs、SVsの除去を目的に本プラットフォームのゲノムデータを取得。
- 特にCoding領域について非同義的置換の多くからGermline SNP由来のものを差し引くことができ、またSVsについてもGermline由来のものの多くを除去することができた。

3

希少・難治性疾患原因遺伝子^②バリエーションの地域特異性の解明

概要

- ある常染色体潜性遺伝性疾患Xの日本人患者群において見出された病的バリエーションYについて、世界的地域集団に加えAMEDデータ利活用プラットフォームが提供している日本人地域集団におけるアレル頻度の比較を行ったところ、バリエーションYは日本人集団でのみ確認されかつZ地域において有意に頻度が高いことが明らかとなった。
- さらに疾患Xの日本人患者群および同プラットフォームから入手したコントロール群のゲノムデータを用いてバリエーションY近傍のハプロタイプ解析を行ったところ、バリエーションY保有アレルには共通するパターンが見出され、バリエーションYが創始者効果によるものであることが強く示唆された。
- 疾患Xは保険適用された遺伝学的検査の対象となっているが、地域特異性を有するバリエーションYはその遺伝子パネルには収載されておらず、このようなエビデンスの蓄積が今後のパネル改定の機会に活かされることが期待される。

4

小児慢性疾患の新たな疾患感受性遺伝子および日本人特異的バリエーションの発見

概要

- ある小児慢性疾患Xの日本人患者群および非患者群を対象とした既存のSNPアレイデータに対して、AMEDデータ利活用プラットフォームが提供している大規模日本人参照パネルを用いた遺伝型インビュテーションにより高密度かつ高精度に推定された遺伝情報を用いたGWAS解析を行ったところ、複数の疾患感受性遺伝子および病的バリエーションを新たに発見することに成功し、日本人集団で高頻度に見られるバリエーションの存在も明らかとなった。
- 本研究は日本における疾患Xの発症機序の解明や治療薬の開発等に大きく寄与することが期待される。

① 3大バイオバンク

日本における主要なバイオバンクのうち、バイオバンク・ジャパン(BBJ)、東北メディカル・メガバンク計画(TMM)、ナショナルセンター・バイオバンクネットワーク(NCBN)を指す。

② バリエーション

個人のDNAの塩基配列とヒトゲノム計画等において構築された参照配列とを比較した際に見られる配列上の違いのこと。

③ アレル頻度

バリエーションにおける個々の染色体上のDNA配列をアレルといい、集団の中でのアレルの頻度をアレル頻度という。

④ JGA

Japanese Genotype-phenotype Archiveの略称。

⑤ スキーマ

データの構造や要素、属性を定義したもの。

⑥ ICD-10

世界保健機関(WHO)が定めた国際疾病分類第10版のこと。