

Supplementary Fig.1 Number of variants per clinical significance

a In the Clinical significance box, the number of variants per clinical significance is displayed for 1,052,579 ClinVar variants in TogoVar. Note that the number is counted for each significance when two or more clinical significances are assigned to one variant.

b The number per clinical significance for 123,160 variants included in the five datasets of JPN_UNION is displayed.

a

Statistics / Filters Reset

Dataset		^
<input type="checkbox"/>	All	1,052,579
<input type="checkbox"/>	WGS GEM-J WGA	138,718
<input type="checkbox"/>	WES JGA NGS	56,079
<input type="checkbox"/>	SNP JGA SNP	19,190
<input type="checkbox"/>	WGS ToMMo 8.3KJPN	146,704
<input type="checkbox"/>	WES HGVD	51,670
<input type="checkbox"/>	WGS gnomAD genomes	422,465
<input type="checkbox"/>	WES gnomAD exomes	521,805
<input checked="" type="checkbox"/>	Disease ClinVar	1,052,579

Alternative allele frequency v

Variant calling quality v

Variant type v

Clinical significance ^

<input checked="" type="checkbox"/>	All	1,052,579
<input checked="" type="checkbox"/>	Not in ClinVar	0
<input checked="" type="checkbox"/>	Pathogenic	154,665
<input checked="" type="checkbox"/>	Likely pathogenic	77,153
<input checked="" type="checkbox"/>	Uncertain significance	601,500
<input checked="" type="checkbox"/>	Likely benign	404,249
<input checked="" type="checkbox"/>	Benign	206,181
<input checked="" type="checkbox"/>	Conflicting interpretations of pathogenicity	18,709
<input checked="" type="checkbox"/>	Drug response	2,228
<input checked="" type="checkbox"/>	Association	365
<input checked="" type="checkbox"/>	Risk factor	771
<input checked="" type="checkbox"/>	Protective	74
<input checked="" type="checkbox"/>	Affects	192
<input checked="" type="checkbox"/>	Other	1,213
<input checked="" type="checkbox"/>	Not provided	17,833
<input checked="" type="checkbox"/>	Association_not found	2

Consequence ^

b

Statistics / Filters Reset

Dataset		^
<input type="checkbox"/>	All	123,160
<input checked="" type="checkbox"/>	WGS GEM-J WGA	101,183
<input checked="" type="checkbox"/>	WES JGA NGS	31,297
<input checked="" type="checkbox"/>	SNP JGA SNP	12,964
<input checked="" type="checkbox"/>	WGS ToMMo 8.3KJPN	104,709
<input checked="" type="checkbox"/>	WES HGVD	47,402
<input type="checkbox"/>	WGS gnomAD genomes	94,795
<input type="checkbox"/>	WES gnomAD exomes	87,336
<input type="checkbox"/>	Disease ClinVar	123,160

Alternative allele frequency v

Variant calling quality v

Variant type v

Clinical significance ^

<input type="checkbox"/>	All	123,160
<input type="checkbox"/>	Not in ClinVar	0
<input checked="" type="checkbox"/>	Pathogenic	5,466
<input checked="" type="checkbox"/>	Likely pathogenic	2,519
<input checked="" type="checkbox"/>	Uncertain significance	56,789
<input checked="" type="checkbox"/>	Likely benign	49,484
<input checked="" type="checkbox"/>	Benign	92,481
<input checked="" type="checkbox"/>	Conflicting interpretations of pathogenicity	3,159
<input checked="" type="checkbox"/>	Drug response	446
<input checked="" type="checkbox"/>	Association	176
<input checked="" type="checkbox"/>	Risk factor	235
<input checked="" type="checkbox"/>	Protective	43
<input checked="" type="checkbox"/>	Affects	49
<input checked="" type="checkbox"/>	Other	262
<input checked="" type="checkbox"/>	Not provided	1,515
<input checked="" type="checkbox"/>	Association_not found	2

Consequence ^

Supplementary Fig.2 How to access individual-level genome data in JGA

a Identify the allele frequency dataset (GEM-J WGA) that contains the variant of interest. **b** Focus on hum0184 among the six individual-level genomic datasets of GEM-J WGA. **c** Obtain an overview of the hum0184 dataset in the NBDC human database. **d** Submit a data-use application.

a

TogoVar A comprehensive Japanese genetic variation database **GRCh37** Home **Datasets** Downloads API Terms Contact About History Help

TogoVar datasets (GRCh37)

Variant frequencies for which you can apply for use of individual-level data*¹ to the NBDC human databases*²

Click the links at the Included controlled-access datasets to apply for use of individual-level data

Variant dataset name	Analysis method	Target population	Healthy subjects	Affected subjects	Sample size	Number of variants (# of sites)	Included controlled-access datasets
GEM Japan Whole Genome Aggregation (GEM-J WGA) Panel	WGS	Japanese	✓		7,609	95,863,463 (90,280,248)	6 datasets
JGA-NGS	WES	Japanese	✓	✓	125	4,679,025	7 datasets
JGA-SNP	SNP-Chip	Japanese	✓	✓	183,884	1,249,724	3 datasets

*1 : fastq/bam/cel files and/or lists of genotype data etc.
 *2 : [Japanese Genotype-phenotype Archive \(JGA\)](#) / [AMED Genome group sharing Database \(AGD\)](#)

b

WGS datasets used for joint variant calling [Controlled access]

If you would like to use the datasets, apply for data use of them whose ID begins with "JGAD" and "AGDS" to [the NBDC Human database](#) and [the AMED group sharing database](#), respectively.

Dataset ID (NBDC research ID)	Study title	Participants	Sample size	Data provider
JGAD000220 (hum0014)	The Tailor-made Medical Treatment Program (BioBank Japan: BBJ)	The cohort participants registered in the BBJ from 2003 to 2007	768	BioBank Japan
AGDS_00000000005 (agd0008)	バイオバンク・ジャパンの運営・管理と個別化医療の実現に向けた疾患バイオマーカー探索 (English page is under construction)	心筋梗塞、胃がん (非腫瘍組織)、認知症	2,089	BioBank Japan
JGAD000117 (hum0103)	To investigate genomic alterations of Japanese biliary tract cancers	Biliary tract cancer (non-tumor tissue)	17	RIKEN Center for Integrative Medical Sciences
JGAD000228 (hum0158)	To investigate genomic alterations of Japanese liver cancers	Liver cancer (non-tumor tissue)	220	RIKEN Center for Integrative Medical Sciences
JGAD000233 (hum0160)	To investigate genomic alterations of Japanese esophageal squamous cell carcinomas	Esophageal squamous cell carcinoma (non-tumor tissue)	20	RIKEN Center for Integrative Medical Sciences
JGAD000338 (hum0184)	Construction of Japanese Whole-Genome database	General residents	4,495	Tohoku Medical Megabank Organization
Total			7,609	

Note: Those datasets above provide fastq/bam file formatted data. The result data will be shown in our database soon. The sample size of each dataset indicates the sample number after quality control in this current study.

c

Human data NBDC Human Database Japanese Site Search

Home **Data Use** Data Submission Guidelines Off-premise Server Data Access Committee Publications Contact us

NBDC Research ID: hum0184.v1

SUMMARY

Aims: Tohoku University Tohoku Medical Megabank Organization (ToMMo) and Iwate Tohoku Medical Megabank Organization (IMM) were founded to establish an advanced medical system to foster the reconstruction from the Great East Japan Earthquake. These organizations are developing a biobank that includes medical and genome information for supporting health and welfare in the Tohoku area. In the first stage, the part of our mission was to sequence the 4,000 individuals to construct Japanese whole-genome reference panel.

Methods: Whole genome sequencing

Participants/Materials: 4,566 Japanese general residents

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

Data Set ID	Type of Data	Criteria	Release Date
JGAS000239	NGS (WGS)	Controlled Access (Type II)	2020/09/01

*Release Note
 * Data users need to apply the [Form 2 \(Application Form for Using NBDC Human Data\)](#) to reach the Controlled Access Data. [Learn more](#)

MOLECULAR DATA

JGAS000239

Participants/Materials:	4,566 Japanese general residents
Targets	WGS
Target Loci for Capture Methods	-
Platform	Illumina [HiSeq 2500, NovaSeq 6000]

d

NBDC Application system for the NBDC Human Database Japanese English

Menu
 IDP
 Application for data use
 Application for data submission
 Create Group
 My Data User Groups
 My Data Submitter Groups

New Application

Group: PI Research collaborator Study Ethics review Attachments Overview

Information on the principal investigator (PI)

* Last Name (in Japanese): * First Name (in Japanese):

Name:

* Telephone number:

E-mail address:

* Affiliated institution (in Japanese):

* Division (in Japanese):

Affiliated institution:

Division:

* Job title (in Japanese):

* Job title:

Mailing address:

Information on the head of the affiliated institution

* Name: * Job title:

* Telephone number: * E-mail address: