

日本オンライン情報検索ユーザー会 (OUG)
ライフサイエンス分科会データベース
2022年9月15日 14:00-

ヒト関連情報の共有

科学技術振興機構 NBDC事業推進部
川嶋 実苗

目次

1. 背景
2. NBDCHITデータベース群の紹介
3. TogoVarの紹介

目次

1. 背景

2. NBDCHITデータベース群の紹介

3. TogoVarの紹介

NBDC事業推進部の活動

NBDC

ポータルサイト内を検索

English

サービス イベント ファンディング 研究開発 NBDCについて お問い合わせ

データベース統合を通じて新たな知識へ

NBDCについて

Catalog Cross search Archive Human data TOGOVAR NBDC RDF Portal

すべてのサービス

新着情報

8/25 AJACS オンライン12

パスウェイ&画像データベースを知って・学んで・使う (AJACSオンライン12)

2022年8月25日

オンライン

参加受付中 イベント

トーゴーの日シンポジウム2022 (10/5) オンライン

登録締切 8/5

発表募集

ポスター

トーゴーの日シンポジウム2022 ポスター発表募集

2021年8月5日

オンライン

参加受付中 イベント

NBDC

検索



<FAIR 原則>

Findable

(見つけられる)

Accessible

(アクセスできる)

Interoperable

(相互運用可能)

Re-usable

(再利用できる)

データの価値の最大化、データ駆動型科学の促進

NBDCの活動



ポータルサイト内を検索



English



ヒト

ヒト試料の解析データ全般を再利用

> NBDCヒトデータベース [🔗](#)



概要

参加プロジェクト内等において一般公開前に共有可能なデータを検索

> NBDCグループ共有データベース [🔗](#)



概要

日本人ヒトゲノムのバリエーションを検索

> TogoVar [🔗](#)



概要

ヒト遺伝子アノテーションをキーワードで検索

> HOWDY-R [🔗](#)



概要

遺伝子発現情報、転写開始点情報の検索（ヒト・マウス）

> DBKERO [🔗](#)



概要



3. 大規模解析者向け（一括ダウンロード）

プログラムを駆使して大量のデータを利用する

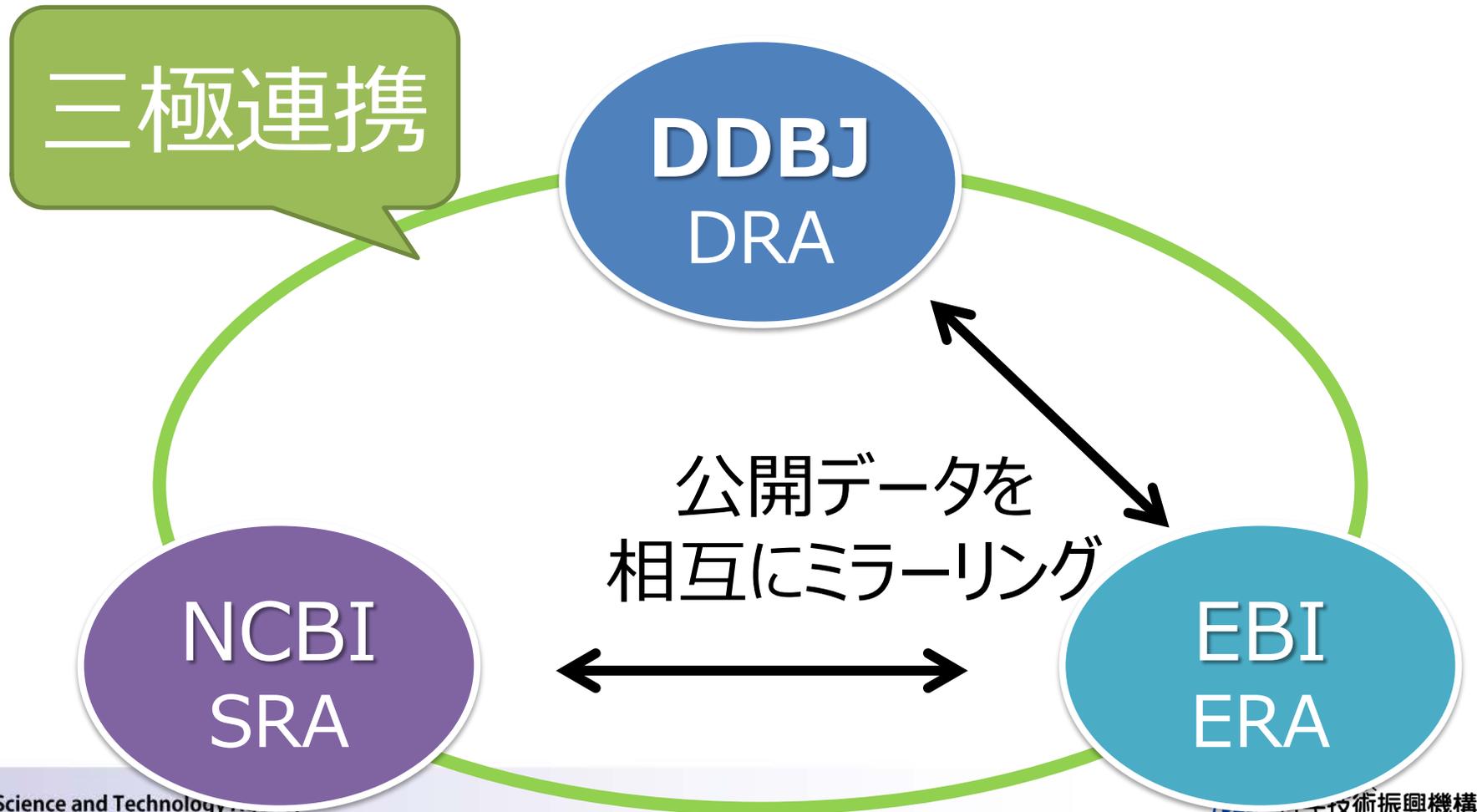


4. データ所有者向け（公開・変換）

手持ちのデータを公開または形式変換する

国際的なデータ共有（非制限公開データ）

International Nucleotide Sequence Database Collaboration (INSDC)



三極連携体制

INSDC

Data type	ROIS-DDBJ	EMBL-EBI	NIH-NCBI
Next generation reads	Sequence Read Archive	European Nucleotide Archive (ENA)	Sequence Read Archive
Capillary reads	Trace Archive		Trace Archive
Annotated sequences	DDBJ		GenBank
Samples	BioSample		BioSample
Studies	BioProject		BioProject
Functional Variations	Genomic Expression Archive (GEA)	Array Express	Gene Expression Omnibus (GEO)
Structural Variations	JVar-SV	Database of Genomic Variants archive (DGVa)	dbVar
Short Genetic Variations	JVar-SNP	European Variants archive (EVA)	dbSNP
Controlled-access Human Data	Japanese Genotype-phenotype Archive (JGA)	European Genome-phenome Archive (EGA)	database of Genotypes and Phenotypes (dbGaP)

国際的な研究データの共有（制限公開データ）

DACが存在しないため、NBDCがその役割を担う

DDBJ
JGA

NIHの各Instituteの
DACが審査

NCBI
dbGAP

各データ提供機関が
指定するDACが審査

EBI
EGA

目次

1. 背景

2. NBDCHITデータベース群の紹介

3. TogoVarの紹介

『NBDCヒトデータベース群』の役割

- 公的資金によるヒト関連研究の成果の受け皿
- 論文に記載可能なアクセッション番号を発番

- 個人由来のデータ共有と個人情報保護の両立を可能にする（匿名化したデータ共有）

NBDCグループ共有データベース
では発番されず、公開系DBへ移行の際に発行されます。

- 匿名化されたデータ共有
- ガイドラインに基づき、NBDCに設置された Data Access Committee (DAC) が実施

※ 1 : 個人情報保護に関する法律、人を対象とする生命科学・医学系研究に関する倫理指針、など

『NBDCヒトデータベース群』の役割

- 公的資金によるヒト関連研究の成果の受け皿
- 論文に記載可能なアクセッション番号を発番
- 個人由来のデータ共有と個人情報保護の両立を可能に（日本の法令※₁に沿ったデータ共有）
 - データ共有のためのガイドラインを策定
 - **NBDCヒトデータ共有ガイドライン**
 - **NBDCヒトデータ取扱いセキュリティガイドライン**
 - ガイドラインに沿って申請内容の審査をNBDCに設置されたData Access Committee (DAC) が実施

※ 1 : 個人情報の保護に関する法律、人を対象とする生命科学・医学系研究に関する倫理指針、など

NBDCヒトデータベース群でのデータ共有状況 (2022/9/14時点)

【NBDCヒトデータベース】

データ提供申請：413件 + データ更新・追加申請：396件

>40万人のデータが格納

公開：242 Studies（制限公開：207、非制限公開：54）

データ利用申請：273件

【NBDCグループ共有データベース】

□ AGD

データ提供申請：15件 + データ更新・追加申請：4件

13,249名のデータが格納（内、3415名は公開系へ移行済み）

サイト公開（日本語のみ）：9件 ➡ 3件は全データ公開系へ移行済み

データ利用申請：3件

□ SHD

データ提供申請：1件 + データ更新・追加申請：2件

642名のデータが格納

データ利用申請：11件

データの利用にあたって遵守すべき基本的事項

- ✓ データ利用者の限定（申請された研究代表者および研究代表者と同一機関に所属する研究分担者、および受託者に限る）
- ✓ 利用目的の明示
- ✓ 申請した利用目的以外への使用の禁止
- ✓ 研究・開発利用への限定
- ✓ 武器開発・軍事への利用禁止
- ✓ 個人同定の禁止
- ✓ 再配布の禁止

+ 制限事項

設定可能な制限事項

✓ 同意内容の反映

- ・研究で得られた成果は、他のがん研究にも利用させていただきます。
➡がん研究に限定

✓ 事業等におけるデータ共有方針

- ・AMEDの『ゲノム医療実現のためのデータシェアリングポリシー』
➡NBDCのガイドラインとの調整を行い、“別表”として異なる点を示す

データ利用のポイント

- ✓ 利用を希望するデータに関係した研究の経験があること。
- ✓ 学術研究もしくは公衆衛生の向上に貢献する研究であること。
- ✓ NBDCHITDBに登録されたデータを研究に使用することについて記載された研究計画書が、倫理審査委員会の承認、および所属機関の長の許可を受けていること。
- ✓ 申請するデータの制限レベルに適したセキュリティ対策がなされていること（NBDCHITデータ取扱いセキュリティガイドラインの遵守）。
- ✓ 利用者要件（データセットに付与された制限事項） を満たすこと。

論文に記載されるアクセッション番号

【制限公開データ】

- ・JGAS + 6桁の数字、JGAD + 6桁の数字

【非制限公開データ】

- ・NGS配列データ：DRA + 6桁の数字
- ・発現データ：GEAD + 3桁の数字
- ・メタボロミクスデータ：MTBKS + 自然数
- ・GWAS統計など：hum + 4桁の数字



NBDCヒトデータベース

(<https://humandbs.biosciencedbc.jp/>) へ！

アクセッション番号の見つけ方

NBDCヒトデータベースについて

ヒトに関するデータは、次世代シーケンサーをはじめとした解析技術の発達に伴って膨大な量が産生されつつあり、それらを整理・格納して、生命科学の進展のために有効に活用するためのルールや仕組みが必要です。

国立研究開発法人科学技術振興機構(JST)バイオサイエンスデータベースセンター(NBDC)では、個人情報の保護に配慮しつつヒトに関するデータの共有や利用を推進するために、ヒトに関する様々なデータを共有するためのプラットフォーム『NBDCヒトデータベース』を設立するとともに、国立遺伝学研究所 DNA Data Bank of Japan (DDB)と協力して、ヒトに関するデータを公開しています。

本Webサイトを通じて、ヒトに関するデータの利用及びヒトに関するデータの提供を行なうことができます。

なお、本データベースの目的・意義、扱うデータの種類、データ利用者の範囲、責任者についてはこちらをご覧ください。

新着情報

- 2022/01/25
千葉県がんセンター研究所からの制限公開データ (Type I) を公開しました (hum0320)
- 2022/01/25
制限公開データ (Type I) 1件が追加されました (hum0014.v25)

▶ ニュース一覧へ

Search NBDC Human Database Beacon for Alternative Alleles [\[API help\]](#)

NBDC Human Database Beacon is a member of GA4GH Beacon Network.

GRCh37 Example: ALDH2 Variant (GRCh37, 12:112241766 A)

利用可能な研究データ一覧

データ利用方法はこちらをご覧ください。

全 204 件

一覧内検索:

Research ID	研究題目	公開日	データの種類	研究方法	手法	参加者 (対象集団)	提供者	アクセス制限
hum0323.v1 JGA S000489	大腸癌の診断マーカーの探索研究	v1:2022/01/24	NGS (CAGE-seq)	転写開始点 同定	Illumina (HiSeq 2500)	大腸癌: 45症例 (日本人)	向後 泰司	制限 (Type I)
hum0320.v1 JGA S000480	悪性腫瘍を有する患者を対象とした免疫モニタリング研究	v1:2022/01/25	NGS (scRNA-seq, TCR-seq)	発現	MGI (DNBSEQ-G400)	肺腺癌: 2症例 頭頸部癌: 3症例 (日本人)	富樫 庸介	制限 (Type I)
hum0319.v1 JGA S000487	遺伝子多型を用いた前立腺癌診断マーカーの開発 日本人の前立腺がん発症に関わるレアバリアントの同定	v1:2022/01/21	NGS (Target Capture)	配列決定	Illumina (HiSeq 2500)	前立腺癌の確定診断のために前立腺生検を行う患者: 1,336症例 (日本人)	赤松 秀輔	制限 (Type I)

疾患名や解析手法などで検索

NBDCヒトデータベースについて

ヒトに関するデータは、次世代シーケンサーをはじめとした解析技術の発達に伴って膨大な量が産生されつつあり、それらを整理・格納して、生命科学の進展の
国立研究開発法人科学技術振興機構(JST)バイオサイエンスデータベースセンター(NBDC)では、個人情報の保護に配慮しつつヒトに関するデータの共有や利用を推
プラットフォーム『NBDCヒトデータベース』を設立するとともに、国立遺伝学研究所 DNA Data Bank of Japan (DDB)と協力して、ヒトに関するデータを公開
本Webサイトを通じて、ヒトに関するデータの利用及びヒトに関するデータの提供を行なうことができます。
なお、本データベースの目的・意義、扱うデータの種類、データ利用者の範囲、責任者についてはこちらをご覧ください。

新着情報

疾患名 (ICD10疾患分類コードでの検索も可能)、解析手法、研究費番号、研究グループ名、などによるフリーテキスト検索

Search NBDC Human Database Beacon for Alternative Alleles (API help)

NBDC Human Database Beacon is a member of GA4GH Beacon Network.

GRCh37 e.g. 12:112241766 A Search Example: ALDH2 Variant (GRCh37, 12:112241766 A)

利用可能な研究データ一覧

データ

全 204

Copy CSV Excel

一覧をDLして手元でフィルタリング可能

一覧内検索:

Research ID	研究題目	公開日	データの種類	研究方法	手法	参加者 (対象集団)	提供者	アクセス制 限
hum0323.v1 JGA S000489	大腸癌の診断マーカーの探索研究	v1:2022/01/24	NGS (CAGE-seq)	転写開始点 同定	Illumina (HiSeq 2500)	大腸癌: 45症例 (日本人)	向後 泰司	制限 (Type I)
hum0320.v1 JGA S000480	悪性腫瘍を有する患者を対象とした免疫モニタリング研究	v1:2022/01/25	NGS (scRNA-seq, TCR-seq)	発現	MGI (DNBSEQ-G400)	肺腺癌: 2症例 頭頸部癌: 3症例 (日本人)	富樫 庸介	制限 (Type I)
hum0319.v1 JGA S000487	遺伝子多型を用いた前立腺癌診断マーカーの開発 日本人の前立腺がん発症に関わるレアバリアントの同定	v1:2022/01/21	NGS (Target Capture)	配列決定	Illumina (HiSeq 2500)	前立腺癌の確定診断のために前立腺生検を行う患者: 1,336症例 (日本人)	赤松 秀輔	制限 (Type I)

利用したいデータが見つかったら・・・

制限公開データ (利用可能な研究データ一覧のアクセス制限欄に”制限”と表記。データの利用には利用申請が必要。)

== 2020年9月30日より、メールでの申請書類受付は終了し、申請システムによる手続きへ移行しました ==

※ 過去のデータ利用申請の際に添付いただいた倫理関係書類の取り込み中です。

利用申請方法は以下の通りです。

1. [NBDCヒトデータ共有ガイドライン](#)の最新版を熟読し、権利や責務、利用者要件について確認してください。
2. [利用可能な研究データ一覧](#)から各研究内容を確認 (Research IDをクリック) のうえ、利用したいデータセットのDataset IDを控えてください。データ利用申請時に必要になります。その際、データセット毎に付与されている利用者要件 (利用にあたっての制限事項に記載されているPolicyの内容) に合致するか確認してください。1回の申請で、複数のDatasetを申請することが可能です。
3. データ利用申請のために必要な以下の情報を準備してください。
 - 研究代表者およびデータ利用を希望する研究分担者の情報 (氏名、職名、メールアドレス、研究者ID、倫理講習受講の有無)、ならびに解析等受託先がある場合は受託者の情報 (所属機関、決まっている場合は担当者名・職名・メールアドレス・研究者ID・倫理講習受講の有無) ※同一機関に所属する研究分担者、ならびに受託者を1回の申請で登録することが可能です。
 - 研究代表者によって利用申請される研究内容に関連した研究に従事したことがわかるエビデンス (論文や学会発表等の情報)
 - 所属機関の長の氏名、職名、メールアドレス ※所属機関の長とは、倫理審査委員会によって承認された研究計画の実施を許可する者を指します。
 - 取得したデータを使用する研究の研究計画書 (倫理審査申請書など)
 - 所属機関等の倫理審査委員会の審査・承認を受けた上で、所属機関の長が研究実施について許可したことが分かる書面 (承認通知書など) の写し
4. [NBDCヒトデータ取扱いセキュリティガイドライン \(データ利用者向け\)](#) の内容を確認し、使用したいデータのアクセス制限レベル (Type IまたはType II) に応じて必要な対策を実施してください。[NBDCヒトデータ取扱いセキュリティガイドラインチェックリスト](#)をダウンロードし、実施状況を記載してください。データの保管・解析に機関外サーバ^{*1}の利用を予定されている場合は、各機関外サーバ運用機関のサイトから公開されているセキュリティチェックリスト (機関外サーバ運用責任者向け) を入手した上で、当該機関外サーバがNBDCヒトデータ取扱いセキュリティガイドラインを満たしていることを確認してください。
5. データ利用申請には、研究代表者、ならびに、研究分担者の内、研究代表者によって指名された申請者 (研究代表者に代わってデータ利用申請手続きを実施する者) やデータダウンロード担当者のD-wayアカウントが必要ですので、[アカウント登録](#)を行ってください。所属機関の長のアカウントは不要です。

目次

1. 背景

2. NBDCHITデータベース群の紹介

3. TogoVarの紹介

TogoVar

- **日本人ゲノム多様性統合**データベース
genomic **Var**iations Togo
- NBDCヒトデータベースへ登録、公開されたデータについて、個人特定されない形（頻度情報）で提供。データセットの概要把握の手助けになることを目指す。
- **日本**や海外で公開されている頻度情報、**ゲノム多様性**と疾患との関連情報を**統合**、ワンストップで検索可能に。
- 2018年6月7日公開

TogoVar

多種多様なデータベースに散在して収録されてきたGenotypeやPhenotypeに関連する情報を整理統合し、バリエーションを解釈するための情報をワンストップで提供

7番染色体



▲ 注目するバリエーション

ClinVar (NCBI)

位置 : chr7:127254587
関連する疾患 : 2型糖尿病
疾患感受性 : あり

gnomAD (ブロード研究所)

位置 : chr7:127254587
アレル頻度 : 0.000304573

TogoVarID: tgv30913364

位置 : chr7:127254587
関連する疾患 : 2型糖尿病
疾患感受性 : あり
アレル頻度(iJGVD 3.5KJPN) : 0.0233
アレル頻度(HGVD) : 0.0272809
アレル頻度(gnomAD) : 0.000304573
関連論文 : A missense mutation of Pax4 gene ...

<https://togovar.biosciencedbc.jp/variant/tgv30913364>



iJGVD 3.5KJPN
(東北メディカル・メガバンク機構)

位置 : chr7:127254587
アレル頻度 : 0.0233

HGVD(京都大学)

位置 :
chr7:127254587
アレル頻度 : 0.0272809

↑ ワンストップ検索



TogoVar

TOGOVAR A comprehensive Japanese genetic variation database

Home Datasets Downloads Terms Contact About History Help Configuration

Search for disease or gene symbol or rs...

Disease: Breast-ovarian cancer, familial 2 Gene: ALDH2 refSNP: rs114202595 TogoVar: tgv421843 HGVS: NM_000690:c.1510G>A HGVS: p.ALDH2:p.Glu504Lys Position (GRCh37/hg19): 16:48258198 Region (GRCh37/hg19): 10:73270743-73376976

Results The number of available variations is 10,000 out of 320,351,666.

TogoVar ID	RefSNP ID	Position	Ref / Alt	Type	Gene	Alt frequency	Consequence	SIFT	PolyPhen	Clinical significance
	rs1457723673	1: 10129	CCCT... 19bp	Deletion			Intergenic variant			
tgv83272255	rs1385251551	1: 10135	CCCT... 13bp	Deletion			Intergenic variant			
tgv122011874	rs1557426741	1: 10141	CCCT... 7bp	Deletion			Intergenic variant			
tgv122011875		1: 10144	TAAC... 8bp	Deletion			Intergenic variant			
tgv83272256	rs144773400	1: 10145	A	Deletion			Intergenic variant			
	rs1557426742	1: 10145	A > T	SNV			Intergenic variant			
tgv122011876		1: 10146	AC >	Deletion			Intergenic variant			
tgv67071948	rs779258992	1: 10147	C	Deletion			Intergenic variant			
	rs1557426745	1: 10147	CCTA... 5bp	Insertion			Intergenic variant			
tgv83272257	rs1286868604	1: 10150	CT	Deletion			Intergenic variant			
tgv122011877	rs371194064	1: 10150	C > T	SNV			Intergenic variant			
tgv83272259		1: 10151	T	Deletion			Intergenic variant			
tgv83272260	rs1557426747	1: 10152	A	Deletion			Intergenic variant			
tgv122011879		1: 10153	C	Insertion			Intergenic variant			
tgv122011880	rs1331781659	1: 10153	A > G	SNV			Intergenic variant			
tgv122011882	rs1264289758	1: 10156	CTAA... 22bp	Deletion			Intergenic variant			
tgv83272261	rs1487252449	1: 10157	TAAC... 24bp	Deletion			Intergenic variant			
tgv122011883	rs1570391843	1: 10157	T > C	SNV			Intergenic variant			
tgv83272262		1: 10157	T > G	SNV			Intergenic variant			
	rs1557426750	1: 10158	AACC... 5bp	Deletion			Intergenic variant			

Filters Reset

Dataset

- All 320,351,666
- WGS GEM-J WGA 81,560,747
- WES JGA NGS 4,626,776
- SNP JGA SNP 1,249,724
- WGS ToMMo 8.3KJPN 86,019,804
- WES HGVD 554,461
- WGS gnomAD genomes 232,877,353
- WES gnomAD exomes 15,199,750
- Disease ClinVar 773,085

Alternative allele frequency

0 ~ 1 Invert range

0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

for all datasets for any dataset

Variant calling quality

Exclude filtered out variants in all datasets

Variant type

- All 320,351,666
- SNV 283,830,077
- Insertion 14,496,129
- Deletion 21,670,386
- Indel 98,668
- Substitution 256,406

togovar

検索

<https://togovar.biosciencedbc.jp/>



検索方法：Advanced Search

AND/ORを複雑に組み合わせた検索式をGUIで直感的に作成可能に！

例：集団間でAlternative Allele頻度に差があるバリエーションを検索

((frequency(GEM-J WGA) >= 0.3) **OR** (frequency(ToMMo 8.3KJPN) >= 0.3))

👉日本人集団で高頻度

AND

(frequency(gnomAD genomes) <= 0.05) **OR** (frequency(gnomAD exomes) <= 0.05))

👉世界中の集団の平均値で低頻度

The screenshot shows the TOGOVAR Advanced Search interface. The search conditions are:

- Alternative allele frequency/count: GEM-J WGA (0.3 ~ 1)
- Alternative allele frequency/count: ToMMo 8.3KJPN (0.3 ~ 1)
- Alternative allele frequency/count: gnomAD genomes (0 ~ 0.05)
- Alternative allele frequency/count: gnomAD exomes (0 ~ 0.05)

The results table shows the following data:

TogoVar ID	RefSNP ID	Position	Ref / Alt	Type	Gene	Alt frequency	Consequence	SIFT
lgy2158	rs201365517	1: 612688	CTC	Deletion			Intron variant	
lgy67086890	rs1553174279	1: 974792	GC	Insertion	AGRN		Intron variant	
lgy22068	rs57527288	1: 1092563	G T	SNV			Intergenic variant	
lgy23923	rs35247267	1: 1123284	AA	Deletion	ITLL10		Intron variant	
lgy30173	rs112572774	1: 1223626	T C	SNV	SCNN1D		Intron variant	
lgy30244	rs1977348	1: 1225579	G C	SNV	SCNN1D		Intron variant	
lgy30526	rs2273276	1: 1226292	G A	SNV	SCNN1D		Synonymous variant	
lgy31621	rs2297234	1: 1233580	G T	SNV	ACAP3		Non coding transcript exon variant	

Statistics:

Dataset	Count
All	30,350
WGS GEM-J WGA	27,846
WES JGA NGS	18,641
SNP JGA SNP	2,987
WGS ToMMo 8.3KJPN	30,341
WES HGVD	320
WGS gnomAD genomes	30,338
WES gnomAD exomes	646
Disease ClinVar	194

👉検索式ビルダー
(Add conditionから追加)

TogoVar

TOGOVAR A comprehensive Japanese genetic variation database

Home Datasets Downloads Terms Contact About History Help Configuration

Search for disease or gene symbol or rs...

Disease: Breast-ovarian cancer, familial 2 Gene: ALDH2 refSNP: rs114202595 TogoVar: tgv421843 HGVS: NM_000690:c.1510G>A HGVS: p.ALDH2:p.Glu504Lys Position (GRCh37/hg19): 16:48258198 Region (GRCh37/hg19): 10:73270743-7337697

Results The number of available variations

TogoVar ID	RefSNP ID	WES	JGA NGS	SNP	JGA SNP
	rs14577236	<input checked="" type="checkbox"/>	4,626,776	<input checked="" type="checkbox"/>	1,249,724
tgv83272255	rs13852515	<input type="checkbox"/>		<input type="checkbox"/>	
tgv122011874	rs15574267	<input type="checkbox"/>		<input type="checkbox"/>	
tgv122011875	rs14477340	<input type="checkbox"/>		<input type="checkbox"/>	
tgv83272256	rs15574267	<input type="checkbox"/>		<input type="checkbox"/>	
tgv122011876	rs15574267	<input type="checkbox"/>		<input type="checkbox"/>	
tgv67071948	rs779258992	<input type="checkbox"/>		<input type="checkbox"/>	
	rs1557426745	<input type="checkbox"/>		<input type="checkbox"/>	
tgv83272257	rs1286868604	<input type="checkbox"/>		<input type="checkbox"/>	
tgv122011877	rs371194064	<input type="checkbox"/>		<input type="checkbox"/>	
tgv83272259	rs1557426747	<input type="checkbox"/>		<input type="checkbox"/>	
tgv83272260	rs1557426747	<input type="checkbox"/>		<input type="checkbox"/>	
tgv122011879	rs1557426747	<input type="checkbox"/>		<input type="checkbox"/>	
tgv122011880	rs1331781659	<input type="checkbox"/>		<input type="checkbox"/>	
tgv122011882	rs1264289758	<input type="checkbox"/>		<input type="checkbox"/>	
tgv83272261	rs1487252449	<input type="checkbox"/>		<input type="checkbox"/>	
tgv122011883	rs1570391843	<input type="checkbox"/>		<input type="checkbox"/>	
tgv83272262	rs1557426750	<input type="checkbox"/>		<input type="checkbox"/>	

Filters

Dataset

- All 320,351,666
- WGS GEM-J WGA 81,560,747
- WES JGA NGS 4,626,776
- WES JGA SNP 1,249,724
- WGS ToMMo 8.3KJPN 86,019,804
- WES HGVD 554,461
- WGS gnomAD genomes 232,877,353
- WES gnomAD exomes 15,199,750
- Disease ClinVar 773,085

Alternative allele frequency

0 ~ 1 Invert range

for all datasets for any dataset

Variant calling quality

Exclude filtered out variants in all datasets

Variant type

- All 320,351,666
- SNV 283,830,077
- Insertion 14,496,129
- Deletion 21,670,386
- Indel 98,668
- Substitution 256,406

togovar

検索

<https://togovar.biosciencedbc.jp/>

TogoVar検索例

NFKB1

Disease: Breast-ovarian cancer, familial 2 Gene: ALDH2 refSNP: rs114202595 TogoVar: tgv421843 HGVS: NM_000690:c.1510G>A HGSP: A Region: (GRCh37/hg19):10:73270743-73376976

Results The number of available variations is 27 out of 27.

TogoVar ID	RefSNP ID	Position	Ref / Alt	Type	Gene	Alt frequency	Consequence	SIFT	PolyPhen	Clinical significance
tgv18081589	rs1609993	4: 103514658	T > C	SNV	NFKB1		Synonymous variant		0	B not provided
tgv18081608	rs4648049	4: 103514737	C > T	SNV	NFKB1		Intron variant		0	B not provided
tgv18081610	rs4648050	4: 103514741	T > C	SNV	NFKB1		Intron variant		0	B not provided
tgv18081611	rs192833592	4: 103516041	C > T	SNV	NFKB1		Splice region variant		0	B not provided
tgv18081612	rs1648065	4: 103517463	C > T	SNV	NFKB1		Missense variant	0.35	0	B not provided
tgv18081613	rs1648065	4: 103517463	C > T	SNV	NFKB1		Missense variant	0.09	0.031	B Immunodeficiency, common variable, 12
tgv18081614	rs1648065	4: 103517463	C > T	SNV	NFKB1		Missense variant	0.09	0.006	B not provided
tgv18081615	rs1648065	4: 103517463	C > T	SNV	NFKB1		Missense variant	0.239	0	B Immunodeficiency, common variable, 12
tgv18081616	rs1648065	4: 103517463	C > T	SNV	NFKB1		Missense variant	0.003	0	B Immunodeficiency, common variable, 12
tgv18081617	rs1648065	4: 103517463	C > T	SNV	NFKB1		Missense variant	0.152	0	B not provided

タンパク質機能への影響予測

疾患との関連

個々のバリエーション情報へのリンク

Variant report: tgv18081608 rs4648049

Variant type: SNV Ref / Alt: C > T Position: 4:103514737 (GRCh37) figs: 4:g.103514737C>T

Other overlapping variants: No other variants on the same location

Dataset	Population	alt count	Total	Frequency	Genotype count	Filter	Quality
		Alt	Total		Alt / Ref / Alt		status score
GEM-J WGA	Japanese	574	15,164	0.038		PASS	208.799
JGA NGS	Japanese	13	250	0.052		PASS	524.77
ToMMo 8.3KJPN	Japanese	668	16,760	0.040		PASS	-
HGVD	Japanese	81	1,674	0.048		PASS	-
gnomAD Genomes	Total	2,073	31,352	0.066		PASS	947002.1
	African American/African	998	8,698	0.115		-	-
	Ashkenazi Jewish	2	288	0.007		-	-
	East Asian	87	1,554	0.056		-	-
	Finnish	208	3,462	0.060		-	-
	Latino	27	848	0.032		-	-
	Non-Finnish European	697	15,414	0.045		-	-
	Other	54	1,088	0.050		-	-
gnomAD Exomes	Total	1,217	243,016	0.046		PASS	15641394

DB毎の頻度 (NBDCヒトDB登録データ含む)



Preview

Detailed variant report page

Genes

Symbol: NFKB1
Alias: KBF1, p105, NFKB-p50, p50, NF-kappaB, NFkappaB, NF-kB1

External links

refSNP: rs4648049
ClinVar: VCV001168687

Alternative allele frequencies

Dataset	Alt	Total	Frequency
WGS GEM-J WGA	574	15,164	3.80e-2
WES JGA NGS	13	250	5.20e-2
SNP JGA SNP	/	/	/
WGS ToMMo 8.3KJPN	668	16,760	3.99e-2
WES HGVD	81	1,674	4.84e-2
WGS gnomAD genomes	2,073	31,352	6.61e-2
WES gnomAD exomes	1,217	243,016	4.62e-2

TogoVar検索例

NFKB1

Disease: Breast-ovarian cancer, familial 2 Gene: ALDH2 refSNP: rs114202595 TogoVar: tgv421843 HGVS: NM_000690:c.1510G>A HGVSp: ALDH2:p.Glu504Lys Position(GRCh37/hg19):16:48258198 Region(GRCh37/hg19):10:73270743-73376976

Results The number of available variations is 27 out of 27.

TogoVar ID	RefSNP ID	Position	Ref / Alt	Type	Gene	Alt frequency	Consequence	SIFT	PolyPhen	Clinical significance
tgv18081589	rs1609993	4: 103514658	T > C	SNV	NFKB1		Synonymous variant			(B) not provided
tgv18081608	rs4648049	4: 103514737	C > T	SNV	NFKB1		Intron variant			(B) not provided
tgv18081610	rs4648050	4: 103514741	T > C	SNV	NFKB1		Intron variant			(B) not provided
tgv18081641	rs192833592	4: 103516041	C > T	SNV	NFKB1		Splice region variant			(B) not provided
tgv18081713	rs4648065	4: 103517463	C > T	SNV	NFKB1		Missense variant	0.35 (+)	0 (+)	(B) not provided
tgv18081760	rs4648072	4: 103518700	A > G	SNV	NFKB1		Missense variant	0.09 (+)	0.031 (+)	(B) Immune system variable, 12 (+)

Preview

Detailed variant report page

Genes

Symbol: [NFKB1](#)
 Alias: KBF1, p105, NFKB-p50, p50, NF-kappaB, NFkappaB, NF-kB1

External links

refSNP: [rs4648049](#)
 ClinVar: [VCV001168687](#)

Alternative allele frequencies

Dataset	Alt	Total	Frequency
WGS GEM-J WGA	574	15,164	3.80e-2
WES JGA NGS	13	250	5.20e-2
SNP JGA SNP	/	/	/
WGS ToMMo	668	16,760	3.99e-2
WES HGVD	81	1,674	4.84e-2
WGS gnomAD genomes	2,073	31,352	6.61e-2
WES gnomAD exomes	11,217	243,016	4.62e-2

遺伝子情報へのリンク

疾患情報へのリンク

TogoVarのAPIも公開しちゃいます！

[Swagger](#)を利用したAPIドキュメント上で検索例を実行可能

TogoVar API ^{0.9.1} ^{OAS3}

TogoVar API version 0.9.1 specification

Servers

<https://togovar.biosciencedbc.jp/api>

search

POST /search/variant

GET /search/gene

POST /search/gene

GET /search/disease

POST /search/disease

👉 手順1. 検索方法を選択する。

- バリエントの検索
- 遺伝子
- 疾患

search

POST /search/variant

Parameters

Name	Description
------	-------------

pretty	If set, response body will be nicely formatted.
--------	---

boolean	
(query)	<input type="text" value="--"/>

👉 手順4. その検索条件に対して Trueかfalseかを選択可能

Request body

Schema for searching variant

Examples:

List variants by identifiers (tgv, rs)

List variants whose class is any of insertion, deletion and indel.

List variants by identifiers (tgv, rs)

List variants exist on an exact position

List variants exist in a certain range

List variants range from 0.05 to 0.95 in frequency in JGA-NGS.

List variants exist on ALDH2 (Aldehyde Dehydrogenase 2 family member).

List variants associated with C2675520 (Breast-ovarian cancer, familial 2).

List variants associated with MONDO_0003582 (hereditary breast ovarian cancer syndrome).

paginate result over 10,000 records.

手順2. 「Try it out」をクリック

👉 手順3. 何で検索するかを選択
ここではtgv, rsidで検索するクエリ例を選択

TogoVarのAPIも公開しちゃいます！

[Swagger](#)を利用したAPIドキュメント上で検索例を実行可能

TogoVar API 0.9.1 OAS3
TogoVar API version 0.9.1 specification

Servers

search

POST /search/variant

Parameters

Name	Description
pretty	If set, response body will be nicely formatted.

pretty
boolean
(query)

Request body

Schema for searching variant

Examples:
List variants by identifiers (tgv, rs)

```
{
  "query": {
    "id": [
      "tgv421843",
      "rs114202595"
    ]
  }
}
```

Execute

👉 手順 5. Executeボタンを押下

Responses

Curl

```
curl -X 'POST' \
  'https://togovar.biosciencedbc.jp/api/search/variant' \
  -H 'accept: application/json' \
  -H 'content-type: application/json' \
  -d '{
  "query": {
    "id": [
      "tgv421843",
      "rs114202595"
    ]
  }
}'
```

curlコマンドの例

Request URL

Server response

Code Details

200

Response body

```
{
  "data": [
    {
      "id": "tgv421843",
      "type": "SO_0001483",
      "chromosome": "1",
      "position": 1277989,
      "start": 1277984,
      "stop": 1277985,
      "reference": "T",
      "alternate": "C",
      "existing_variations": [
        "rs180765"
      ],
      "symbols": [
        {
          "name": "FRANK77",
          "id": 2812,
          "synonyms": []
        }
      ]
    }
  ]
}
```

検索結果はJSONで返される

Response headers

```
access-control-allow-methods: GET,POST,HEAD
access-control-allow-origin: *
access-control-expose-headers:
www-authenticate: WWW-Authenticate: Basic
```

まとめ

- 個人毎の臨床情報やヒト由来試料の解析データの共有時には、国内の法令や倫理指針等の遵守が必要。
- 研究対象者へのインフォームドコンセントの際の同意内容に沿った共有が必要であり、利用者要件として「制限事項」に示される。
- 制限公開データならびにグループ共有データは、データ利用申請が必要。
- TogoVarでは、今後も利便性の向上のための仕組みを検討し、開発を続ける。また、他のデータベース等との連携も計画している。

ご清聴いただき、誠にありがとうございました。

ご不明点がありましたら、NBDCヒトデータ審査委員会事務局までお問い合わせください。

電話：03-5214-8491

メール：humandbs@biosciencedbc.jp

