

TogoVarを使ってヒトゲノムに存在する バリアントに関する情報を調べる

豊岡 理人

バイオサイエンスデータベースセンター(NBDC)
国立研究開発法人 科学技術振興機構

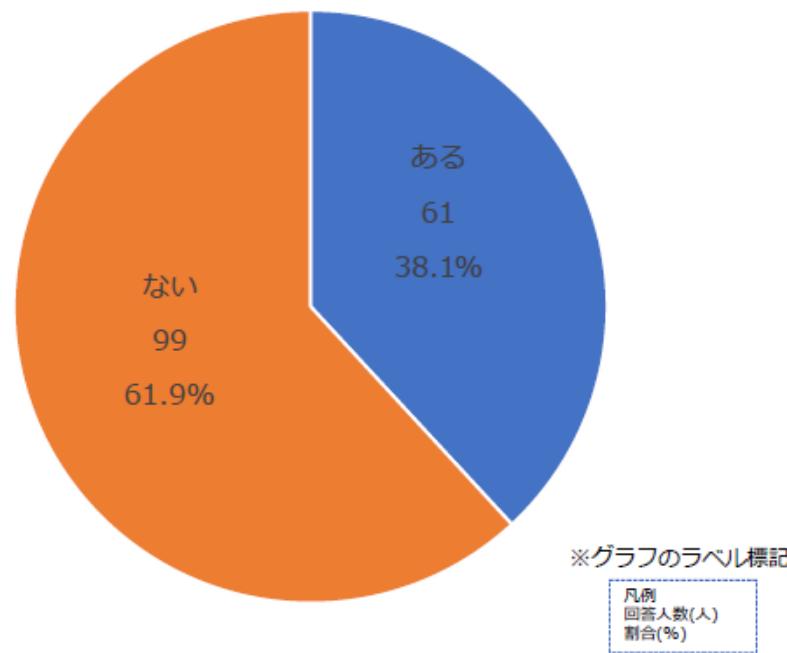
2021年12月16日
AJACS オンライン9



事前アンケート

Ⅱ.各講義別 質問事項への回答 ※任意回答

【講義1-1】バリエント(dbSNP/gnomAD/jMorp/HGVD/TogoVar等)のデータベースを利用したことがありますか？



TogoVar以外のデータベースについても紹介します。

多型/バリアントとは

- ・ 多型(polymorphism) / バリアント (variant)
 - ヒトゲノムにおいて観察されるreferenceと異なる塩基配列
- ・ セントラルドグマ
ゲノム → トランスクリプト → 蛋白質 → 形質
(疾患、薬剤副反応等)
- ・ ゲノムと形質との関連を調べる研究において重要な基礎データ

- 一塩基多型

Single Nucleotide Polymorphism: SNP

Single Nucleotide Variation: SNV

- コピー数多型

Copy Number Variation: CNV

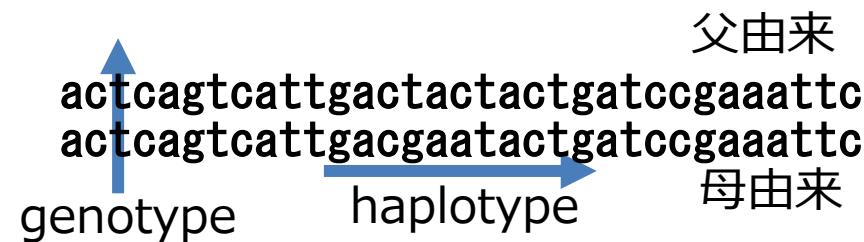
- タンデムリピート、トランスポゾン、レトロトランスポン

- 構造多型

Structural Variation: SV

- サンガーシークエンサー

- 配列情報



- SNP Chip

- 60-90万箇所のgenotype情報
 - haplotype情報は推定
 - 安価

- 次世代シークエンサー (Next Generation Sequencer: NGS)

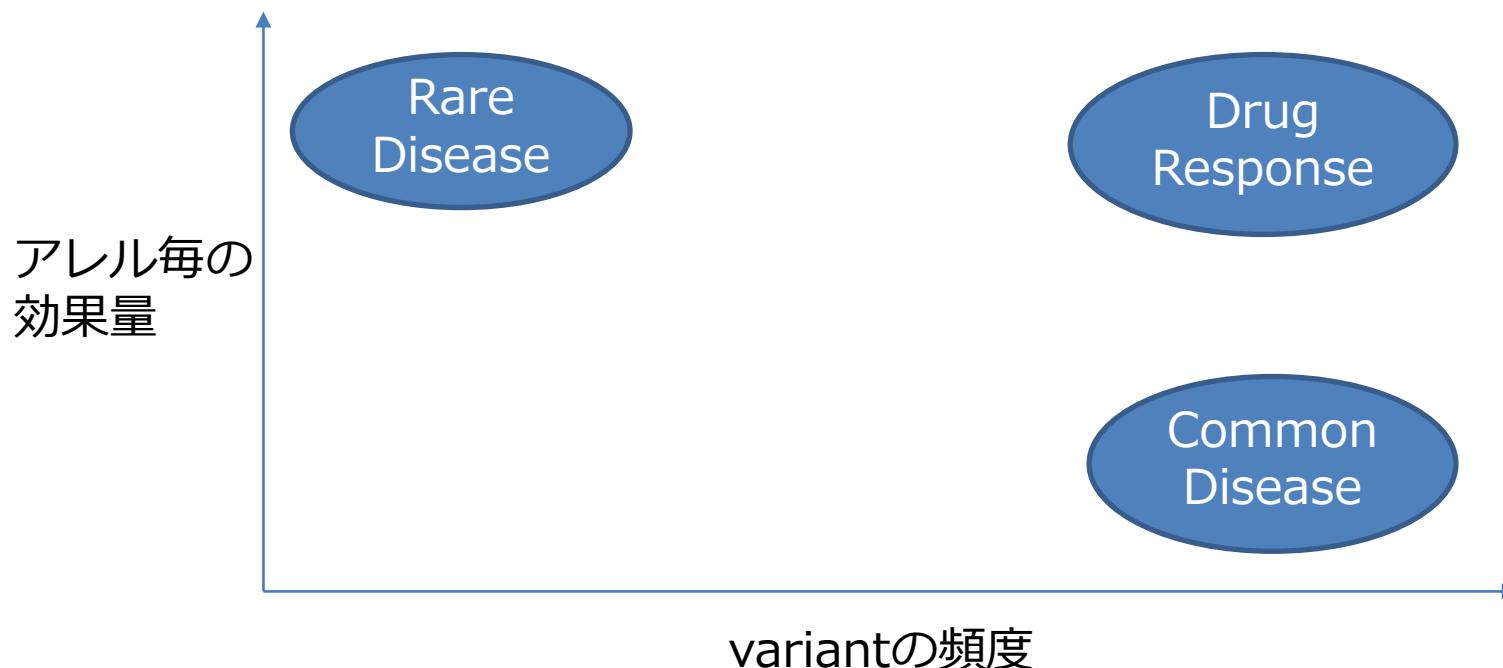
- genotype情報、haplotype情報
 - 全exome領域や全genome領域などあらゆる領域が検知対象

- NGS (long read)

- genotype情報、haplotype情報
 - 全exome領域や全genome領域などあらゆる領域が検知対象

バリエントの頻度とアレルごとの効果量 および形質の関係

- Common Disease Common Variant仮説



Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nature Reviews Genetics* 9, 367 (2009) より改変

疾患と関連する遺伝子同定の研究における バリアント頻度情報の利用

- 対象疾患：ありふれた疾患(common disease)
RareのバリアントはPrimaryの探索から外す。
- 対象疾患：稀少疾患(rare disease)
CommonのバリアントはPrimaryの探索から外す。
 - Trio解析
 - 発症者（子）+未発症（父母）の3人から発症者におけるde novoなvariantを抽出
 - Trioではde novo のvariantであっても、既知のvariantは探索対象から除外可能

一般集団のvariant頻度情報は基礎的情報として有用

バリアントデータベースの使い方

- 把握すべきバリアントの情報
 - rs番号
 - 新しいrs番号が割り振られることもある
 - HGVS表記
 - NM_000690:c.1510G>A
 - ポジションでの表記
 - chr16:48258198
 - 注意！） 参照ゲノム配列を確認
 - hg18, hg19=GRCh37, GRCh38

バリアントデータベースの種類

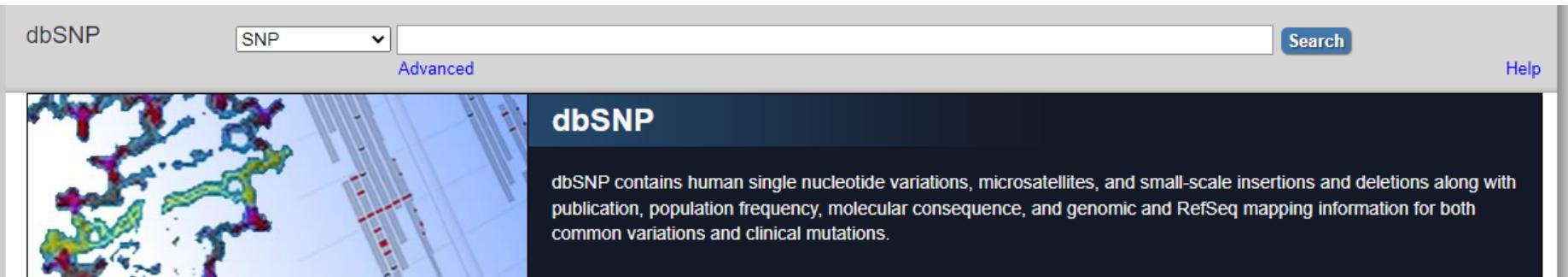
主たる情報	説明	例
頻度	集団に存在するバリアントの頻度情報を提供	dbSNP, gnomAD, jMorp, HGVD, TogoVar
GWASの解析結果	GWASの結果を提供 GWAS-catalogはp値 1×10^{-5} 以下ののみ掲載	GWAS catalog, Jenger, jMorp
疾患や形質との関連	疾患や形質と関連が報告されたバリアント情報を提供	ClinVar, MGeND
eQTLの結果	遺伝子発現に関連するバリアントの情報を提供 バリアント別組織別遺伝子別にeQTLの結果を提供	GTEEx

頻度データベースの紹介 : dbSNP

dbSNP

- 運営元： 米国NCBI
- URL: <https://www.ncbi.nlm.nih.gov/snp/>
- 特徴
 - 様々の研究において検知された多型/バリアントの情報を受け付け、ss番号を付与。その後、ss番号をまとめてrs番号を付与
 - ss番号 → rs番号
 - 50bp以下の多型を対象に情報を収集
 - 50bpよりも大きい多型 dbVarに収載

dbSNPの画面および操作



dbSNP

SNP Advanced Help

dbSNP

dbSNP contains human single nucleotide variations, microsatellites, and small-scale insertions and deletions along with publication, population frequency, molecular consequence, and genomic and RefSeq mapping information for both common variations and clinical mutations.

Getting Started

- [dbSNP 20th Anniversary](#)
- [Overview of dbSNP](#)
- [About Reference SNP \(rs\)](#)
- [Factsheet](#)
- [Entrez Updates \(May 26, 2020\)](#)

Submission

- [How to Submit](#)
- [Hold Until Published \(HUP\) Policies](#)
- [Submission Search](#)

Access Data

- [Web Search](#)
- [eUtils API](#)
- [Variation Services](#)
- [FTP Download](#)
- [Tutorials on GitHub](#)

i ALFA Project Release 2 with over 900M variants from 192K subjects is now [available](#) (January 6, 2021)
The goal is to provide allele frequency from more than 1 million dbGaP subjects with regular updates. Visit the project [page](#) for more information or view the introduction video below. Please provide your feedback by completing this short 3 min [survey](#).

rs番号やHGVS表記を入力可能。
Advancedをクリックすると、様々な条件での検索が可能。

NBDC dbSNPの検索結果(*PDCD1* rs11568821)

NCBI Resources How To Sign in to NCBI

dbSNP SNP PDCD1 rs11568821 Create alert Advanced Help

Annotation Cited in PubMed PubMed nucleotide Global MAF Custom range... Validation Status by-cluster by-frequency Clear all Show additional filters

Display Settings: Summary, Sorted by SNP_ID Send to: Filters: Manage Filters

Find related data Database: Select Find items

検索式の表示

Search details PDCD1 [All Fields] AND rs11568821 [All Fields] Search See more...

過去の検索内容の表示

recent activity Turn Off Clear PDCD1 rs11568821 (2) SNP PDCD1 (3020) SNP 12 : 111803962 (5) SNP rs671 (5) SNP See more...

検索候補が表示

rs11568821 [*Homo sapiens*] Variant type: SNV Alleles: C>G,T Chromosome: NT_187527.1:63858 Gene: LOC105373977 (GeneView), PDCD1 (GeneView) Functional Consequence: downstream_transcript_variant,genic_downstream_transcript_variant,upstream_transcript_variant,intron_variant Clinical significance: risk-factor Validated: by frequency,by cluster T=0.0409/205 (1000Genomes) T=0.0711/2196 (GnomAD) T=0.0746/9373 (TOPMED) T=0.1157/446 (ALSPAC) T=0.1162/431 (TWINSUK) HGVS: NC_000002.12:g.241851760C>G, NC_000002.12:g.241851760C>T, NC_000002.11:g.242793912C>G, NC_000002.11:g.242793912C>T, NG_012110.1:g.12147G>C, NG_012110.1:g.12147G>A, NT_187527.1:g.63858C>G, NT_187527.1:g.63858C>T PubMed Varview

rs606231173 has merged into rs11568821 [*Homo sapiens*] Variant type: SNV Alleles: C>G,T Chromosome: NT_187527.1:63858 Gene: LOC105373977 (GeneView), PDCD1 (GeneView)

dbSNPの検索結果（詳細画面）

dbSNP Short Genetic Variations

Search for terms

Search

Examples: rs288, BRCA1 and more

Advanced search



Welcome to the Reference SNP (rs) Report

All alleles are reported in the Forward orientation. Click on the Variant Details tab for details on Genomic Placement, Gene, and Amino Acid changes. HGVS names are in the HGVS tab.

Reference SNP (rs) Report

[Switch to classic site](#)
 [Download](#)
 [f](#)
 [t](#)
 [g](#)
 [?](#)

rs11568821

Organism	Homo sapiens
Position	chr2:241851760 (GRCh38.p13) ?
Alleles	C>A / C>G / C>T
Variation Type	SNV Single Nucleotide Variation
Frequency	T=0.076843 (10769/140142, GnomAD) T=0.07461 (3349/44884, ALFA) T=0.0409 (205/5008, 1000G) (+9 more)

Clinical Significance	Reported in ClinVar
Gene : Consequence	PDCD1 : Intron Variant LOC105373977 : 2KB Upstream Variant
Publications	39 citations LitVar 66
Genomic View	See rs on genome

[ClinVarへのリンク](#)

FEEDBACK

Variant Details

Clinical Significance

Frequency

HGVS

Submissions

History

Publications

Flanks

Genomic Placements

Sequence name	Change
GRCh37.p13 chr 2	NC_000002.11:g.242793912C>A
GRCh37.p13 chr 2	NC_000002.11:g.242793912C>G
GRCh37.p13 chr 2	NC_000002.11:g.242793912C>T
GRCh38.p13 chr 2	NC_000002.12:g.241851760C>A
GRCh38.p13 chr 2	NC_000002.12:g.241851760C>G
GRCh38.p13 chr 2	NC_000002.12:g.241851760C>T
GRCh38.p13 chr 2 alt locus HSCHR2_3_CTG15	NT_187527.1:g.63858C>A
GRCh38.p13 chr 2 alt locus HSCHR2_3_CTG15	NT_187527.1:g.63858C>G

旧表示形式
での表示も可能

参照配列毎の位置
ClinVarの情報
頻度情報
Submission情報
rs番号の変遷
文献情報
をTabで切り替え

<https://biosciencedbc.jp/>

- ALFA(Allele Frequency Aggregator)
 - dbGaPに格納されたデータを再解析
 - 79 Studies 192,710 Subjects
 - 2,515 East Asian
- 8.3KJPN
 - 東北メディカルメガバンクの収集したWGSの解析結果である
頻度情報がFrequency欄に8.3KJPNとして表示
 - 日本人 約8,300 Subjects

頻度データベースの紹介： gnomAD

gnomAD

- 運営元： ブロード研究所
- URL: <https://gnomad.broadinstitute.org/>
- ExACの後継プロジェクト
 - ExAC
 - 60,706名のWESデータ
 - v2.1 (リファレンスはGRCh 37)
 - 125,748名のWESデータ, 15,708名のWGSデータ
 - 日本人は76 subjects
 - v3 (リファレンスはGRCh 38)
 - 76,156名のWGSデータ
- 同じ解析プロトコールでジョイントコール

gnomAD browser

Search by gene, region, or variant

About Downloads Terms Contact Jobs FAQ

This is a new version of the gnomAD browser. The old version is available at <http://gnomad-old.broadinstitute.org>

gnomAD

genome aggregation database

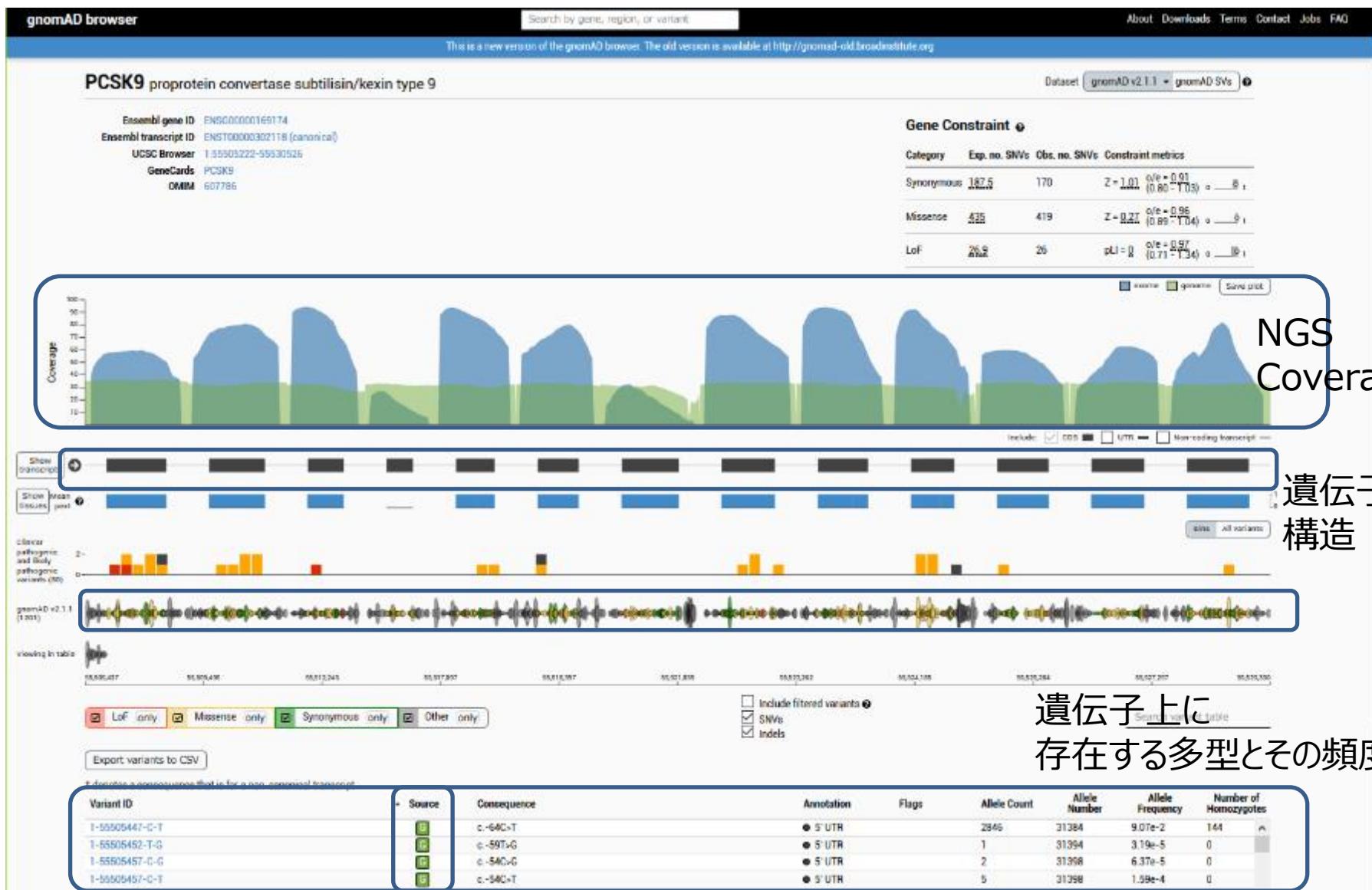
Search by gene, region, or variant

Examples - Gene: PCSK9, Variant: 1-55516888-G-GA

遺伝子名またはバリアントを指定
バリアントは、rs番号、HGVS、染色体番号-位置-参照アリル-代替アリルの様式で指定

The Genome Aggregation Database (gnomAD) is a resource developed by an international coalition of investigators, with the goal of aggregating and harmonizing both exome and genome sequencing data from a wide variety of large-scale sequencing projects, and making summary data available for the wider scientific community.

The data set provided on this website spans 125,748 exome sequences and 15,708 whole-genome



gnomADの検索結果(バリアント)

gnomAD browser gnomAD v2.1.1 Search About News Downloads Terms Publications Feedback Changelog Help

Single nucleotide variant: 1-11856378-G-A(GRCh37) Copy variant ID

Dataset gnomAD v2.1.1 ?

Filters

	Exomes	Genomes	Total
Allele Count	Pass	Pass	
Allele Number	79177	8057	87234
Allele Frequency	251468	31316	282784
Popmax Filtering AF (95% confidence)	0.3149	0.2573	0.3085
Number of homozygotes	0.4965	0.4313	
Mean depth of coverage	14637	1182	15819
Mean depth of coverage	97.5	32.4	

External Resources

- dbSNP (rs1801133)
- UCSC
- ClinVar (3520)
- ClinGen Allele Registry (CA170990)

Feedback

Report an issue with this variant

Population Frequencies

Population	Allele Count	Allele Number	N
Latino/Admixed American	17788	35440	
Ashkenazi Jewish	4740	10370	
European (non-Finnish)	43635	129108	
Other	2325	7220	
East Asian	5790	19944	
European (Finnish)	5799	25116	
South Asian	4443	30616	
African/African American	2714	24970	
XX	40861	129446	389 0.1451
XY	46373	153338	158 0.1087
Total	87234	282784	7576 0.3157
			8243 0.3024
			15819 0.3085

* Allele frequencies for some sub-continental populations are not included in this table.

Include: Exomes Genomes

Population 每、性別 每の
頻度情報

頻度の再計算が可能

- 全サンプル
- コントロールサンプル
- 非がんサンプル
- 非神経疾患サンプル
- TOPMedに含まれないサンプル

gnomADの検索結果 (バリアント)

Related Variants

Liftover

This variant lifts over to the following GRCh38 variant:

- 1-11796321-G-A
[View variant in gnomAD v3.1.2](#)

Nearby Variants

[View variants located within 20 bases of this variant.](#)

Variant Effect Predictor

This variant falls on 4 transcripts in 1 gene.

Note The gene symbols shown below are provided by VEP and may differ from the symbol shown on gene pages.

missense

1. MTHFR

- ENST00000376592.1
 Ensembl canonical transcript for MTHFR
 HGVS: p.Ala222Val
 Domains: [PF02219 \(Pfam\)](#), and 4 more
 Polyphen: ● probably_damaging
 SIFT: ● deleterious

- ENST00000376583.3

- HGVS: p.Ala263Val
 Domains: [PF02219 \(Pfam\)](#), and 4 more
 Polyphen: ● probably_damaging
 SIFT: ● deleterious

- ENST00000376585.1

- HGVS: p.Ala263Val
 Domains: [PF02219 \(Pfam\)](#), and 4 more
 Polyphen: ● probably_damaging
 SIFT: ● deleterious

- and 1 more

ClinVar

ClinVar Variation ID 3520

Conditions

Gastrointestinal stromal tumor, Gastric cancer, not provided, Neural tube defects, folate-sensitive, Homocystinuria due to methylene tetrahydrofolate reductase deficiency, Homocystinuria due to MTHFR deficiency, none provided, MTHFR deficiency, thermolabile type, not specified, Thrombophilia due to thrombin defect, cyclophosphamide response - Toxicity/ADR, carboplatin response - Efficacy, methotrexate response - Dosage, Efficacy, Toxicity/ADR

Clinical significance

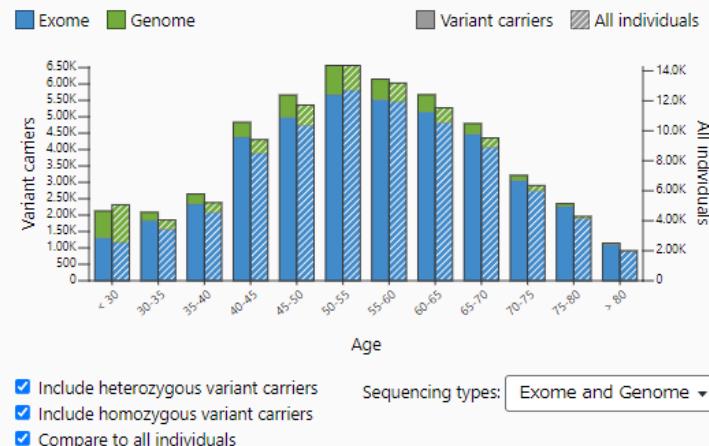
drug response

Variant Effect Predictor (VEP)の
 アノテーション情報が取得可能。
 Exonのバリアントについては、
 生物学的なインパクトを評価する
 SIFT、Polyphenのスコアを参照できる。

ClinVar(疾患とバリアントとの関連を収載したデータベース)
 からの情報も参照可能。

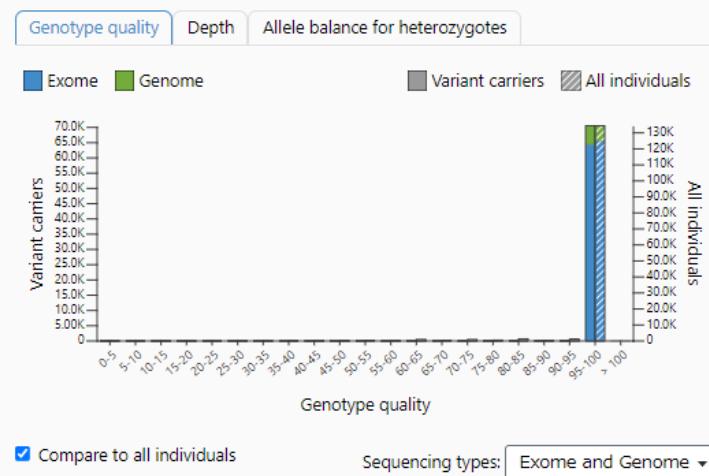
gnomADの検索結果（バリアント）

Age Distribution

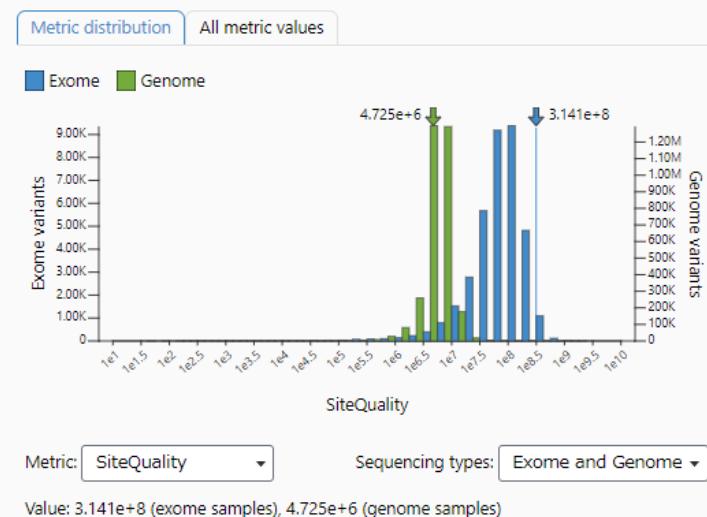


年齢の分布や
バリアントコールの際の各評価値を
参照可能。

Genotype Quality Metrics



Site Quality Metrics



gnomADの検索結果 (バリアント)

Read Data

This interactive IGV.js visualization shows reads that went into calling this variant. Reads may not be available for every sample carrying this variant.

These are reassembled reads produced by GATK HaplotypeCaller --bamOutput, so they accurately represent what HaplotypeCaller was seeing when it called this variant.

Note Reads shown here may include low quality genotypes that were excluded from allele counts.



あるサンプルにおけるNGSのリードのマッピング状況を参照可能。
Alternative alleleをヘテロまたはホモで持つサンプルを追加できる。

TO GO TV

Top About Videos Illustrations Training Contact Search for videos

病患に関連するバリアントや遺伝子発現の情報を調べる

◀ Previous 17. gnomAD v2.1.1 を使ってヒトバリアントの集団中のアレル頻度を調べる Next ▶

gnomAD v2.1.1 を使ってヒトバリアントの集団中のアレル頻度を調べる

TO GO TV

gnomAD v2.1.1 を使ってヒトバリアントの集団中のアレル頻度を調べる

20210818版

DBCLS

見る YouTube

この動画が再生されない場合は、YouTubeでご覧ください。 <https://youtu.be/SjXGgBs-fGo>

2021.08.23 11:27 Author: Ryo Niwa Editor: 井本 逸勢, 才津 浩智, 吉橋 博史, 小野 浩雅(ONO, Hiromasa)

gnomAD v2.1.1 を使ってヒトバリアントの集団中のアレル頻度を調べる

igsr-1000 genomes....tsv

Highlights

- 00:09 1. gnomADの紹介
- 00:27 2. gnomADのホームページを開く
- 00:58 3. gnomADデータセットのバージョンによる違い
- 02:10 4. gnomADで遺伝子を検索する
- 02:52 5. Constraintを確認する
- 03:54 6. シーケンスデータのカバレッジプロット
- 04:33 7. RNAアイソフォームを確認する
- 05:13 8. pext scoreを確認する
- 05:38 9. ClinVar variantsを確認する
- 06:42 10. gnomAD variantsを確認する
- 09:00 11. 個別のバリアントの詳細を確認する
- 09:37 12. Variant Effect Predictorについて
- 10:05 13. サンプルの詳細情報を確認する

Tags for this video

す

<https://togotv.dbcls.jp/en/20210823.html>

頻度データベースの紹介： jMorp、HGVD

Japanese Multi Omics Reference Panel (jMorp) とは

- 運営元： 東北メディカル・メガバンク機構
- [URL:https://jmorp.megabank.tohoku.ac.jp/202109/](https://jmorp.megabank.tohoku.ac.jp/202109/)
- 2021年12月8日にデータが更新。
 - 日本人（主に東北在住）の約1万4千名の住民コホートのWGSデータで検知されたバリアントの頻度データ
 - SVの頻度データ
 - 日本人参照ゲノム配列

jMorp 14KJPNのデータの内訳

プラットフォーム	サンプル数		
	男性	女性	計
Illumina HiSeq 2500	1,637	1,930	3,567
Illumina HiSeq X Five	33	36	69
Illumina NovaSeq 6000	1,829	3,271	5,100
MGI DNBSeq G400	643	719	1,362
MGI DNBSeq T7	1,893	2,138	4,031
計	6,035	8,094	14,129

カテゴリ	バリアント数	
	SNVs	INDELS
Autosome	106,705,823	13,130,321
chrX (PAR1+PAR2+XTR)	4,074,917	526,406
chrMT	3,832	
計	110,784,572	13,656,727

<https://jmorp.megabank.tohoku.ac.jp/202112/variants/statistics/> より改変

URL: <https://jmorp.megabank.tohoku.ac.jp/202109/>

Welcome to
Japanese Multi Omics
Reference Panel.

Datasets available on jMorp

	Phenome	Metagenome
	Metabolome	PGx
	Proteome	Metabolome
	Transcriptome	Proteome
	Methylome	ToMMo ISO-Seq
	Genome Variation	IMM Transcriptome
		IMM Methylome
		14KJPN (short-read based SNV/INDEL)
		JSV1 (long-read based SV)
		8.3KJPN-SV (short-read based SV)

jMorp release 202112

December 8th, 2021

Major Genome/Methylome update

[Genome Sequence]: We released Japanese reference genome JG2.1.0 as a successor of JG2.0.0 beta. In JG2.1.0, GRCh38-derived sequences are patched for undetermined regions. Resource bundles for WGS are available.

[Short-read based SNP analysis (Genome Variation)]: 14KJPN, allele and genotype frequency panels from about 14,000 Japanese individuals were released.

[Short/Long-read based structural variation analysis (Genome Variation)]: We released JSV1, a structural variation panel utilizing long-read sequencing technology of 333 individuals composing 111 trios. We also released 8.3KJPN-SV, a structural variation panel based on short read technology from about 8,300 individuals.

[Genetic Map]: Result of estimation of genome-wide recombination rate from linkage disequilibrium information of 300 haploid genomes is available from Downloads page.

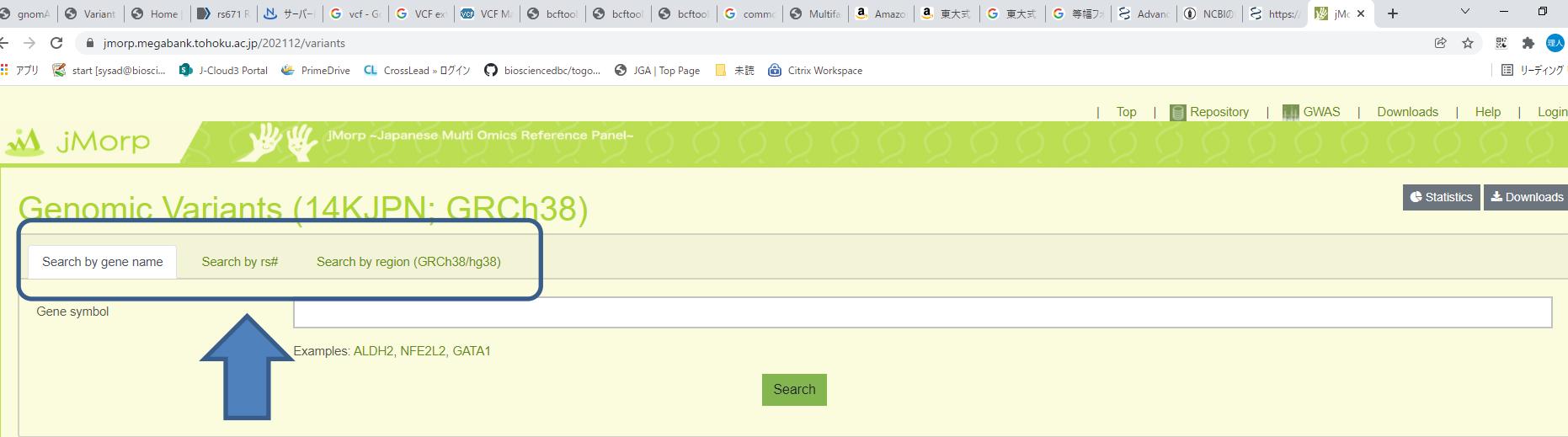
[Methylome]: New tracks for DNA methylation levels of cord blood and nucleated RBC were added to Genome Browser.

More

jMorp Publication

Tadaka et al., "jMorp updates in 2020: large enhancement of multi-omics data resources on the general Japanese population", Nucleic Acids Research 2020 Nov 12 okaa1034 doi: 10.1093/nar/oka1034 ↗

jMorpの検索方法



Genomic Variants (14KJPN: GRCh38)

Search by gene name Search by rs# Search by region (GRCh38/hg38)

Gene symbol Examples: ALDH2, NFE2L2, GATA1

Search

タブで検索対象を切り替え（遺伝子名、rs番号、領域で検索可能）

jMorp release 202112 / Last Update: December 8th, 2021 / Conditions of Use
Tohoku Medical Megabank Organization, Tohoku University
Contact: jmorp [at] omics.megabank.tohoku.ac.jp

This website uses cookies to ensure you get the best experience on our website. ([Privacy Policy](#)).
Use of this website and information stored in this database is subjected to [Conditions of Use](#).

Got it!

jMorpの検索結果

jMorp ~Japanese Multi Omics Reference Panel~

| Top | Repository | GWAS | Downloads | Help | Login |

Search by gene: ALDH2 (14KJPN) ToMMoにおける頻度

11258 variants found

Filter by keyword

Type	Position	Ref/Alt	rs#	Annotation	Gene	ClinVar Annotation	ToMMo 14KJPN	gnomAD AFR.	gnomAD AMR.	gnomAD EAS.	gnomAD NFE
SNV	chr12:111762007	A/T	rs1445408349	upstream_gene_variant	ALDH2		0.00025				
SNV	chr12:111762008	A/G	rs1280389202	upstream_gene_variant	ALDH2			0.00007			0.00001
SNV	chr12:111762009	G/A		upstream_gene_variant	ALDH2						0.00001
SNV	chr12:111762009	G/C		upstream_gene_variant	ALDH2						0.00003
SNV	chr12:111762011	T/C	rs974072569	upstream_gene_variant	ALDH2		0.00002		0.00007		
SNV	chr12:111762013	A/C		upstream_gene_variant	ALDH2			0.00002			
SNV	chr12:111762015	T/A	rs920558698	upstream_gene_variant	ALDH2		0.00002				
SNV	chr12:111762024	G/T	rs1022625657	upstream_gene_variant	ALDH2		0.00080	0.00020			0.00001
SNV	chr12:111762034	C/T	rs1410913722	upstream_gene_variant	ALDH2						0.00001
SNV	chr12:111762060	C/G		upstream_gene_variant	ALDH2			0.00007			
SNV	chr12:111762064	G/C	rs1476612156	upstream_gene_variant	ALDH2						0.00001
SNV	chr12:111762068	C/G	rs928863013	upstream_gene_variant	ALDH2		0.00002				0.00007
SNV	chr12:111762072	C/G		upstream_gene_variant	ALDH2						0.00003
SNV	chr12:111762072	C/T		upstream_gene_variant	ALDH2		0.00005				0.00001
SNV	chr12:111762073	G/A	rs1209311024	upstream_gene_variant	ALDH2						0.00001
SNV	chr12:111762080	T/C	rs539004647	upstream_gene_variant	ALDH2		0.00060				

gnomADにおける頻度

Statistics | Browser | Plot

« < < 1 2 3 4 ... > »

jMorpの検索結果

Table configuration

of rows: 50

Show/Hide columns:

- Type
- Position
- Ref/Alt
- rs#
- Annotation
- HGVS.c
- HGVS.p
- Gene
- ClinVar Annotation
- ToMMo 14KJPN
- ToMMo 14KJPN (Male)
- ToMMo 14KJPN (Female)
- ToMMo 14KJPN (Illumina HiSeq 2500 162PE)
- ToMMo 14KJPN (Illumina NovaSeq 6000 150PE)
- ToMMo 14KJPN (Illumina NovaSeq 6000 161PE)
- ToMMo 14KJPN (MGII DNBSeq T7 150PE)
- gnomAD AFR
- gnomAD AMR
- gnomAD EAS
- gnomAD NFE

歯車マークで設定を変更すれば
男女別
プラットフォーム別
の頻度も参照可能

The screenshot shows the jMorp search results page with a modal dialog titled "Table configuration". The dialog contains settings for the number of rows (set to 50) and a list of columns to show/hide. The "Show/Hide columns" section includes checkboxes for various genomic and population frequency annotations. A large text box on the right side of the dialog contains Japanese text explaining how to use the gear icon to change settings for gender, platform, and frequency.

Human Genome Variation Database (HGVD)とは

- 運営元： 京都大学
- URL: <http://www.hgvd.genome.med.kyoto-u.ac.jp/>
- 1208名のWESデータからの頻度情報 および
3,248名のSNP chipからの頻度情報
- 長浜コホートからのサンプルが中心

Human Genetic Variation Database

Home About Statistics Link Download Repository Contact How to Use Login



NGS Bioinformatics
Kyoto Course



Page views: 3,289,095
Total downloads: 6,259

Welcome to Human Genetic Variation Database

Search database

Gene name/ID GO

dbSNP rsID GO

Pathogenic Variation GO

Chromosome

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y

What's New?

- 08/02/2017 HGVD version 2.3 is now downloadable.
- 04/17/2017 New HLA typing software 'HLA-HD' is released. [link](#)

HGVDの詳細画面(バリエント)

Human Genetic Variation Database

[Home](#) [About](#) [Statistics](#) [Link](#) [Download](#) [Repository](#) [Contact](#) [How to Use](#) [Login](#)

dbSNP

SNP Position	rsID	View
12:112241766	rs671	View



Exome

View ボタンを押すと、ゲノムブラウザへ

Allele Color Code: A (Yellow), G (Blue), C (Green), T (Red), Others (Grey)

Variation Color Code: █:nonsense, █:splice, █:missense, █:synonymous, █:indel, █:non-coding, █:intron, █:other

Variant position	rsID	Ref/Alt	Frequency of alternative allele	Genotype count			Number of Samples Covered	Average sample read depth	Alleles	Gene	mRNA Accession#	Function				Org	Platform		
				Ref/Ref	Ref/Alt	Alt/Alt						Codon		AA					
				Ref	Alt	Ref						Ref	Alt	Ref	Alt				
chr12:112241766	rs671	G/A	0.2386	563	313	69	945	20.94±8.87	0.7614 0.2386	ALDH2	NM_000690	GAA	AAA	Glu	Lys	All			
											NM_001204889	GAA	AAA	Glu	Lys				
chr12:112241766	rs671	G/A	0.2667	162	116	22	300	20.63±6.81	0.7333 0.2667	ALDH2	NM_000690	GAA	AAA	Glu	Lys	KU	HiSeq + SOLiD		
											NM_001204889	GAA	AAA	Glu	Lys				
chr12:112241766	rs671	G/A	0.2282	253	140	26	420	10.37±6.01	0.6555 0.2282	ALDH2	NM_000690	GAA	AAA	Glu	Lys	YCHU	HIS		

6施設におけるバリエントの頻度情報 eQTL情報 (Top20のみ表示)

HGVDの詳細画面(ゲノムブラウザ)



Coverageやバリアントの頻度、e-QTLの情報をゲノムブラウザ上に配置可能
 左側からの青線枠内から欲しいTrackを右側へドラッグアンドドロップ。
 Trackを削除するときは、赤線枠内の×で削除

GWASデータベースの紹介： GWAS Catalog, JENGER

GWASの結果を集めた多型データベース GWAS Catalog

- 運営元： 米国国立ヒトゲノム研究（NHGRI）
および 欧州バイオインフォマティクス研究所
(EBI)
- URL: <https://www.ebi.ac.uk/gwas/>
- 今までGWASによって同定された形質（疾患、形質、薬剤副反応等）と関連が報告されたSNPをマニュアルキュレーション



GWAS Catalog

The NHGRI-EBI Catalog of published genome-wide association studies

Search the catalog

Examples: breast carcinoma, rs7329174, Yao, 2q37.1, HBS1L, 6:16000000-25000000

feedback

疾患名、rs番号、著者名、遺伝子名、領域で検索可能

Take a tour of the NEW search interface...

Search

Search the Catalog in a number of ways, including by trait, SNP identifier, publication, gene and genomic location.

Diagram

Explore an interactive visualisation of all SNP-trait associations with genome-wide significance ($p \leq 5 \times 10^{-8}$).

Download

Download a full copy of the GWAS Catalog in spreadsheet format as well as current and older versions of the GWAS diagram in SVG format.

GWASデータベースの紹介 GWAS Catalog

Show SNPs for

Digestive system disease	333
Cardiovascular disease	1042
Metabolic disease	403
Immune system disease	1879
Nervous system disease	1269
Liver enzyme measurement	159
Lipid or lipoprotein measurement	993
Inflammatory marker measurement	448
Hematological measurement	4955
Body weights and measures	1726
Cardiovascular measurement	955
Other measurement	11855
Response to drug	325
Biological process	2155
Cancer	1415
Other disease	1633
Other trait	1281



例) breast carcinomaと入力

The screenshot shows the GWAS Catalog search interface. At the top, there is a navigation bar with links to Home, Diagram, Download, Documentation, About, EMBL-EBI, and NIH. On the left, there is a sidebar titled "Refine search results" with buttons for "Traits" (10) and "Genes" (2). Below this is a section titled "Catalog stats" listing various statistics: Last data release on 2019-05-03, 3989 publications, 90031 SNPs, 138312 associations, Genome assembly GRCh38.p12, dbSNP Build 151, and Ensembl Build 96. The main content area has a large title "GWAS Catalog" and a subtitle "The NHGRI-EBI Catalog of published genome-wide association studies". A search bar contains the query "breast carcinoma" with a magnifying glass icon. Below the search bar is an example text: "Examples: breast carcinoma, rs7329174, Yao, 2q37.1, HBS1L, 6:16000000-25000000". The search results are displayed in a card-based format. The first result is "breast carcinoma" (EFO_0000305), described as a carcinoma arising from the breast, most commonly the terminal ductal-lobular unit. It is the most common malignant tumor in females. Risk factors include country of birth, family history, menstrua... Show more >. It has 1640 associations and 112 studies. The second result is "HER2 positive breast carcinoma" (EFO_1000294), described as a biologic subset of breast carcinoma defined by high expression of HER2, GRB7, and TRAP100, and by lack of expression of estrogen receptor (ER). It has 0 associations and 1 study.

GWAS Catalog

The NHGRI-EBI Catalog of published genome-wide association studies

breast carcinoma

Examples: breast carcinoma, rs7329174, Yao, 2q37.1, HBS1L, 6:16000000-25000000

GWAS / Search / breast carcinoma

Refine search results

T Traits 10

G Genes 2

Catalog stats

- Last data release on 2019-05-03
- 3989 publications
- 90031 SNPs
- 138312 associations
- Genome assembly GRCh38.p12
- dbSNP Build 151
- Ensembl Build 96

Search results for *breast carcinoma*

breast carcinoma EFO_0000305

A carcinoma arising from the breast, most commonly the terminal ductal-lobular unit. It is the most common malignant tumor in females. Risk factors include country of birth, family history, menstrua... Show more >

Associations 1640 Studies 112

HER2 positive breast carcinoma EFO_1000294

A biologic subset of breast carcinoma defined by high expression of HER2, GRB7, and TRAP100, and by lack of expression of estrogen receptor (ER).

Associations 0 Studies 1

GWAS Catalog 詳細画面(形質で検索)

Associations 1409

ある形質で行われた関連解析の結果

Variant and risk allele	P-value	P-value annotation	RAF	OR	Beta	CI	Mapped gene	Reported trait	Trait(s)	Study accession	Location
rs2111836-G	7×10^{-6}		0.22	1.32	-	[1.17-1.44]	MTOL1	Breast cancer	breast carcinoma	GCST002735	18:8716323
rs13025833-A	1×10^{-7}		0.52	1.2	-	[1.13-1.33]	CTNNA2	Breast cancer	breast carcinoma	GCST002735	2:79716982
rs1078806-C	2×10^{-6}	(Pooled P value)	0.39	1.43	-	[NR]	FGFR2	Breast cancer	breast carcinoma	GCST001831	10:121579461
rs166870-T	3×10^{-7}	(HR+ HER2-)	0.13	2.3	-	[1.67-3.15]	AC026826.2, RNU6-667P	Disease-free survival in breast cancer	breast carcinoma, disease free survival	GCST002847	15:79777420
rs10825036-G	4×10^{-7}	(HR- HER2-)	0.32	2.26	-	[1.34-3.81]	RNA5SP318, AL365496.1	Disease-free survival in breast cancer	disease free survival, breast carcinoma	GCST002847	10:53506471

Showing 1 to 5 of 1409 rows 5 rows per page 1 2 3 4 5 ... 282 >

関連解析の研究概要（サンプルの集団も記載あり）

GWASの結果を集めたデータベースの紹介： JENGER

GWASの結果を集めたデータベース JENGER

- 運営元： 理化学研究所
- URL: <http://jenger.riken.jp:8080/>
- 理研が実施したGWASおよびPheWASの統計値を公開
- 13 category, 244 phenotypeの結果が公開

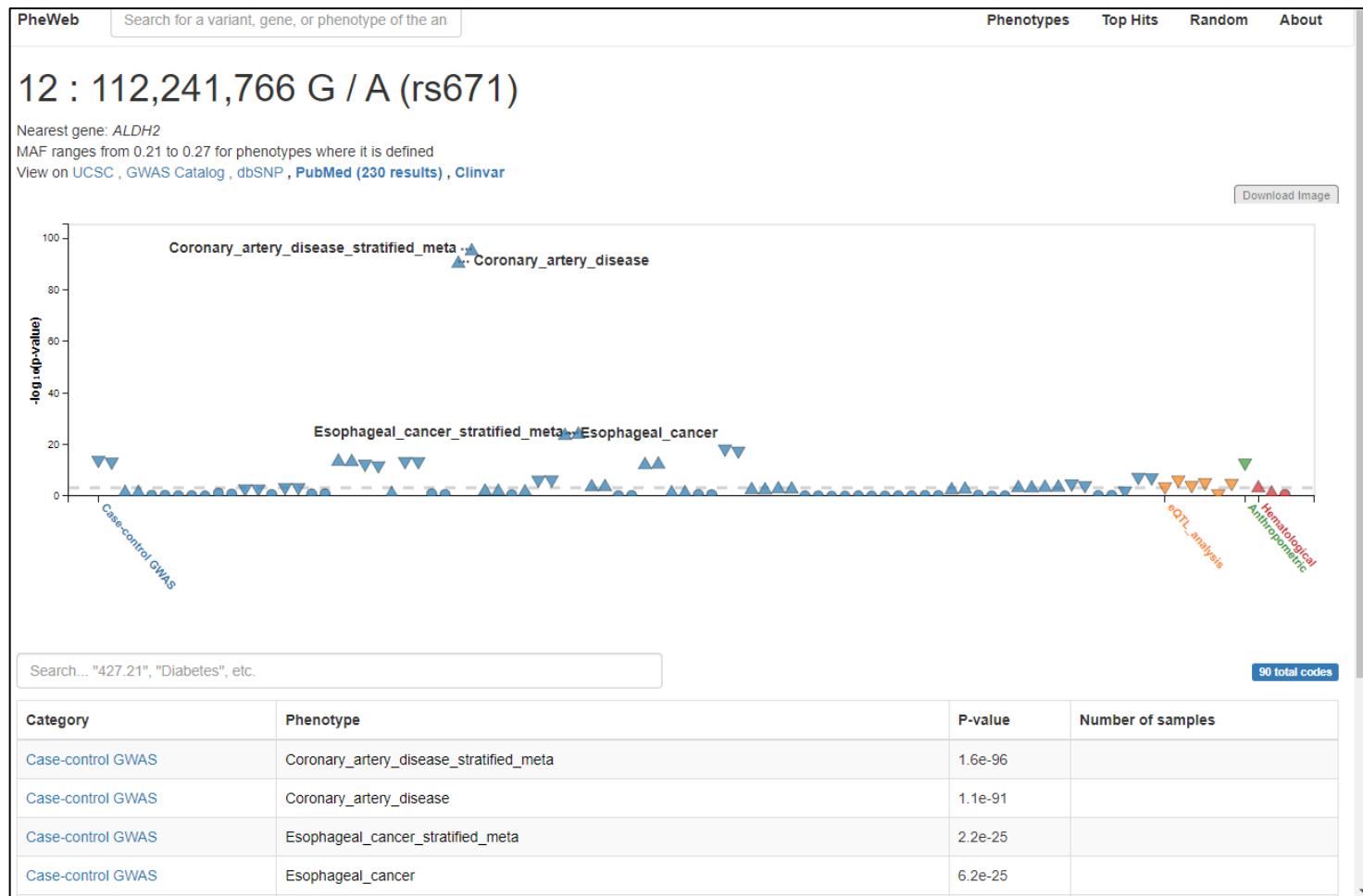
Phenotypes Top Hits Random About

PheWeb for JENGER

Search for a variant, gene, or phenotype of the analysis published in JENGER

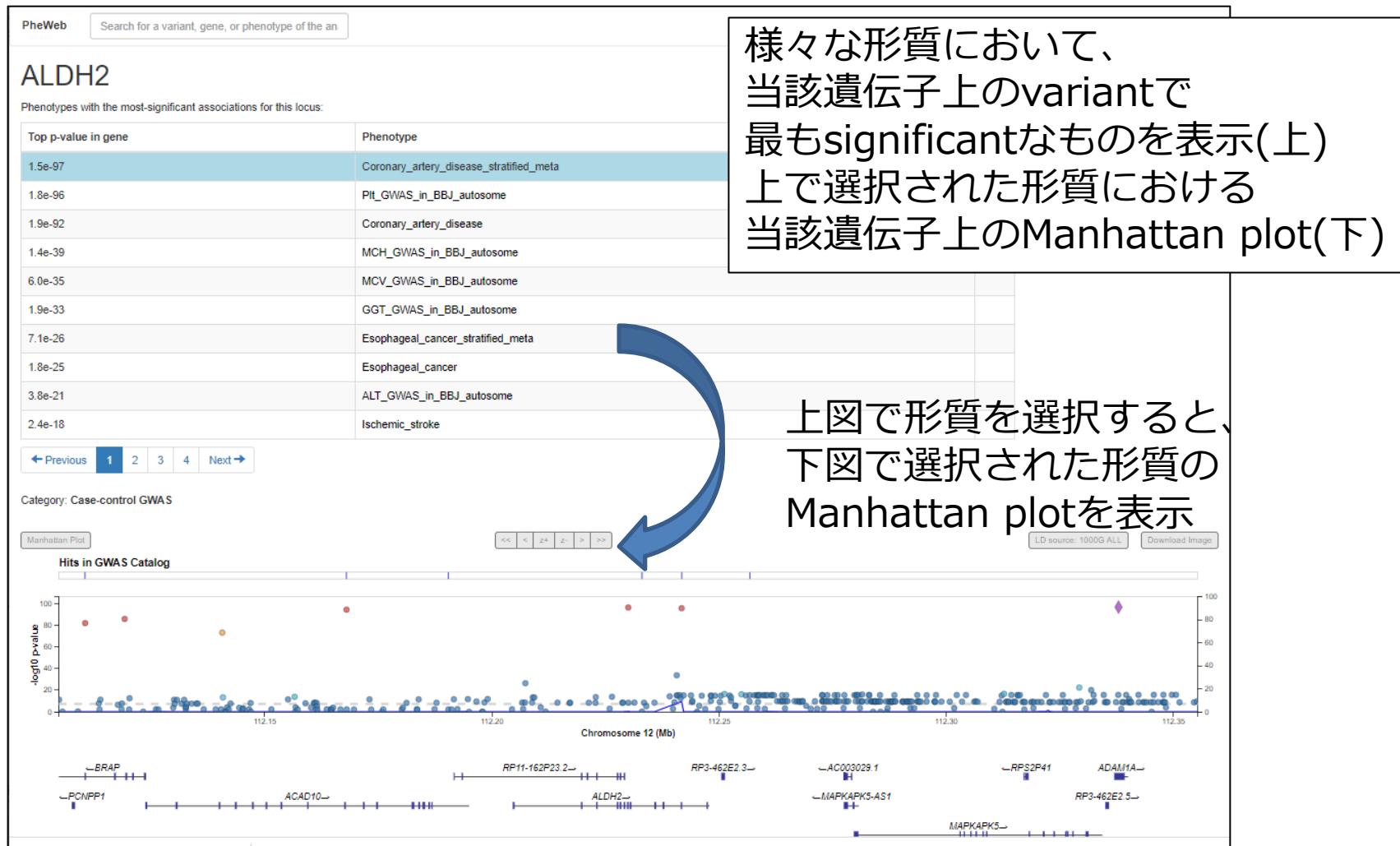
疾患名、rs番号、遺伝子名、染色体番号:位置-Refアリル-Altアリルで検索可能

JENGER詳細画面(バリアントで検索時)



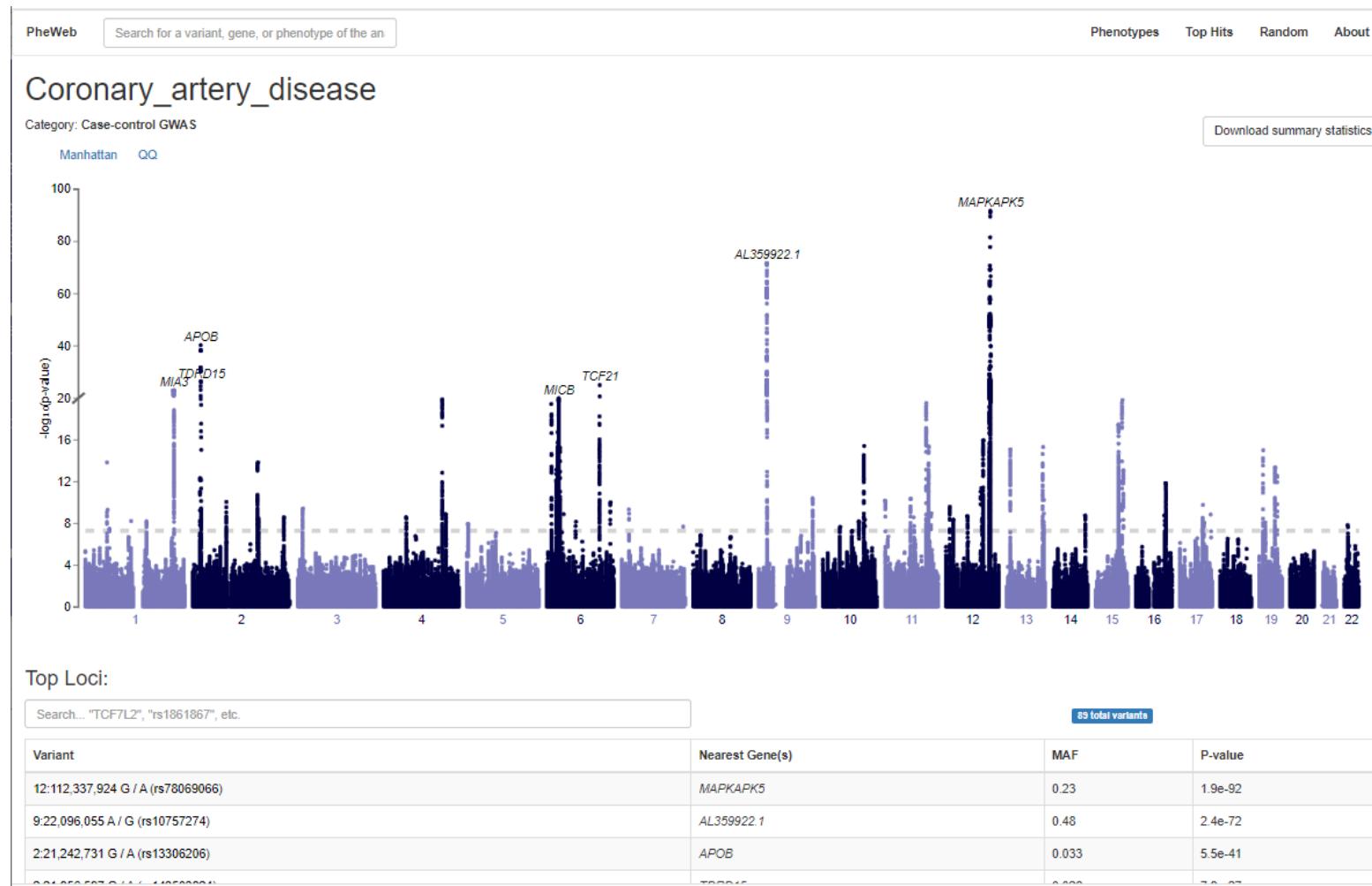
縦軸が $-\log_{10}(p\text{-value})$ で、横軸が phenotype あるバリアントについて複数のphenotypeとの関連を確認できる

JENGER詳細画面（遺伝子名で検索時）



様々な形質において、
当該遺伝子上のvariantで
最もsignificantなものを表示(上)
上で選択された形質における
当該遺伝子上のManhattan plot(下)

上図で形質を選択すると、
下図で選択された形質の
Manhattan plotを表示



検索した形質におけるManhattan Plot, トップヒットのSNV, QQ-plotが表示

- 運営元 : NBDC, JST
- URL:
<https://togovar.biosciencedbc.jp/>
- 「NBDCヒトデータベース」を基に個人特定されない加工データ（頻度情報）を提供、データの概要を把握可能に
- 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
- 現状germline variantのみ、somaticなし

NBDCヒトデータベースとは？

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活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 (DDJ)**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

- 2019/05/28 球化研究所 生命医科学研究センターからの制限公開データ (Type I) を公開しました (hum0160)
- 2019/05/09 大阪大学免疫学フロンティア研究センター 実験免疫学からの制限公開データ (Type I) を公開しました (hum0141)

[ニュース一覧へ](#)

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

Human Database Beacon is a member of [GA4GH Beacon Network](#).

RCh37 e.g. 12:112241766 A Search Example: ALDH2 Variant (GRCh37, '12:112241766 A')

可能な研究データ一覧

再利用方法は[こちら](#)をご覧下さい。

14 件 [Copy](#) [CSV](#) [Excel](#)

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
m0173.v1 AS000000000171	胎児期に始まる子どもの健康と発達に関する調査	v1:2019/04/23	NGS (RRBS)	メチル化 解析	Illumina (NextSeq)	臍帯組織(正常出 産) : 30検体 (日本人)	森 千里	制限 (Type I)
m0165.v1	家族性骨髓異形成症候群	v1:2019/05/07	NGS	配列決定	Illumina	家族性骨髓異形成症候 群	古屋 淳史	制限

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

研究者の産出したデータを、他の研究者が再利用できるサービス

NBDCヒトデータベースの必要性 研究データ公開・再利用の世界的潮流から

- オープンサイエンス
学術論文のオープン化 + 研究データの公開

2014年米国NIHのGenomics Data Sharing Policy
ヒトゲノム研究データのデータリポジトリへの登録を要求
2018年3月 AMEDデータマネージメントプラン提出の義務化
- FAIR原則
Findable(見つけられる)、Accessible(アクセスできる)、
Interoperable(相互運用できる)、Reusable (再利用できる)
- Global Alliance for Genomics and Health(GA4GH)
国際的な研究情報共有や統一プロトコルでのゲノム解析の仕組みや、
研究者認証等についての枠組みを討議するアライアンス

NBDCヒトデータベースの必要性

制限公開データベースの必要性（プライバシーの担保）

- 研究データのオープン化の懸念
 - プライバシー保全とのバランス
 - データ公開・利用についての適切な審査
 - →制限公開データベース
- 制限公開データベースの発足
 - 2007年米国NIH
 - Database of Genotypes and Phenotypes (dbGaP)
 - 2008年欧州EMBL-EBI
 - European Genome-phenome Archive (EGA)
 - 2013年10月 NBDCおよび国立遺伝学研究所DDBJセンター
 - NBDCヒトデータベース/Japanese Genotype-phenotype Archive(JGA)

非制限公開データ Open data

- ・頻度や統計量
 - ・特定の個人由来では無い試料の解析結果
- Webからダウンロード可能

制限公開データ Controlled-Access data

- ・個人ごとのゲノムデータ
- 利用者を限定
(国内外のアカデミアや企業内研究者)

二重匿名化

各プロジェクト・実施機関

NBDCヒトデータベースのデータ例 バイオバンクジャパン (BBJ) : hum0014

心筋梗塞：1666症例、対照：3198名

健常者：934名 (JSNP)

35疾患：各約190症例 (JSNP)

食道癌：182症例 (JSNP)

ALS：92症例 (JSNP)

T2DM：9817症例、対照：6763名

T2DM：5646症例、対照：19,420名

AD：1472症例、対照：7966名

AF：8180症例、対照：28,612名

BMI：158,284名

Genotypeデータ：182,505名

POAG：3980症例、対照：18,815名

58臨床検査値：162,255名

初潮年齢データを有する女

性：67,029名

閉経年齢データを有する女

性：43,861名

BBJ第1コホート：1,026名

(日本人)

登録された対象疾患および形質：

心筋梗塞、2型糖尿病、アトピー性皮膚炎、
心房細動、BMI、開放隅角緑内障、
初潮・閉経年齢等々

プラットフォーム：

インベーダー法 (Hologic Japan社)
Human610-Quad BeadChip(Illumina社) 等の
Genotyping Arrayデータ

昨年8月

Illumina社 HiSeq2500による
BBJ第一コホート 1,026名の
Whole Genome Sequencingを公開

NBDC Research ID: hum0043.v1

研究内容の概要

目的 : アルツハイマー病（AD）治療薬の薬効評価基準の最適化を行うために、ADの病態を忠実に反映するサロゲートマーカー (surrogate marker)を定めておく必要がある。J-ADNI研究は、MRIやPETなどの画像サロゲートマーカーの長期的変化に関する一定の基準値を作成するとともにその妥当性を証明するために臨床/神経心理検査データ、血液・脳脊髄液サンプルを並行して収集することを目的とする。

対象 :

537名

(高齢者健常群 149名、軽度認知機能障害 234名、
早期アルツハイマー病 154名)

- 臨床・心理学検査結果

- 検査結果

ApoE遺伝子型、末梢血液一般検査

- MRI画像、PET画像

NBDCヒトデータベースの必要性

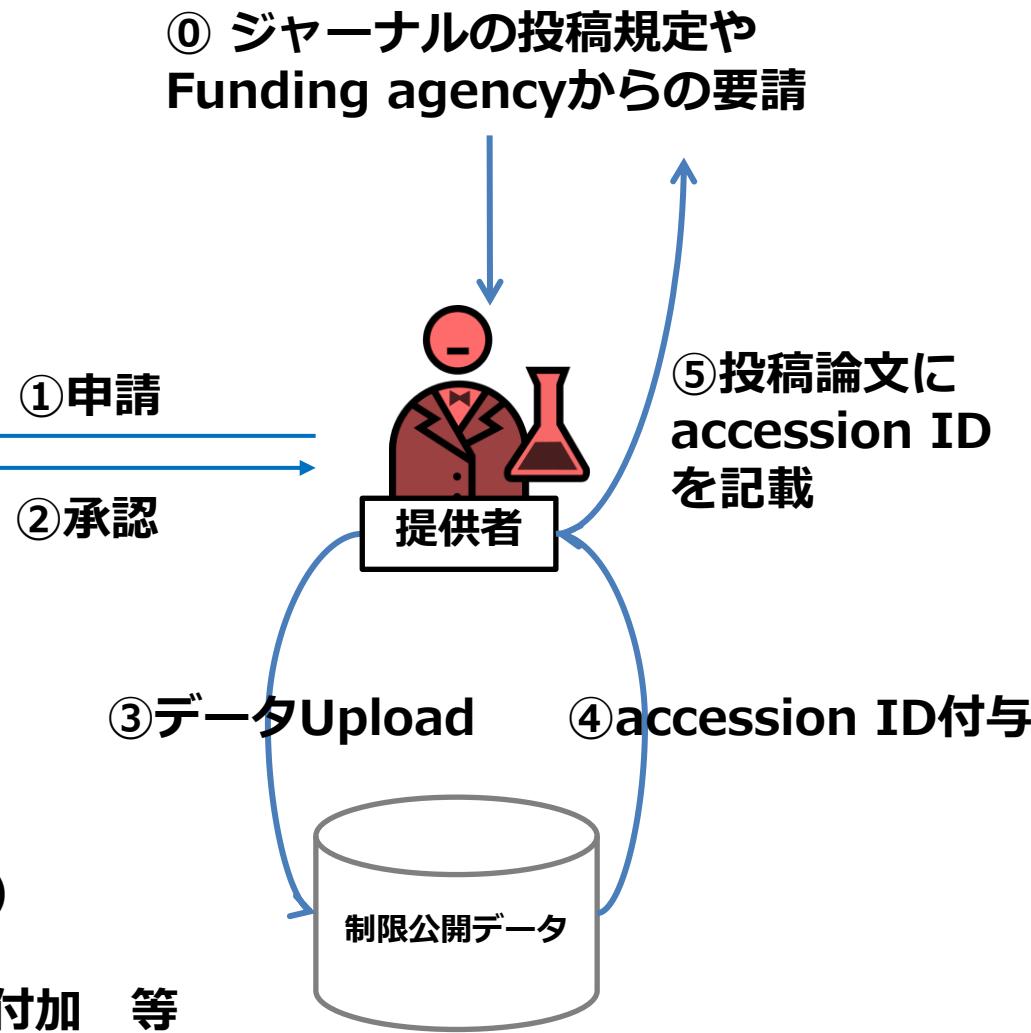
制限公開データベースの必要性（プライバシーの担保）

- 研究データのオープン化の懸念
 - プライバシー保全とのバランス
 - データ公開・利用についての適切な審査
 - →制限公開データベース
- 制限公開データベースの発足
 - 2007年米国NIH
 - Database of Genotypes and Phenotypes (dbGaP)
 - 2008年欧州EMBL-EBI
 - European Genome-phenome Archive (EGA)
 - 2013年10月 NBDCおよび国立遺伝学研究所DDBJセンター
 - NBDCヒトデータベース/Japanese Genotype-phenotype Archive(JGA)

NBDC

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

ヒトデータ審査委員会



Human data NBDCヒトデータベース English サイト内検索 検索

ホーム データの利用 データの提供 ガイドライン NBDCヒトデータ審査委員会 成果発表 お問い合わせ FAQ

データの提供 <https://humandbs.biosciencedbc.jp/data-submission>

i 2017年5月30日より改正個人情報保護法および改正倫理指針が施行されます。データ提供を行う際には、新指針に沿った手続きを実施する必要がありますのでご留意ください。詳細はFAQをご参照ください。

NBDCヒトデータベースへのヒトに関するデータ（遺伝情報、臨床情報、画像情報など）の提供手続きは以下の通りです。

1. [NBDCヒトデータ共有ガイドライン](#)を熟読して下さい。
2. 以下の情報を準備して下さい。
 - 提供するデータについての情報：研究内容の概要（目的、方法、対象、発表論文など）、データの種類（名称・量、アクセス制限レベルの分類に関する情報、公開可能日など）
 - 研究代表者の氏名、所属情報、連絡先
 - 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。
 - 研究計画書（倫理審査申請書）（PDF形式 5 MB以内）
 - 同意文書及び説明文書のフォーム（PDF形式 5 MB以内） ※実際に署名されたものではなくフォーム（書式）をご提出ください。
 - 研究内容についての倫理審査の承認通知書（PDF形式 5 MB以内）

研究計画書、インフォームドコンセント(IC)の説明・同意文書にて
『データの提供・利用(共有)』について記載しておくことが重要。

- ・データベースの具体名は必須ではないが、**公的データベースの公開を記載。**
- ・国内研究機関だけではなく、製薬企業等の**民間企業や海外の機関に所属する研究者**ともデータを共有する可能性があることについて記載。
- ・データベースから公開されるデータは、将来、**他の研究にも利用**される可能性があることを記載。
- ・データベースではデータ種類に応じたアクセスレベル（**制限公開、非制限公開**）があり、非制限公開データは不特定多数の者に利用され、制限公開データ（個人毎のゲノムデータ）は、データ利用の審査を経て承認された研究者に利用されることを記載。

データ提供に必要な手続きは

ホーム | データの利用 | データの提供 | ガイドライン | NBDCヒトデータ審査委員会 | 成果発表 | お問い合わせ | FAQ

<https://humandbs.biosciencedbc.jp/faq>

よくあるご質問 (FAQ)

19. インフォームドコンセントの説明文書にどのように記載すればよいのか?

2017年5月30日に施行された改正個人情報保護法や改正研究倫理指針の変更点を反映したモデル書式を先進ゲノム支援との協力の下、作成しました。参考にして下さい。



先進ゲノム支援とは 公募要項/申請受付 支援技術紹介 倫理関連情報 FAQ 支援課題一覧 成果論文一覧

倫理関連情報

ヒト由来試料のゲノム研究のためのモデル書式等の改訂について (2018年4月27日)

「先進ゲノム支援」ではゲノム科学全体の発展のためにデータの公開・共有を積極的に進めることを方針としております。このために、支援で得られたヒト由来試料のゲノムデータは、我が国における米国のdbGAPや欧州のE加カウンターパートである科学技術振興機構（JST）バイオサイエンスデータベースセンター（NBDC）内のNBDCデータベースJGA（Japanese Genotype-phenotype Archive, <http://humandbs.biosciencedbc.jp/>）に制限なしでデータとして登録することいたします。登録時点では、支援依頼者の承諾なしに公開・共有されることはありません。その後、論文発表等により支援依頼者からの承諾が得られましたら、制限公開に移り、JGAのData Access CommitteeであるNBDCヒトデータ審査委員会による審査を経て承認された研究者を対象に登録データの共有を行うことになります（データの制限公開日は支援依頼者が指定することができます。また、公開前に必ず支援依頼者に確認を行います）。

JGAに登録するためには、個人情報保護法および研究倫理指針等に沿った適切な倫理面の対応が必要です。そのため、「先進ゲノム支援」ではヒト由来試料を用いるゲノム解析研究のための「説明文書および同意文書のモデル書式」等を提示するとともに、各研究機関における倫理審査申請時の支援を実施して参りました。昨年7月に改訂モデル書式などを掲載しましたが、その後の検討を経てマイナーな改訂を行いましたので、以下を利用ください。

- ▶ IC文書等改訂のお知らせ (2017年7月)
- ▶ 先進ゲノム支援同意説明文書のひな形【2018年度版】 (2017年度版からの変更履歴)
- ▶ 先進ゲノム支援_記載すべき内容のリスト【2018年度版】 (2017年度版からの変更履歴)

<https://www.genome-sci.jp/ethic>

データ公開に適したIC
のサンプルは？

同意説明文書の雛形をダウンロード可
不明点は事前にご相談ください



NBDC

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

ヒトデータ審査委員会

①申請

②承認

レプリケーション
データセット
が欲しい④利用状況報告
(年1回および終了時)

利用者

③データへアクセス

制限公開データ

1. データ利用に必要な手続きは

Human data NBDCヒトデータベース

English サイト内検索 検索

ホーム データの利用 データの提供 ガイドライン NBDCヒトデータ審査委員会 成果発表 お問い合わせ FAQ

https://humandbs.biosciencedbc.jp/data-use

まつき。

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下さい。

- 研究代表者およびデータ利用を希望する研究分担者全員の氏名、所属機関情報 ※同一機間に所属する研究分担者を一回の申請で登録することができます。
- 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）（PDF形式 5 MB以内）
- 倫理審査の承認通知書（PDF形式 5 MB以内）

4. NBDCヒトデータ取り扱いセキュリティガイドライン（利用者向け）の内容を確認し、書式5) NBDCヒトデータ取り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ます。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行ない、NBDCヒトデータ取り扱いセキュリティガイドラインチェックリストを記載して下さい（XLS/XLSX形式 5 MB以内）。

5. 研究分担者がいる場合は、書式7) 研究分担者一覧をダウンロードし、必要事項を記載して下さい。

6. 書式2) NBDCヒトデータベースデータ利用申請書（制限公開データ用）に必要事項を入力し、研究計画書（倫理審査申請書）、承認通知書、NBDCヒトデータ取り扱いセキュリティガイドラインチェックリスト、研究分担者一覧（該当者のみ）とともにNBDCヒトデータ審査委員会事務局(humandbs@biosciencedbc.jp)までメールで送付し、利用申請を行なって下さい。なお、一つの申請で複数のデータIDについて申請することができます。実際の申請作業は、代表者もしくは代表者が指名する者が行なって下さい。

研究計画書に、NBDCヒトデータベースのデータ利用について記述し、倫理審査委員会の承認および機関の長の研究実施の許可を得る。

- 自研究機関

自機関で大規模なサーバが用意できない場合は、
機関外サーバで区画を借りて解析可能

- 機関外サーバ
 - 国立遺伝学研究所のスパコンの個人ゲノム解析区画
 - 東北メディカル・メガバンク機構のスパコン
 - 一般社団法人柏の葉オーミクスゲート (KOG)
 - 商用利用可
- 国立遺伝学研究所と解析パイプラインを開発中
 - NGSのvariant call
 - GWASの imputation解析

日本人多様性データベース: TogoVar

- 運営元 : NBDC, JST
- URL:
<https://togovar.biosciencedbc.jp/>
- NBDCヒトデータベースを基に個人特定されない加工データ（頻度情報）を提供、データの概要を把握可能に
- 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
- 現状germline variantのみ、somaticなし

『NBDCヒトDB』の概要を把握

 Human data NBDCヒトデータベース

 DDBJ

Japanese Genotype-phenotype Archive

 TOGO VAR

個人別の情報（個人情報）

個人別の情報でない（集計情報）

制限公開（提供・利用審査あり）

非制限公開

データ
提供者



研究プロジェクトA



同じ手法で
再解析

研究プロジェクトB



研究プロジェクトC



NGSデータ由來
頻度情報
(JGA-NGS)

SNP-Chipデータ由來
頻度情報
(JGA-SNP)

日本人大規模
バリアント頻度

① 概要を
把握

データ
利用者



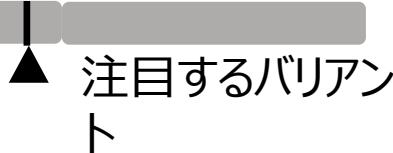
出典元 (© 2016 DBCLS TogTV)

② 利用申請

ワンストップ検索

ゲノム配列の個人による違い（バリアント）に関する
さまざまな条件を用いて、国内外のDBや文献情報などのワンストップ検索を可能に

7番染色体



ClinVar (NCBI)

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

ExAC (ブロード研究所)

位置 : chr7:127254587
アレル頻度 : 0.000304573



jMorp 8.3KJPN

(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

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(jMorp) : 0.0233

アレル頻度(HGVD) : 0.0272809

アレル頻度(ExAC) : 0.000304573

関連論文 :

A missense mutation of Pax4 gene ...

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



ワンストップ検索



出典元 (© 2016 DBCLS TogoTV)

TogoVar／検索の対象データベース		
データベース名 (運営組織)	説明	対象人数
NBDCヒトデータベース (JST-NBDCと国立遺伝学研究所DDBJセンターの共同運営) • JGA-NGSデータセット (全エクソーム) • JGA-SNPデータセット (SNP Chip)	主に日本の研究者からの個人ゲノムデータのリポジトリ	125人 (全エクソーム) 183,884人 (既知SNP)
Japanese Multi Omics Reference Panel(jMorp) (東北メディカル・メガバンク機構)	ゲノムコホート (東北地方中心)	8,380人 (全ゲノム)
Human Genetic Variation Database (HGVD) (京都大学)	ゲノムコホート (滋賀県長浜市を中心)	1,208人 (全エクソーム)
gnomAD (ブロード研究所)	配列決定プロジェクトの再解析データ (約20プロジェクト)	125,748人 (全エクソーム) 15,708人 (全ゲノム)
ClinVar (NCBI)	バリアントの疾患関連性	
PubTator (NCBI) Colil (DBCLS)	バリアント (rs番号) が出現する文献情報	

JGA-NGS/JGA-SNPデータの由来

JGA-NGS

https://togovar.biosciencedbc.jp/doc/datasets/jga_ngs

集約されたデータの由来

JGAID	ヒトDB	研究題目	対象集団	サンプルサイズ	データ提供者
JGAD000000000004	hum0006	脳腫瘍のゲノム・遺伝子解析とその臨床病理学的意義の解明	脳腫瘍 アストロサイトーマ	6	齊藤 延人 (P-DIRECT)
JGAD000000000106	hum0006	脳腫瘍のゲノム・遺伝子解析とその臨床病理学的意義の解明	脳腫瘍 オリゴデンドログリオーマ	16	齊藤 延人 (P-DIRECT)
JGAD000000000112	hum0006	脳腫瘍のゲノム・遺伝子解析とその臨床病理学的意義の解明	脳腫瘍 小脳グリオーマ	17	齊藤 延人 (P-DIRECT)
JGAD000000000014	hum0021	精神神経疾患の原因解明および診断法・治療法の開発に関する研究	健常一卵性双生児	6	加藤 忠史 (ゲノム支援)
JGAD000000000036	hum0035	固形腫瘍における遺伝子異常の網羅的解析	多数の固形腫瘍性疾患検体	23	滝田 順子 (P-DIRECT)
JGAD000000000038	hum0040	ヒト胎盤におけるインプリント制御領域およびインプリント遺伝子の同定	ヒト胎盤および母体血	48	有馬 隆博 (IHEC)
JGAD000000000060	hum0066	癌の再発・転移に関するnon-coding RNAの同定とその機序の解明	大腸がん	9	三森 功士
合計					125

JGA-SNP

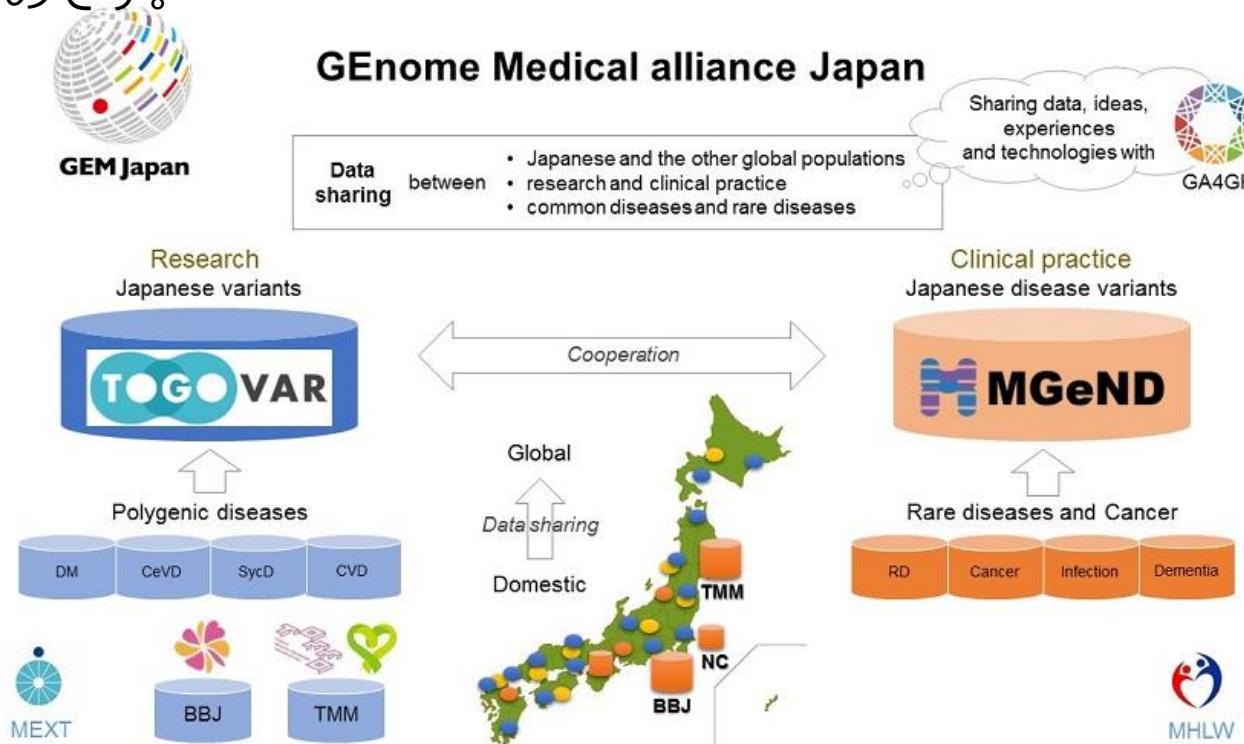
https://togovar.biosciencedbc.jp/doc/datasets/jga_snp

集約されたデータの由来

JGAID	ヒトDB	研究題目	対象集団	サンプルサイズ	データ提供者
JGAD00000000123	hum0014	オーダーメイド医療の実現プログラム	健常者及び罹患者	182,557	久保 充明 (BBJ)
JGAD00000000018	hum0028	オーダーメイド医療の実現プログラム	健常者	908	久保 充明 (BBJ)
JGAD00000000130/131	hum0082	日本人健常者におけるゲノム全塗のSNP解析	健常者	419	徳永 勝士
Total				183,884	

Genome Medical alliance Japan (GEM Japan)

「GEM Japan」は、データシェアリングを進めながらゲノム医療の実現を目指すAMEDの各事業に関わる大学、研究所、病院等と日本全国規模で協力体制を築き、臨床情報と個人ゲノム情報のデータシェアリングと研究利用を促進し、ゲノム医療の実現を目指すものです。



(https://www.amed.go.jp/aboutus/collaboration/ga4gh_gem_japan.html より)

日本人全ゲノム解析に基づく日本人アレル頻度情報の公開

	SNV（一塩基多様性）		INDEL（挿入欠失配列）	
	総数	新規検知数 (内数)	総数	新規検知数 (内数)
常染色体	76,768,387	35,660,425	10,202,908	4,152,671
X 染色体	2,898,518	1,420,888	410,435	164,077

※ 新規検知数: dbSNP152に含まれないバリアントの数

PCAでPlatform毎やサイト毎のクラスタリングが観察
 → VCFのFILTER値にNotHighConfidenceRegionのフラグを付与

検索ボックス

- rs番号
- 位置検索、範囲検索 (hg19)
- 遺伝子名 (あいまい検索)
- 関連疾患名 (あいまい検索)

検索結果

TogoVar ID	RefSNP ID	Position	Consequence	SIFT	PolyPhen	Clinical significance
tgv67071948	rs7792	1: 10231	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071949		1: 10248	Deletion	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071950	rs13663	1: 10248	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071951	rs14094	1: 10248	AAAC... Del	Deletion	Intergenic variant	Intergenic variant
tgv67071952	rs20173260	1: 10248	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071953	rs775928745	1: 10231	Deletion	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071954	rs1267031179	1: 10248	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071955		1: 10248	AAAC... Del	Deletion	Intergenic variant	Intergenic variant
tgv13	rs145427775	1: 10291	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv14	rs060413313	1: 10297	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv15	rs5756	1: 10321	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv16		1: 10327	Deletion	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071957	rs112750067	1: 10327	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071958		1: 10328	Deletion	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071960	rs1035171912	1: 10330	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071959	rs150969722	1: 10330	Deletion	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071962	rs1357792689	1: 10334	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071961	rs1351390918	1: 10334	Deletion	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071963	rs1193008993	1: 10336	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071964	rs1471210572	1: 10340	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071965		1: 10343	CCCT... Del	Deletion	Intergenic variant	Intergenic variant
tgv67071966		1: 10346	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071967	rs1363828207	1: 10348	ACCC... Del	Deletion	Intergenic variant	Intergenic variant
tgv67071968		1: 10349	CCCT... Del	Deletion	Intergenic variant	Intergenic variant
tgv67071969		1: 10352	ACCC... Del	Insertion	Intergenic variant	Intergenic variant
tgv67071970	rs1015856060	1: 10354	SNV	Intergenic variant	Intergenic variant	Intergenic variant
tgv67071971		1: 10384	CCCT... Del	Deletion	Intergenic variant	Intergenic variant
tgv67071972	rs73617372	1: 10396	Deletion	Intergenic variant	Intergenic variant	Intergenic variant

フィルタ機能

- データセット
- データセット+頻度
- バリアントタイプ
- ClinVarの情報

Filters

Dataset

- All 74,698,940
- JGA NGS 4,679,025
- JGA SNP 1,249,724
- 3.5KJPN 64,675,495
- HGVD 354,461
- ExAC 10,195,872
- ClinVar 443,512

Alternative allele frequency

Variant calling quality

Variant type

- All 74,698,940
- SNV 63,066,161
- Insertion 5,383,629
- Deletion 6,240,052
- Indel 7,358
- Substitution 1,740

Clinical significance

- All 74,698,940
- Not in ClinVar 74,255,428
- Pathogenic 112,579
- Likely pathogenic 39,343
- Uncertain significance 353,079
- Likely benign 175,644
- Benign 94,328
- Conflicting interpretations of pathogenicity 6,047

Variant report
tgv47264307
dbSNP
rs671

Variant type SNV

Position 12:112241766 (GRCh37)

Ref / Alt G > A

hgvs 12:g.112241766G>A

Other overlapping variant(s)

No other variants on the same location

Frequency

Dataset	Population	Allele count Alt	Total	Frequency		Genotype count Alt / Alt	Alt / Ref	Ref / Ref	Filter status	Quality score
JGA-NGS	Japanese	52	250	0.208	██████				PASS	451.77
JGA-SNP	Japanese	90,026	365,930	0.246	██████	11,778	66,470	104,717	-	
3.5k JPN	Japanese	1,392	7,104	0.196	██████				PASS	
HGVD	Japanese	451	1,890	0.239	██████				PASS	
ExAC	Total	1,878	88,224	0.021	████				PASS	957348.0

Clinical Significance

Title	Clinical significance	Review status	Last evaluated	Condition(s)
NM_000690.3(ALDH2):c.1510G>A (p.Glu504Lys)	drug response protective	★★★★ no assertion criteria provided	2018-08-09 2010-02-01	• Acute alcohol sensitivity • Alcohol dependence

バリアントの基本的な情報や

ClinVarに登録されたpathogenicityの情報、各データセットにおけるアリル頻度を表示

また、PubMedに登録された論文のabstractに当該バリアントの記述がある論文と関連疾患名を表示。

Transcripts

Transcript ID	Gene symbol	Consequence type	HGVS(cDNA)	HGVS(Amino acid seq.)	SIFT	PolyPhen
ENST00000338516	PAX4	missense_variant			0.000 Deleterious	0.997 Probably Damaging
ENST00000341640	PAX4	missense_variant			0.000 Deleterious	1.000 Probably Damaging
ENST00000378740	PAX4	missense_variant				
ENST00000463946	PAX4	missense_variant				
ENST00000477423	PAX4	non_coding_transcript_exon_variant				
ENST00000483494	PAX4	NMD_transcript_variant missense_variant			0.000 Deleterious	1.000 Probably Damaging
NM_001366110.1		missense_variant			0.000 Deleterious	1.000 Probably Damaging
NM_001366111.1		missense_variant			0.000 Deleterious	0.999 Probably Damaging

トランスクriプト毎の
SIFT, Polyphen2のスコアを表示

Publications

Showing 1 to 10 of 27 entries				Show 10 entries
PMID	Reference	Year	Cited by	
34118937 (PubTatorCentral)	Evaluation of the effect of MTNR1B rs10830963 gene variant on the therapeutic efficacy of nateglinide in treating type 2 diabetes among Chinese Han patients. Song JF, Zhang J, Zhang MZ, Ni J, Wang T, Zhao YQ, Khan NU <i>BMC Med Genomics.</i> 2021-06-12;14(1):156.	2021-06-12	0	
32489678 (PubTatorCentral)	Genetics and Pathophysiology of Maturity-onset Diabetes of the Young (MODY): A Review of Current Trends. Yahaya TO, Ufumora SB <i>Oman Med J.</i> 2020-05;35(3):e126.	2020-05-01	1	
32801813 (PubTatorCentral)	Identification of the First PAX4-MODY Family Reported in Brazil. Abreu GM, Soares CAPD, Tarantino RM, da Fonseca ACP, de Souza RB, Pereira MFC, Cabello PH, Rodacki M, Zajdenverg L, Zembrzuski VM, Campos Junior M <i>Diabetes Metab Syndr Obes.</i> 2020;13():2623-2631.	2020-01-01	0	
31118371	A Variation in the ABCC8 Gene Is Associated with Type 2 Diabetes Mellitus and Repaglinide Efficacy in Chinese Type 2 Diabetes Mellitus Patients.	2010-		

バリエントに関する論文を2つのデータベース(PubTator, LitVar)から取得

解説動画あります (TogoTV)

The screenshot shows the TogoTV website interface. At the top, there's a navigation bar with links to DBCLS, Research, Services, Contact, and About. Below that is the TOGO-TV logo and a sub-header: 生命科学系DB・ツール使い始め系チャンネル. A search bar is present. The main content area has a title: 2018-08-25 TogoVar でヒトゲノムに存在するバリアントに関する情報を調べる. A video player window displays a video titled "TogoVar でヒトゲノムに存在するバリアントに関する情報を調べる". A large blue callout bubble points to the video player with the text: "TogoVar (https://togovar.biosciencedbc.jp/) はヒトゲノムに存在するバリアントに関する情報を、分子生物学的な情報や頻度、関連論文を知ることが出来るサイトです。". Below the video player, there's a link to YouTube and some descriptive text about TogoVar.

TogoVar TogoTV

すべて 地図 動画 画像 ショッピング もっと見る 設定 ツール

約 93 件 (0.40 秒)

togotv - YouTube

<https://im.youtube.com/user/togotv/feed> ▾

3日前 - TogoVar (日本入ゲノム多様性統合データベース) は、国立研究開発法人科学技術振興機構バイオサイエンスデータベースセンター (NBDC)と大学共同利用機関法人情報・システム研究機構 データサイエンス共同利用 ...

TogoVar でヒトゲノムに存在するバリアントに関する情報を ... - Togo...

<https://togovar.biosciencedbc.jp/ja/20180825.html> ▾

3日前 - TogoVar (日本入ゲノム多様性統合データベース) は、国立研究開発法人科学技術振興機構バイオサイエンスデータベースセンター (NBDC)と大学共同利用機関法人情報・システム研究機構 データサイエンス共同利用基盤高度 ワイフ ...

この機能を引出す際は<http://www.w3.org/2013/05>をご利用ください。 DOI:10.7202/TogoTV-2018-158

この記事は、クリエイティブ・コモンズ・ライセンスの下で利用可能ですが、利用した場合は、下記のリンクを必ず掲載してください。
© NBDC, JST | TogoTV

- 操作方法等の解説動画 (8分36秒)
 - Youtube、TogoTV

頻度データ (tsvファイル) の一括ダウンロードURL

- (<https://togovar.biosciencedbc.jp/downloads>)

The screenshot shows the 'Downloads' section of the Togovar website. The page header includes the Togovar logo and navigation links for Home, Datasets, Downloads, Terms, Contact, About, History, Help, and user account information. The main content area is titled 'Downloads' and contains a table with two columns: 'Chromosome' and 'Variation frequency'. The 'Variation frequency' column lists chromosomes 1 through MT with their respective tsv files and sizes. The 'Molecular annotation' column lists chromosomes 1 through MT with their respective tsv files and sizes. A note at the top of the table reads: 'Togovarデータは以下のリンクからダウンロード可能です。詳細なREADME.txtをご覧ください。' (Togovar data can be downloaded from the following links. Please refer to the detailed README.txt for more information.)

Chromosome	Variation frequency	Molecular annotation
1	chr 1 frequency [tsv] (131 MB)	chr 1 molecular annotation [tsv] (119 MB)
2	chr 2 frequency [tsv] (119 MB)	chr 2 molecular annotation [tsv] (119 MB)
3	chr 3 frequency [tsv] (95 MB)	chr 3 molecular annotation [tsv] (101 MB)
4	chr 4 frequency [tsv] (85 MB)	chr 4 molecular annotation [tsv] (81 MB)
5	chr 5 frequency [tsv] (103 MB)	chr 5 molecular annotation [tsv] (90 MB)
6	chr 6 frequency [tsv] (84 MB)	chr 6 molecular annotation [tsv] (75 MB)
7	chr 7 frequency [tsv] (102 MB)	chr 7 molecular annotation [tsv] (90 MB)
8	chr 8 frequency [tsv] (70 MB)	chr 8 molecular annotation [tsv] (71 MB)
9	chr 9 frequency [tsv] (67 MB)	chr 9 molecular annotation [tsv] (56 MB)
10	chr 10 frequency [tsv] (67 MB)	chr 10 molecular annotation [tsv] (62 MB)
11	chr 11 frequency [tsv] (76 MB)	chr 11 molecular annotation [tsv] (75 MB)
12	chr 12 frequency [tsv] (72 MB)	chr 12 molecular annotation [tsv] (73 MB)
13	chr 13 frequency [tsv] (42 MB)	chr 13 molecular annotation [tsv] (35 MB)
14	chr 14 frequency [tsv] (49 MB)	chr 14 molecular annotation [tsv] (49 MB)
15	chr 15 frequency [tsv] (48 MB)	chr 15 molecular annotation [tsv] (49 MB)
16	chr 16 frequency [tsv] (58 MB)	chr 16 molecular annotation [tsv] (59 MB)
17	chr 17 frequency [tsv] (59 MB)	chr 17 molecular annotation [tsv] (50 MB)
18	chr 18 frequency [tsv] (34 MB)	chr 18 molecular annotation [tsv] (35 MB)
19	chr 19 frequency [tsv] (59 MB)	chr 19 molecular annotation [tsv] (53 MB)
20	chr 20 frequency [tsv] (34 MB)	chr 20 molecular annotation [tsv] (30 MB)
21	chr 21 frequency [tsv] (20 MB)	chr 21 molecular annotation [tsv] (19 MB)
22	chr 22 frequency [tsv] (27 MB)	chr 22 molecular annotation [tsv] (25 MB)
X	chr X frequency [tsv] (49 MB)	chr X molecular annotation [tsv] (41 MB)
Y	chr Y frequency [tsv] (265 KB)	chr Y molecular annotation [tsv] (175 KB)
MT	chr MT frequency [tsv] (32 KB)	chr MT molecular annotation [tsv] (33 KB)

Latest release: 20190723

dbSNP:

<https://www.ncbi.nlm.nih.gov/snp/>

dbVar:

<https://www.ncbi.nlm.nih.gov/dbvar/>

jMorp:

<https://jmorp.megabank.tohoku.ac.jp/>

HGVD:

<https://www.hgvd.genome.med.kyoto-u.ac.jp/>

gnomAD:

<https://gnomad.broadinstitute.org/>

TogoVar:

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

GWAS catalog:

<https://www.ebi.ac.uk/gwas/>

JENGER:

<http://jenger.riken.jp/>

ClinVar:

<https://www.ncbi.nlm.nih.gov/clinvar/>

- ・疾患とバリアントの関連を専門家により
マニュアルキュレーションしたデータベース

MGeND:

<https://mgend.med.kyoto-u.ac.jp/>

- ・日本版ClinVar

GTEEx:

<https://gtexportal.org/home/>

- ・遺伝子発現に関連するバリアントの情報を提供

DBKERO :

<http://kero.hgc.jp/>

- ・薬剤応答データを検索可能なデータベース

TogoTV:

<https://togotv.dbcls.jp/>

<https://togotv.dbcls.jp/course.html?id=PL0uaKHgcG00a1DeTTRteZSWzmpDTElmAK> (疾患に関連するバリアントや遺伝子発現の情報を調べる為の関連動画)

1. 調べるバリエントの参照配列を確認
2. 目的に応じたデータベースの利用
 1. 遺伝的背景
 2. 頻度情報 -> gnomAD, jMorp, HGVD, TogoVar
 3. GWASの結果-> GWAS catalog, JENGER
 4. eQTL -> GTEx
 5. 疾患と関連 -> ClinVar, MGeND
3. 生物学のデータベースについて質問があれば、NBDCにお問合せ下さい

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。

メール : humandbs@biosciencedbc.jp