Gene variant databases

sharing information



(LSDBs)



Johan den Dunnen









My affiliations



get all variants/consequences found easily accessible



standards for variant description and databases



software for webbased gene databases



DNA diagnostics

..is based on

SHARING what we know on the relation between variants in genes & phenotypes

without sharing, no DNA diagnostics

..do we share?
..are databases supported?



DNA diagnostics

the work

analyse sample (sequence) identify variants

check what others found draw initial conclusion ..perform additional experiments RNA, protein, cells, computational, animal model, .. plan to publish results

share with colleagues > submit to database





DNA diagnostics

the work

analyse sample (sequence) identify variants



share with colleagues > submit to database

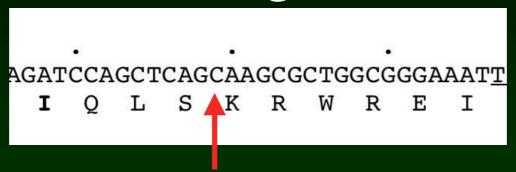
check what others found draw initial conclusion ..perform additional experiments RNA, protein, cells, computational, animal model, .. plan to publish results

It seems so simple...



Example

DMD gene



c.5859C>T r.(?) p(Ser1953=)

found in diagnosis

prenatal at risk family for DMD no definite diagnosis

found in diagnosis

WES, trio analysis male parent 45y, healthy

you may have life saving information, do you realize this??





...privacy



LMNA: c.?

...we have identified a variant in this gene in a patient with a certain phenotype.

For privacy reasons we are not able to share more information.

For details please contact us.

VUS

Variant of Unsufficient Sharing

Organisations

HVP & HGVS

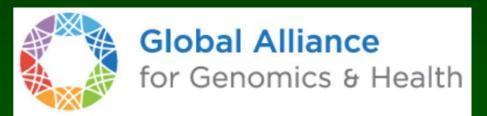
Human Variome Project
Human Genome Variation Society
umbrella organisations promoting collecting
genome variation & establishing standards



GA4GH

Global Alliance for Genomics and Health

recent initiative, similar goals active on many aspects



• EBI & NCBI

reference sequences genome browsers variant databases

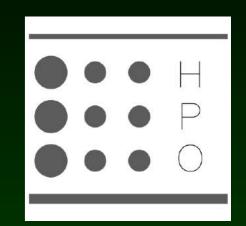






Databases require?

• standard to describe phenotypes HPO (Human Phenotype Ontology)



- standard to describe variants in DNA HGVS nomenclature
- databases sharing information





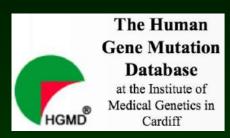




DNA variant DBs

central databases

inch deep, mile wide



dbSNP

HGMD

lists first published report > no frequencies pathogenic only (most)





list firsts report(s) & some interesting cases

dbSNP, EVA

all variants (originally mainly non pathogenic)

many others
COSMIC, DBVar, Decipher, ...





• gene variant databases

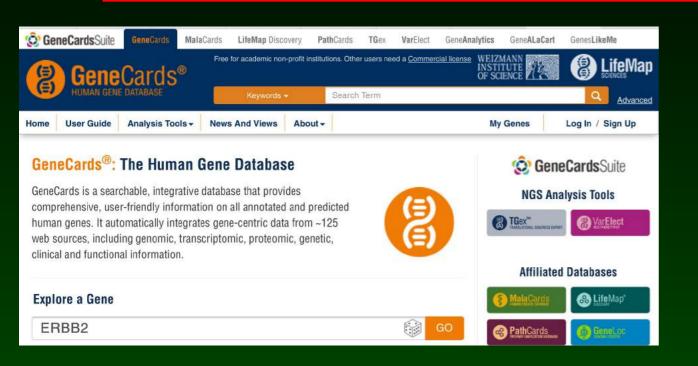
all details per gene all variants unpublished data fi

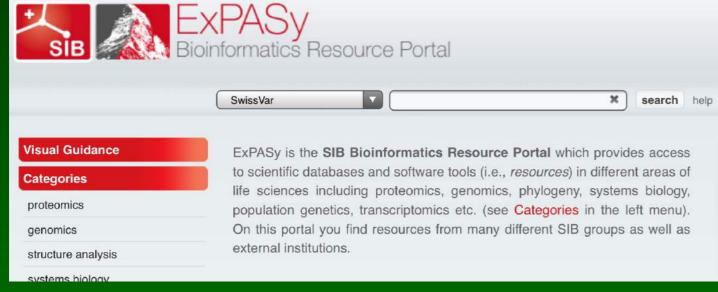
inch wide, mile deep

unpublished data from direct DB submission



Other databases







The portal for rare diseases and orphan drugs

"Rare diseases are rare, but rare disease patients are numerous"

Access our Services





Inventory of orphan drugs



Directory of patient organisations



Directory of professionals and institutions



Directory of expert centres



Directory of medical laboratories providing diagnostic tests



Directory of ongoing research projects, clinical trials, registries and biobanks





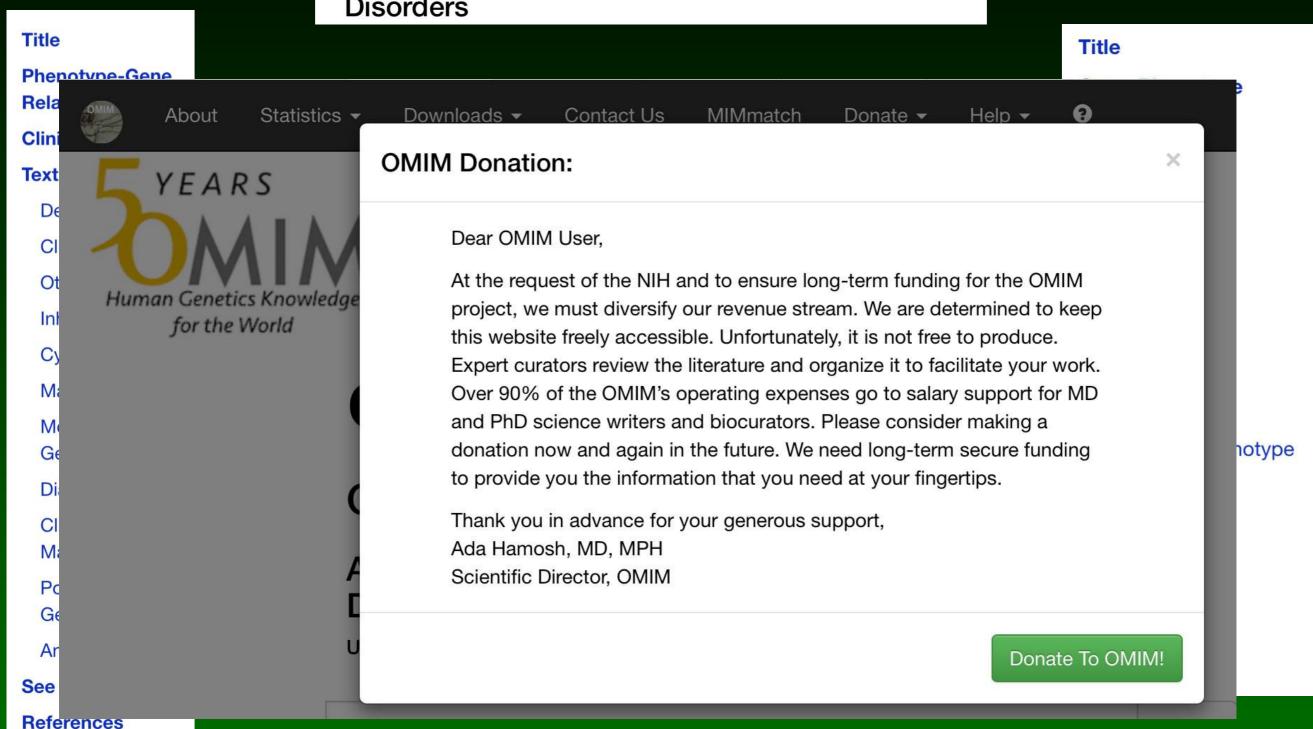


- PROSITE

OMIM[®]

Online Mendelian Inheritance in Man®

An Online Catalog of Human Genes and Genetic Disorders





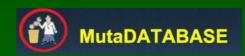
Definition

Locus Specific DataBase (LSDB)

A listing of sequence variants in a specific gene causing a Mendelian disorder or a change in phenotype, curated by an expert in that gene

(gene variant database)









CCAAGAGATATATCT
TCACTTAGACCTCAC
AAGTCAGGGCAGAGC
CTCCTGAGGAGAAGT
GGTTACAAGACAGGT
TGCCTATTGGTCTAT



ClinVar aggregates info



LSDBs

• gene variant databases (LSDBs)

use HGVS description variants link to other relevant sources store ALL variants

from publications AND direct submissions detailed phenotype data (...often limited) active curator searching, asking, adding, ...

variable formats / quality / updating

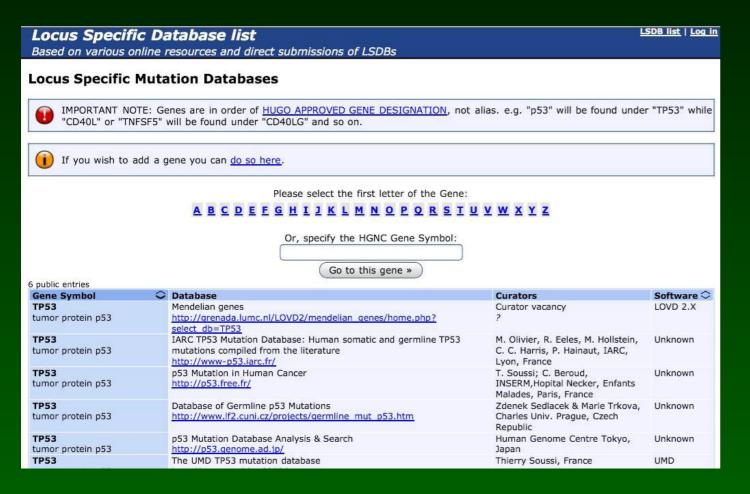
A hit means data on consequences for function are available

...but not a DB for all genes



Is there a DB?

try **GeneSymbol** .variome.org

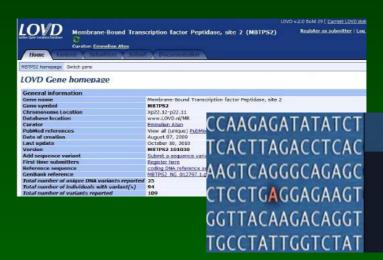


TP53.variome.org



2 second acknowledgement

always a hit in LOVD and ClinVar



ClinVar

ClinVar aggregates in

when not ...?



Collaboration

ACTGATGGTATGGGGCCAAGAGATATATCT CAGGTACGGCTGTCATCACTTAGACCTCAC CAGGGCTGGGCATAAAAGTCAGGGCAGAGC CCATGGTGCATCTGACTCCTGAGGAGAAGT GCAGGTTGGTATCAAGGTTACAAGACAGGT

ClinVar

ClinVar aggregates information about genomic variation and its relationship to human health.

collaboration initiated, exchange data



3 LOVD 3.0 shared installation

LOVD is supported by:

LOVD v.3.0 Build 17 [Current LOVD status] Register as submitter | Log in

Individuals X Diseases X

Screenings

Submit)

Documentation

approaching other databases to merge their data

Your country?

xx.LOVD.org

xx = country code

Malaysia DNA Variant Database

Based on:

MY

Home Variants in individuals from Malaysia

Variants by submitters from Malaysia

This resource automatically retrieves information from our <u>LOVD3 shared installation</u>. The information retrieved is based on the geographic information per country (here for **Malaysia**).

Variants shown are either linked to the country of origin of an *individual* (patient), or to the country of origin of the *submitter* of ta. Note the difference: data from an *individual* in Germany can be submitted by a *submitter* from Belgium.

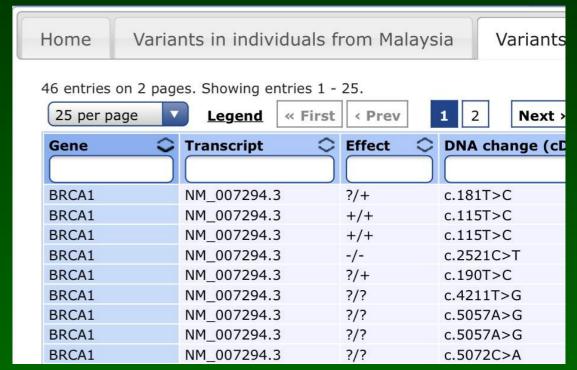
Home Variants in individuals from Malaysia

Variants

Queries can be submitted by using the search boxes in the column's header.

244 entries on 10 pages. Showing entries 1 - 25. 1 2 3 4 25 per page « First < Prev Legend Transcript Gene Effect DNA change (cl ABCC11 NM 032583.3 c.538G>A +/+ ABCC11 NM 032583.3 +/+ c.538G>A ABCC11 NM 032583.3 +/+ c.538G>A ABCC11 NM_032583.3 +/+ c.538G>A APC NM 000038.5 +/+ c.847C>T

244 variants



46 submissions



Your country?

xx.LOVD.org

xx = country code

Singapore DNA Variant Database

Based on:

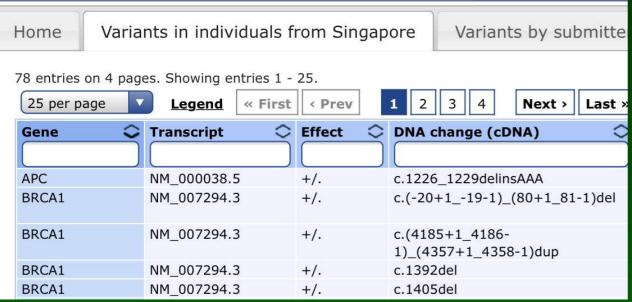
SG

Home Variants in individuals from Singapore Variants by submitters from Singapore

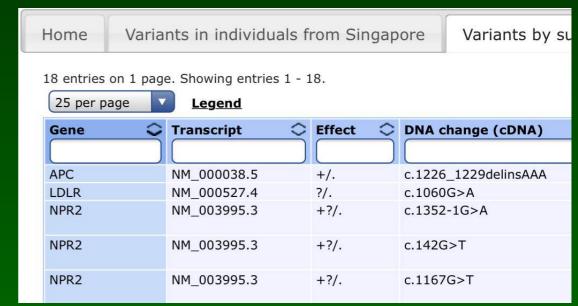
This resource automatically retrieves information from our <u>LOVD3 shared installation</u>. The information retrieved is based on the geographic information per country (here for **Singapore**).

Variants shown are either linked to the country of origin of an *individual* (patient), or to the country of origin of the *submitter* of the data. Note the difference: data from an *individual* in Germany can be submitted by a *submitter* from Belgium.

Queries can be submitted by using the search boxes in the column's header.



78 variants



18 submissions



Home

Variants in individuals from Malaysia

Variants

46 entries on 2 pages. Showing entries 1 - 25.

<u>Legend</u>	« First	< Prev	1	2	Next >the
	<u>Legend</u>	Legend « First	Legend « First < Prev	Legend « First < Prev 1	Legend « First < Prev 1 2

200		100.00	
Gene	Transcript	C Effect C	DNA change (cl
BRCA1	NM_007294.3	?/+	c.181T>C
BRCA1	NM_007294.3	+/+	c.115T>C
BRCA1	NM_007294.3	+/+	c.115T>C
BRCA1	NM_007294.3	-/-	c.2521C>T
BRCA1	NM_007294.3	?/+	c.190T>C
BRCA1	NM_007294.3	?/?	c.4211T>G
BRCA1	NM_007294.3	?/?	c.5057A>G
BRCA1	NM_007294.3	?/?	c.5057A>G
BRCA1	NM 007294.3	?/?	c.5072C>A



Hor

LSDB-in-a-Box

- software
 - Open Source, platform independent freely available > http://www.LOVD.nl follows existing HGVS guidelines LSDB content, description of variants
- fully WWW-based
 display of data
 data management
 access levels incl. manager, curator, submitter
 - free hosting on Leiden servers

©Ivo Fokkema

- runs on any platform
 laptop / PC, intranet, internet
- make any database pre-programmed for DNA / RNA / protein



LOVD connections





(can be country nodes)











contact through central site LSDB list, web service (exome queries)

data flow

lab > gene variant database > central repository



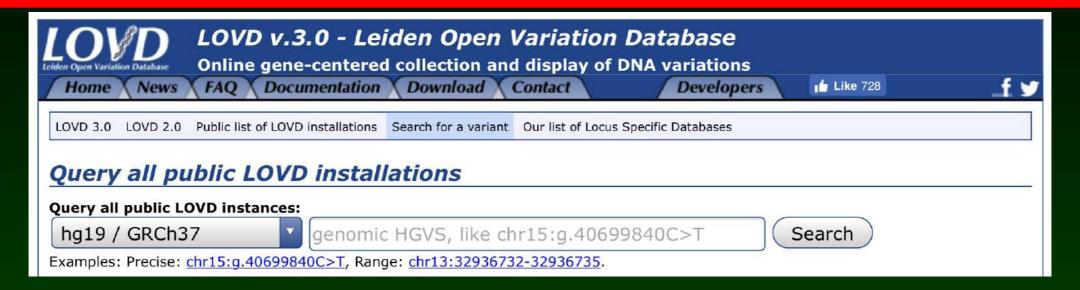


LOVDs world-wide

In total: 292,384,028 variants (3,953,852 unique)	in 763,965 indivi	duals in 85 LOVD installations.	
http://bipmed.iqm.unicamp.br/snparray/	LOVD 3.0-20	17391 genes	Leiden Open Variation Dat
BIPMed SNP Array	A1BG-AS1,A1CF,A2M,A	A2M-AS1,A2ML1,A2MP1,A3GALT2,A4GALT,A4GN	9022/3 unique
http://bipmed.iqm.unicamp.br/	LOVD 3.0-21	20930 genes	66158522 variants
BIPMed WES	A1BG,A1BG-AS1,A1CF	A2M,A2M-AS1,A2ML1,A2MP1,A3GALT2,A4GALT	622610
http://databases.lovd.nl/whole_genome/	LOVD 3.0-20a	22002 genes	1998175
Whole genome datasets	A1BG,A1BG-AS1,A1CF	,A2M,A2M-AS1,A2ML1,A2MP1,A4GALT,A4GNT,A	1998135
https://databases.lovd.nl/shared/	LOVD 3.0-21	22980 genes	757275 '
Global Variome shared LOVD	A1BG,A1BG-AS1,A1CF	,A2LD1,A2M,A2M-AS1,A2ML1,A2MP1,A3GALT2,	173628 Lque
http://proteomics.bio21.unimelb.edu.au/lovd/	LOVD 3.0-07	14772 genes	239690 variants
LOVD - Leiden Open Variation Database	A1BG,A1BG-AS1,A1CF	,A2M,A2ML1,A4GALT,A4GNT,AAAS,AACS,AACSP	152241
http://bipmed.iqm.unicamp.br/cfa/	LOVD 3.0-21	41 genes	134333 GEN2PHEN
Craniofacial anomalies	APOC2,APOC4-APOC2	AXIN2,BCL3,BMP4,CLPTM1,DVL2,ERBB2,FGF22	257 unique
https://ab-openlab.csir.res.in/mitolsdb/	LOVD 2.0-35	37 genes	112662 variants
MitoLSDB	MTATP6,MTATP8,MTCC	01,MTCO2,MTCO3,MTCYB,MTND1,MTND2,MTND3,M	4660 unique
http://www.insight-database.org/	LOVD 3.0-21	17 genes	36826 variants
International Society for Gastrointestinal Hereditary Tumours Database	APC,BMPR1A,EPCAM,F	BXO11,GALNT12,MLH1,MLH3,MSH2,MSH3,MSH6,	5836 unique
http://HCI-LOVD.hci.utah.edu	LOVD 2.0-33	8 genes	26167 variants
LOVD - human mismatch repair genes	MLH1,MLH1_priors,MS	SH2,MSH2_priors,MSH6,MSH6_priors,PMS2,P	26167 unique



Query all LOVDs



Variant location matches your query exactly

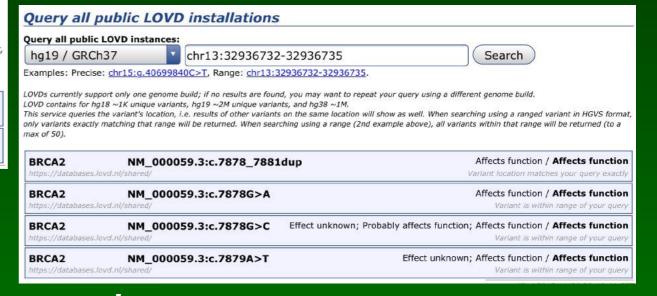
Variant location matches your query exa

Affects function / Probably affects function

using a position

NM 002225.3:c.157C>T

LOVDs have an API



using a range

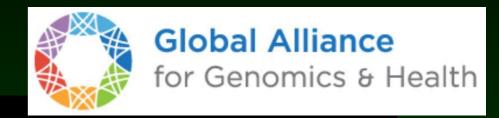


IVD

http://databases.lovd.nl/whole_gen

Beacons





A global search engine for genetic mutations.

GRCh37 →

e.g. 1:100,000 A>C

Search

Quickstart: Search for a BRCA2 variant

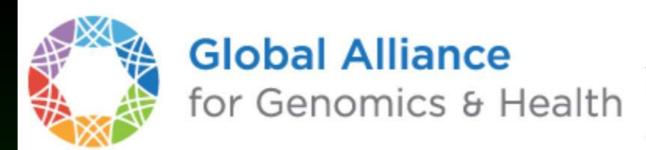
Beacon Project

Being implemented on the website of the world's top genomic organizations to test the willingness of international sites to share genetic data.



there are simply toooooo many databases







brcaexchange.org

BRCA Challenge

The BRCA Challenge aims to advance understanding of the genetic basis of breast and other cancers using data from around the world.



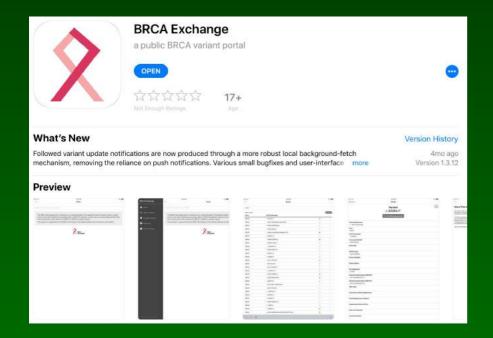
BRCA1

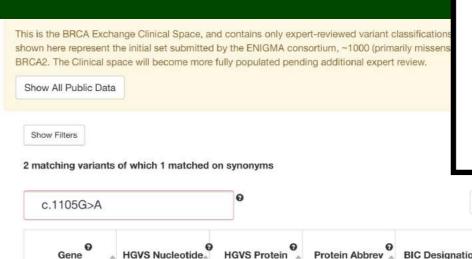
BRCA1

c.1105G>A

c.993G>C

BRCA Exchange **Expert Reviewed** HOME VARIANTS COMMUNITY search for "c.1105G>A", "brca1" or "IVS7+1037T>C"





HGVS Protein

p.(Asp369Asn)

p.(Arg331Ser)

Protein Abbrev

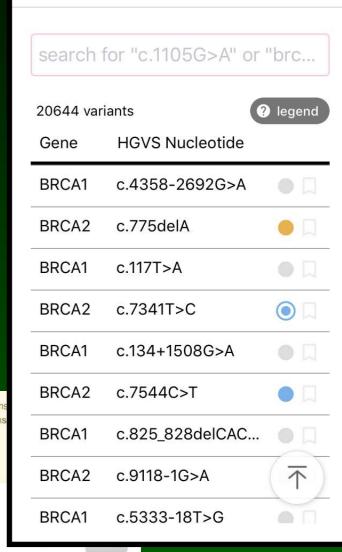
BIC Designation

1224G>A

Significance

Benign / Little Clinical Significance

Not Yet Classified



Search









Shared database DMD (dystrophin) 🎅

Curator: Johan den Dunnen

Genes **Transcripts** **Variants**

Individuals Diseases Screenings

Graphs

Reading frame checker

Submit

Documentation

DMD gene homepage

Unique public DNA variants reported

Individuals with public variants

Hidden variants

Notes

This database is one of the gene variant databases from the Leiden Muscular Dystrophy pages.

, <u> </u>		
General information		
Gene symbol	DMD	Gra
Gene name	dystrophin	Gra
Chromosome	X	Rea
Chromosomal band	p21.2	
Imprinted	Unknown	UCS
Genomic reference	NG_012232.1	Ens
Transcript reference	NM_004006.2	NCE
Exon/intron information	NM_004006.2 exon/intron table	
Associated with diseases	all-in-one, BMD, BMD/DMD, CMD-3B,	DMD
Citation reference(s)		
Refseq URL	Genomic reference sequence	
Curators (1)	Johan den Dunnen	

UCSC Genome Browser Ensembl Genome Browser NCBI Sequence Viewer

Graphical displays and utilities Graphs displaying summary information of all variants in the database » The Reading-frame checker generates a prediction of the effect of wholeexon changes. Active for: NM_004006.2.

> Show variants in the UCSC Genome Browser (full view, compact view) Show variants in the Ensembl Genome Browser (full view, compact view) Show distribution histogram of variants in the NCBI Sequence Viewer

Total number of public variants reported 15850 4925 36693 1468 This database is one of the gene variant databases from

Leiden Muscular Dystrophy pages®

the:

When refering to this database please cite Aartsma-Rus et al. (2006). Muscle Nerve. 34:135-144 and/or White SJ, den Dunnen JT (2006). Cytogenet. Genome Res. 115: 240-246.

Date created July 29, 1997 March 09, 2018 Date last updated DMD:180309 Version

Links to other resources

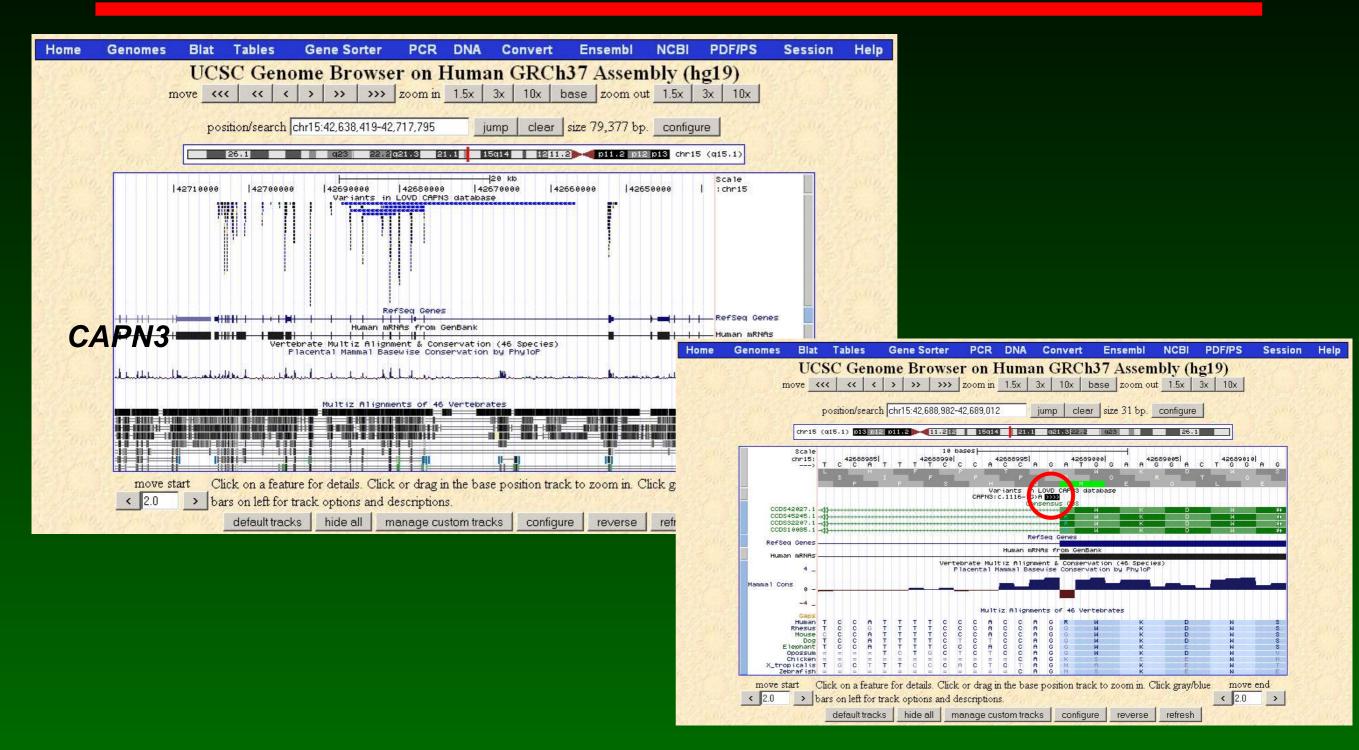
Homepage URL	http://www.LOVD.nl/DMD
External URL	the Leiden Muscular Dystrophy pages Orphanet
HGNC	2928
Entrez Gene	<u>1756</u>
PubMed articles	<u>DMD</u>
OMIM - Gene	300377
OMIM - Diseases	BMD (dystrophy, muscular, Becker type (BMD)) CMD-3B (cardiomyopathy, dilated, type 3B (CMD-3B)) DMD (dystrophy, muscular, Duchenne type (DMD))
HGMD	DMD
GeneCards	<u>DMD</u>
GeneTests	DMD

Copyright & disclaimer

The contents of this LOVD database are the intellectual property of the respective curator(s). Any unauthorised use, without written permission from the curator(s) will lead to copyright infringement with possible ensuing litigation. Co further details, refer to Directive 96/9/EC of the European Parliament and the Council of March 11 (1996) on the lega

We have used all reasonable efforts to ensure that the information displayed on these pages and contained in the dal express or implied, as to its accuracy or that the information is fit for a particular purpose, and will not be held respo inaccuracies or omissions. Individuals, organisations and companies which use this database do so on the understant indirect shall rest upon the curator(s) or any of their employees or agents for the effects of any product, process or n part, notwithstanding that the formulation of such product, process or method may be based upon information here provided.

LOVD to genome





Genome browsers

genome browsers start to accept HGVS descriptions

immediately jumping to position/change in genome



Request:	Genome Browser Response:
chr7	Displays all of chromosome 7
chrUn_gl000212	Displays all of the unplaced contig gl000212
20p13	Displays region for band p13 on chr 20
chr3:1-1000000	Displays first million bases of chr 3, counting from p-arm telomere
chr3:1000000+2000	Displays a region of chr3 that spans 2000 bases, starting with position 1000000
RH18061;RH80175 15q11;15q13	Displays region between genome landmarks, such as the STS markers RH18061 and RH80175, or chromosome bands 15q11 to 15q13, or SNPs
rs1042522;rs1800370	rs1042522 and rs1800370. This syntax may also be used for other range queries, such as between uniquely determined ESTs, mRNAs, refSeqs, etc.
D16S3046	Displays region around STS marker D16S3046 from the
n. 4950a	Genethon/Marshfield maps. Includes 100,000 bases on each side as well.
AA205474	Displays region of EST with GenBank accession AA205474 in BRCA1 cancer gene on chr 17
AC008101	Displays region of clone with GenBank accession AC008101
AF083811	Displays region of mRNA with GenBank accession number AF083811
PRNP	Displays region of genome with HUGO Gene Nomenclature Committee identifier PRNP
NM_017414	Displays the region of genome with RefSeq identifier NM_017414
NP_059110	Displays the region of genome with protein accession number NP_059110
pseudogene mRNA	Lists transcribed pseudogenes, but not cDNAs
homeobox caudal	Lists mRNAs for caudal homeobox genes
zinc finger	Lists many zinc finger mRNAs
kruppel zinc finger	Lists only kruppel-like zinc fingers
huntington	Lists candidate genes associated with Huntington's disease
zahler	Lists mRNAs deposited by scientist named Zahler
Evans, J.E.	Lists mRNAs deposited by co-author J.E. Evans







LOVD - Leiden Open Variation Database

LOVD v.2.0 Build 11 [Current LOVD status] Register as submitter | Log in

Growth Hormone Receptor (GHR)

Home

Variants

Submitters

Submit

Documentation

GHR homepage

Switch gene



LOVD - Leiden Open Variation Database

Growth Hormone Receptor (GHR)

LOVD v.2.0 Build 11 [Current LOVD status]

LOVD v.2.0 Build 11 [Current LOVD status]

Your account | Your submissions | Log out

Home

GHR homepage

Variants

Switch gene

Submitters

Submitters

Submit

Documentation

submitter colleague

Welcome, Cecilia Camacho-Hübner Camacho-Hübner

LOVD - Leiden Open Variation Database

Insulin-like Growth Factor 1 (IGF1)

Configuration

Documentation

curator collaborator

Welcome, Jan Maarten Wit

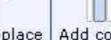
Your account | Log out

Switch gene

Variants

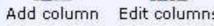






Submit (



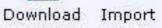












LOVD - Leiden Open Variation Database

Signal Transducer and Activator of Transcription 5B (STAT5B)

LOVD v.2.0 Build 11 [Current LOVD. Welcome, Johan den D Your account | L

variants

Submitters

Submit

Configuration

Setup

Documentation









Updated:2008-09-01

























Edit subs | Lo

LOVD Setup

Installed:2008-06-12

Leiden Open Variation Database

Statistics

Users:12 Submitters:8

Genes:5 Log entries:395

Variants

Total:92 Submitted:24

Non public:2

Marked:8

Public

Show your data

Leiden, Netherlands Carli Tops

Leiden University Medical Center Clinical Genetics

84 entries in MLH1

Path 🛭 🗘 🔿	Exan	DNA change 🗢	DNA_reported	🛇 🗘 RNA change	🛇 🗘 Protein
+/?	01	c.18_34del17	-	r.(?)	p.Val7_ArgfsX18
+/?	01	c.18_34del17	-	r.(?)	p.Val7_ArgfsX18
+/?	01	c.18_34del17	-	r.(?)	p.Val7_ArgfsX18
?/?	01	c.102_103delGA	-	r.(?)	p.Glu34AspfsX3

You have not sent any submissions to the MLH3 gene database yet!

60 entries in MSH2

Path 🙉 🗘	Even 8	DNA change	C D)NA_reported © ≎	>	RNA change	90	Protein
?/?	00	c.(?68)_*+?del	d	lel 1_16		-		p.0?
?/?	00	c.(?68)_*+?del + MSH6 (1)		lel MSH2 ex1_16, lelTACSTD1-15,-27, del MSH6		-		p.0?
?/?	00	c.(?68)_*+?del	d	lel 1_16		-		p.0?

62 entries in MSH6

4	Path. 🛭 🗘	Ехор	DNA change 🗘	DNA_reported 🚳 🗘	RNA change 🛛 🗘	Protein 80	Codon_nr
	?/?	00	c.1-?_*+?del	del MSH2 ex1_16,	-	p.0?	-
			+ MSH2 (1)	delTACSTD1-15,-27, del MSH6			
	?/?	01	c.1-18G>T	-	-	p.?	-
	?/?	01	c.1-?_457+?del	deletion of 21.6 kb around exon	-	p.Met1_Gly153>ValfsX21	-
				1+2			

You have not sent any submissions to the PMS1 gene database yet!

9 entries in PMS2

Path.	Exon & C	DNA change 🔾	DNA_reported 💆 🔾	RNA change 🚳 🤇	Protein
+/?	03	c.219_220dup + MLH1 (2)	-	r.(?)	p.Gly74ValfsX3
?/?	07i	c.?	g.46227+?_47919-?ins	r.(?)	-
?/?	08	c.861_864delACAG + c.1688G>T, MLH1 (1)	c.856delG	r.(?)	p.Arg287SerfsX19





4 databases in 1

Variants \(\) Individuals \(\) Diseases \(\) Screenings

- Individuals
 details on the Individual
 fields fixed per database installation
- Phenotypes
 details on Phenotype(s) per Individual
 one Individual can have several Phenotypes
 fields variable per Phenotype
- Screenings
 details on Screening(s) performed
 fields fixed per database installation
- Variants
 details on Variant(s) identified
 fields variable per Gene



LOVD 3.0 shared installation





4 in 1: Individual

Transcripts Individuals Submit Diseases Screenings Variants Genes Individual #00054908 **Individuals** Screenings Variants Diseases Individual_ID Reference Journal: O'Rawe 2015 Remarks 2-generation family, 3 affected brothers, unaffected heterozygous carrier mother Gender Consanguinity Country Colombia **Population** >9y (later than 9 years) Age/Death VIP 0 Data av **Treatment Panel size** 3 ID Diseases Johan den Dunnen Owner name





4 in 1: Phenotype

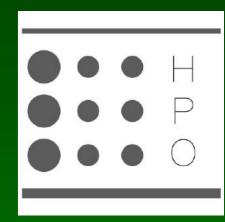
Phenotype #0000041575

Individual ID	00054908
Associated disease	
Diagnosis/Initial	intellectual disability
Diagnosis/Definite	MRXS-33
Inheritance	Familial, X-linked recessive
Age/Examination	09y (9 years)
Birth_Details	36w, caesarian section (HP:0011410), weight 4480 (1.85), height 52 (0.48), OFC 36.5 (1.02)
Age/Diagnosis	
Age/Onset	-
Phenotype/Onset	-
Intellectual_dis	seizures (4 hours after birth), lactic acidosis.; postnatal growth retardation (HP:0008897); delayed gross motor development (HP:0002194); delayed speech and language development (HP:0000750); oral-pharyngeal dysphagia (HP:0200136); prominent supraorbital ridges (HP:0000336); downslanted palpebral fissures (HP:0000494); prominent forehead (HP:0011220); sagging cheeks; long philtrum (HP:0000343); low-set ears (HP:0000369); protruding ear (HP:0000411); thickened helices (large earlobe HP:0009748); long face (HP:0000276); high arched palate (high palate HP:0000218); thin upper lip (thin upper lip vermilion HP:0000219); pointed chin (HP:0000307); broad upturned nose (anteverted nares HP:0000463); hypertelorism (HP:0000316); sacral dimple (HP:0000960); hirsutism (HP:00100797); frequent dermatitis & eczema (eczema HP:0000961); toenail dysplasia (HP:0100797); hearing impairment (HP:0000365); chromic otitis media (HP:0000389); strabismus (HP:0000486); constipation (HP:0002019); gastroesophageal reflux (HP:0002020); microcephaly (HP:0000252); cerebellar atrophy (low cerebral white matter volume); hypoplasia of the corpus callosum (HP:0002079); seizures (HP:0001250); generalized hypotonia (HP:0001290); non-ambulatory; sleep-wake cycle disturbance (HP:0006979); osteopenia (HP:0000988); unusual gluteal crease with sacral caudal remnant/sacral dimple (abnormal sacral segmentatino, HP:0008468), prominent protruding coccyx (HP:0008472); distal joint hypermobility (HP:0001382); kyphosis (HP:0002808); scoliosis (HP:0000729); intellectual disability (HP:0001249); not present -HP:0000490, -HP:0000496, -HP:0000579, -HP:0000739, -HP:0001057, -HP:0001251, -HP:0001257, -HP:0001264, -HP:00007315, -HP:0001151, -HP:0002111, -HP:0002111, -HP:0002111, -HP:0002111, -HP:0002111, -HP:0007018, -HP:0007375
THE RESIDENCE OF THE PROPERTY OF THE PARTY O	
Speech	
Development	- Johan dan Dunnan
Owner name	Johan den Dunnen

HPO

Diseases

HP: -HP: ?HP:



Screenings





LOVD - Leiden Open Variation Database

Insulin-like Growth Factor 1 Receptor (IGF1R)

Submitters

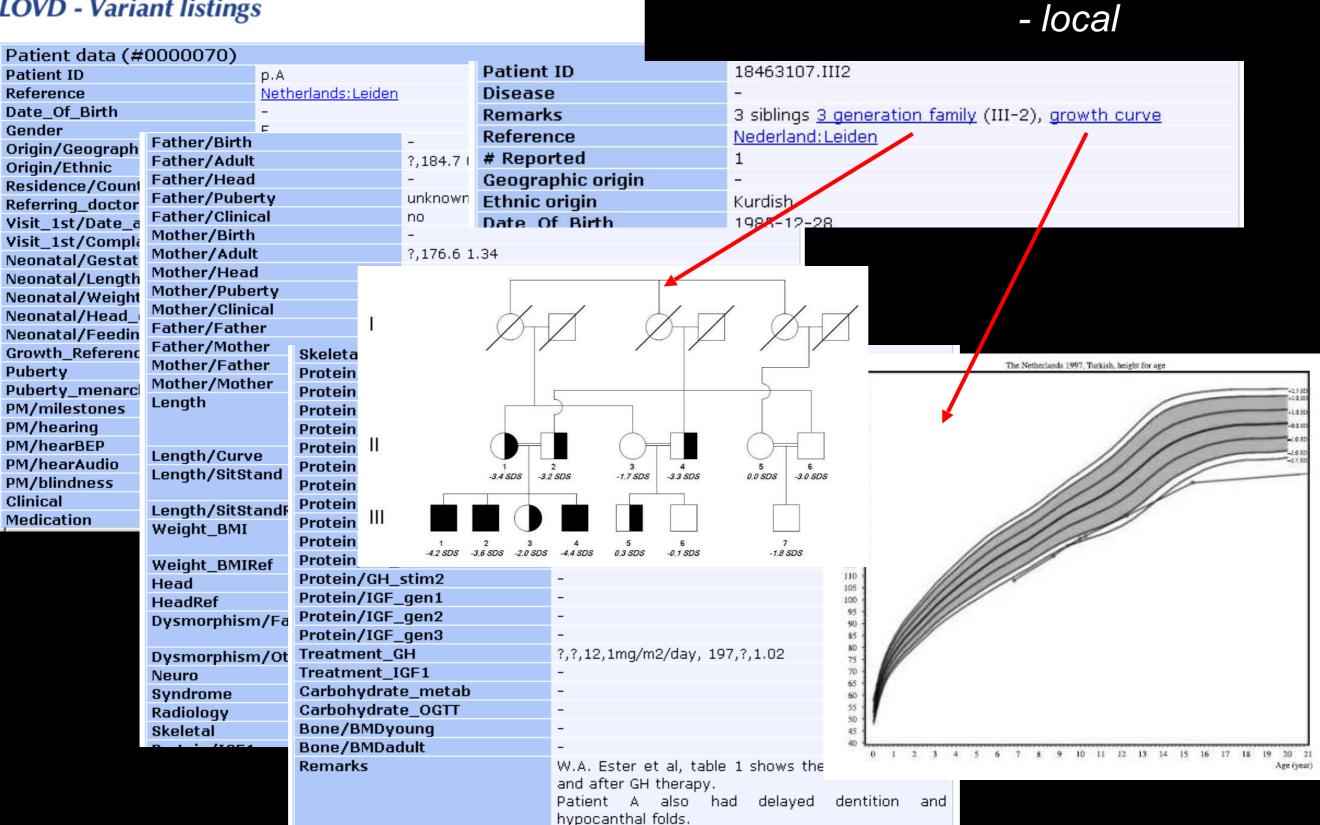
Submit

Configuration

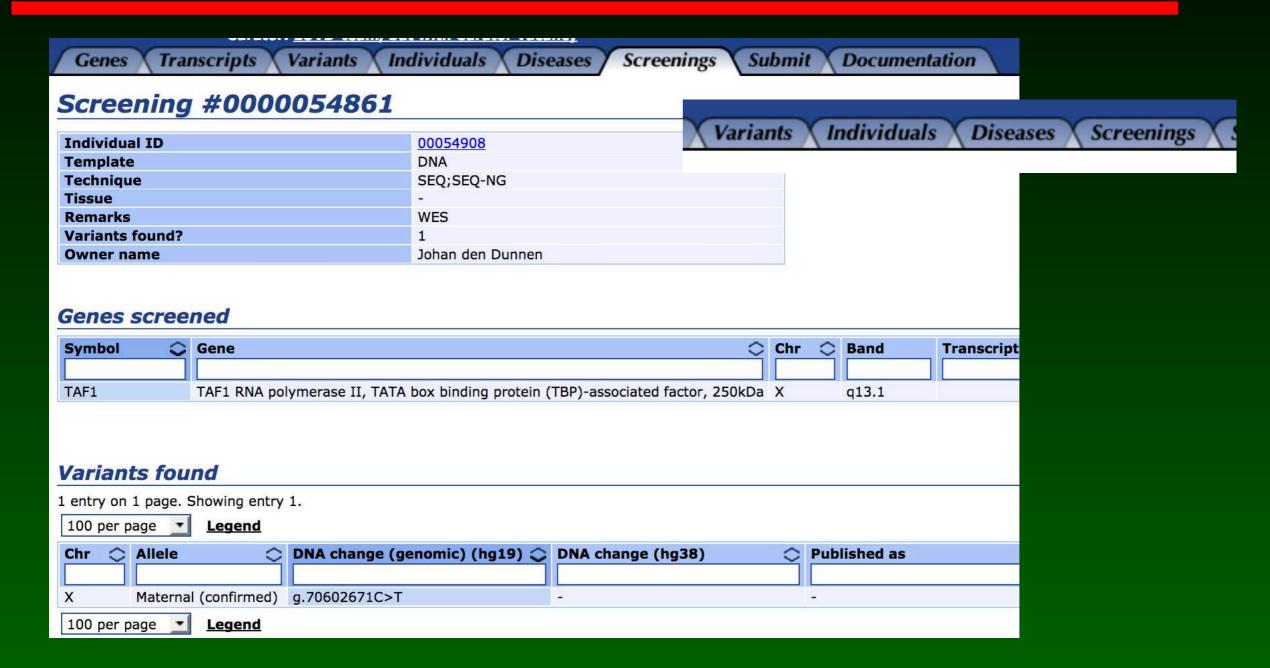
links to files

- internet

LOVD - Variant listings



4 in 1: Screening





4 in 1: Variants

Company of the Compan	
Genes Transcripts Variants Individuals	s X Diseases X Screenings X Submit X Docume
Genomic variant #0000084884	Ī
Genomic Variant #0000084884	•
Individual ID	00054908 Variants Individuals
Chromosome	X
Allele	Maternal (confirmed)
Affects function (as reported)	Affects function
Affects function (by curator)	Not classified
DNA change (genomic) (Relative to hg19 / GRCh37)	g.70602671C>T
DNA change (hg38)	-
Published as	-
ISCN	-
DB-ID	TAF1_000024 See all 4 reported entries
Variant remarks	-
Reference	Journal: O'Rawe 2015
ClinVar ID	-
dbSNP ID	-
Germline/De novo/Somatic	Germline
Segregation	yes
Frequency	-
Re-site	-
VIP	0
Methylation	
Average frequency (large NGS studies)	Variant not found in online data sets
Owner	Johan den Dunnen
Gene	TAF1
Transcript ID	NM_004606.3
Affects function (as reported)	Affects function
Affects function (by curator)	Not classified
Exon	12
DNA change (cDNA)	c.1786C>T
Class.	
RNA change	r.(?)
Protein	p.(Pro596Ser)





Documentation

Diseases

Screenings

1 individual / many variants



DMB-i

Henk Buermans

100 per	page Legend		
Chr 🗢	Allele	DNA change (genomic) (hg19) 🔾	DNA change (hg38)
22	Parent #2	g.42522071C=	g.42126069A>C
22	Parent #2	g.42522071C=	g.42126069A>C
22	Parent #2	g.42522613C>G	g.42126611C>G
22	Parent #2	g.42523003G>A	g.42127001G>A
22	Parent #2	g.42523209C>T	g.42127207C>T
22	Parent #2	g.42523409G=	g.42127407T>G
22	Parent #2	g.42523943A=	g.42127941G>A
22	Parent #2	g.42525132G=	g.42129130C>G
22	Parent #1	g.42525625C>T	g.42129623C>T
22	Parent #2	g.42525772G>A	g.42129770G>A
22	Both (homozygous)	g.42525952C=	g.42129950A>C
22	Parent #2	g.42526484A=	g.42130482C>A
22	Parent #2	g.42526549C=	g.42130547T>C
22	Parent #2	g.42526561_42526562=	g.[42130559T>G;42130560C>G]
22	Parent #2	g.42526567G=	g.42130565A>G
22	Parent #2	g.42526571C=	g.42130569G>C
22	Parent #2	g.42526573T=	g.42130571G>T
22	Parent #2	g.42526580G=	g.42130578C>G
22	Parent #2	g.42527533G=	g.42131531G>A
22	Parent #1	g.42527887delT	g.42131885delT
22	Parent #2	g.42528028C=	g.42132026T>C
22	Parent #1	g.42528030T[21]	g.42132049delT
22	Parent #2	g.42528030T[25]	g.42132047_42132049dup



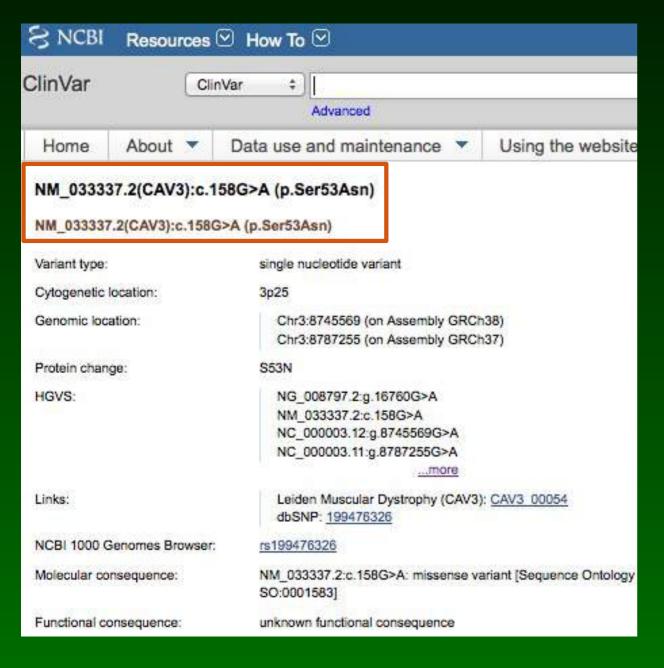
Diseases

Owner name



Databases and RNA

RNA?



HGV Database

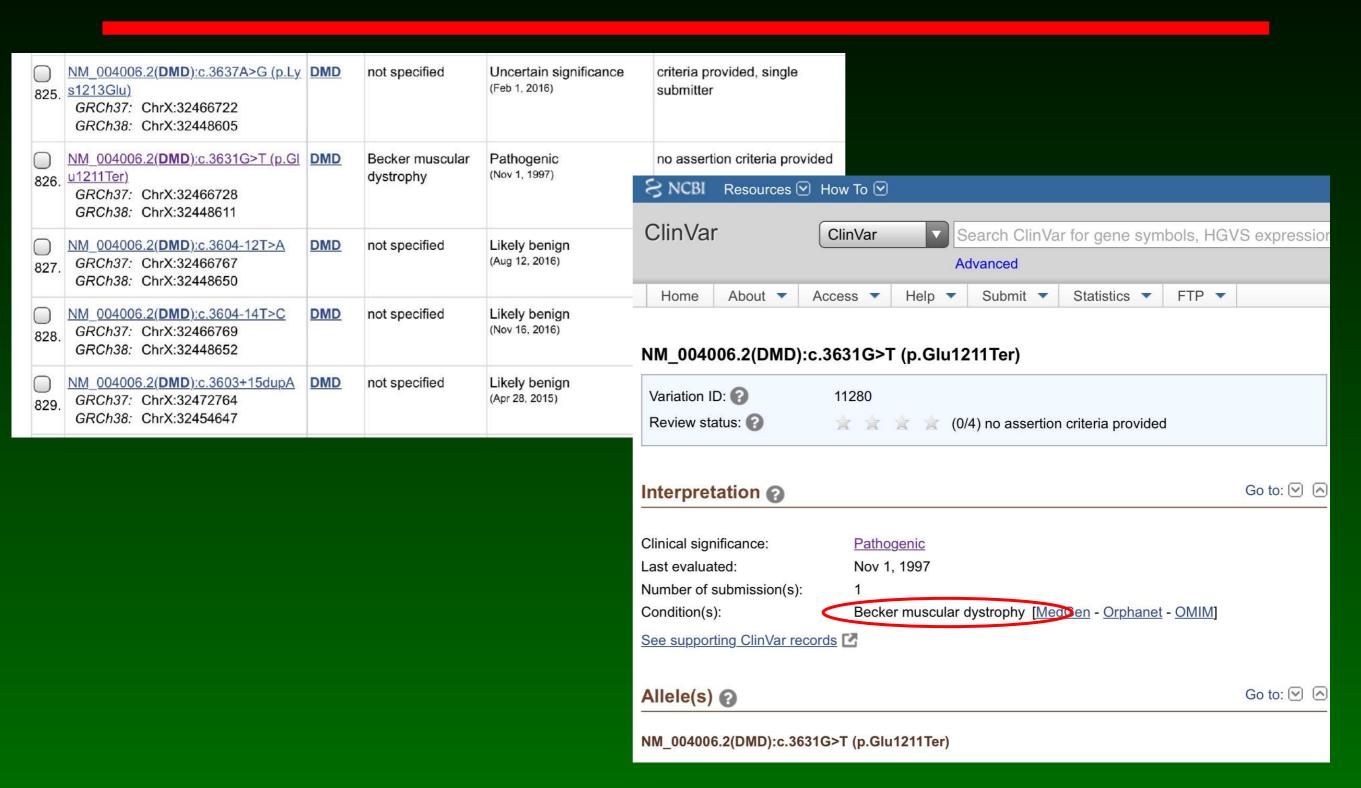
managed by Human Genome Variation powered by figshare

The HGV database is a fully searchable online database of genome variations published in peer-reviewed Data Reports in Human Genome Variation (ISSN 2054-345X).

You can search and filter by different variables, including specific disease,

gene, p	opulation or region.
Gene name	KRT5
Gene/Locus MIM number	148040
Mutation type	missense
Disease / Phenotype	Epidermolysis Bullosa Simplex Koebner ty pe
Phenotype MIM number	131900
GenBank accession number	NM_000424.3
Mutation (HGVS format)	c.T974A
Protein alteration	p.L325H
Codon / Base change	CTC-CAC
Chromosome	12q13.13

ClinVar DMD





LOVD DIVID

Shiga, Takeshima 2010, (OMIM 0074)

DNA, RNA

RT-PCR, SEQ, SSCA

27	c.3628_3665del	-	-			r.3628_3665del		
27	c.3630delA					r.(?)		
27	c.3631G>T	3	3839G>T			r.[3631g>u, 3604_3786del, 3604_40)71del]	
27	c.3679C>T					r(2)		
27	C.3079C21					r.(?)		
27	c.3697delC	Patien	nt data (#0006974)				
27	c.3700G>T	Phenot	type	mı	uscular	dystrophy, Becker (BMD)		
		Phenot	type additional	-				
27	c.3705C>T	Refere		Ja	pan:Kob	<u>e</u>		
	(Reported 2 tilles)	Remark		-				
(a)			phic origin	Ja	pan			
		Ethnic Gender		- M				
		Inherit		M unknown				
		Consanguinity		-	IKIIOWII			
		Fam_P		-				
		# repo		1				0 !
		CK leve		-				825.
		Protein	n data	S.5.				
		Submit	tter	Ma	asafumi	Matsuo		826.
		Varian	nt data					
		Allele		Parent a	#1			827.
			ed pathogenicity	Pathoge				0 1
			Concluded pathogenicity Unknown					828.
		Exon		27				0 1
		DNA ch	1.70	The second second second		ew in <u>UCSC Genome Browser</u> , <u>Ensembl</u>)		829.
		Var_pu		3839G>		04 2796401 2604 40714011		
		RNA ch	range r change			04_3786del, 3604_4071del] rg1202_1262del; Arg1202_1357del]		
		DB-ID	Change	DMD_0		191202_1202uei, Aig1202_133/uei]		
		Variant remarks 10% diff.splice						
		Genet_ori germline (inheri			ited)			
		Segreg		-				
		Defere		Chi 7	T-1	- 2010 (OMIN 0074)		

p.Lys1210*
p.(Glu1211Lysfs*4)
p.[Glu1211*; Arg1202_1262del; Arg1202_1357del]
p.(Gln1227*)
p.(Gln1233Lysfs*4)
p.(Glu1234*)
p.(=)

825.	NM_004006.2(DMD):c.3637A>G (p.Ly s1213Glu) GRCh37: ChrX:32466722 GRCh38: ChrX:32448605	DMD	not specified	Uncertain significance (Feb 1, 2016)	criteria provided, single submitter
826.	NM_004006.2(DMD):c.3631G>T (p.Gl u1211Ter) GRCh37: ChrX:32466728 GRCh38: ChrX:32448611	DMD	Becker muscular dystrophy	Pathogenic (Nov 1, 1997)	no assertion criteria provided
327.	NM_004006.2(DMD):c.3604-12T>A GRCh37: ChrX:32466767 GRCh38: ChrX:32448650	DMD	not specified	Likely benign (Aug 12, 2016)	criteria provided, single submitter
828.	NM_004006.2(DMD):c.3604-14T>C GRCh37: ChrX:32466769 GRCh38: ChrX:32448652	DMD	not specified	Likely benign (Nov 16, 2016)	criteria provided, single submitter
0 829.	NM 004006.2(DMD):c.3603+15dupA GRCh37: ChrX:32472764 GRCh38: ChrX:32454647	DMD	not specified	Likely benign (Apr 28, 2015)	criteria provided, single submitter

Reference

Template Technique

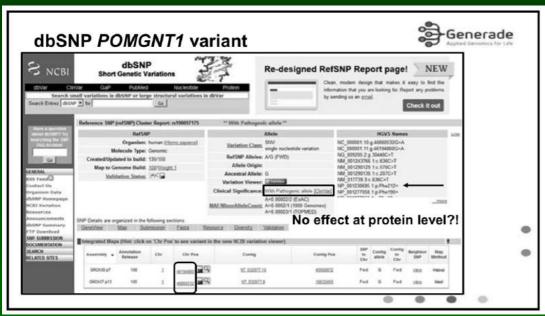
Frequency RE-site



and RNA

POMGNT1 database

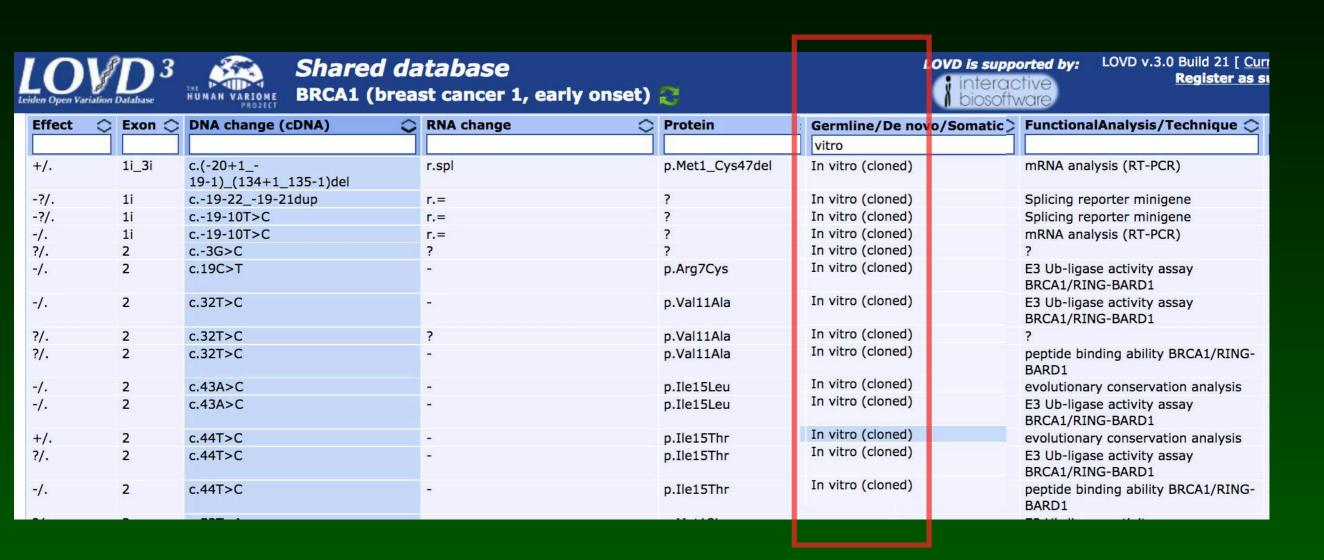
+/.	7	c.594C>G	-	r.(?)	p.(Ser198Arg)	g.46660574G>C
+/.	7	c.630G>T	-	r.(?)	p.(Trp210Cys)	g.46660538C>A
+/+?	7	c.636C>T	=.	r.spl	p.?	g.46660532G>A
+/.	7	c.636C>T	-	r.535_652del	p.Asp179Valfs*23	g.46660532G>A
+/.	7	c.636C>T	-	r.(?)	p.(=)	g.46660532G>A
+/.	7	c.636C>T	-	r.535_652del	p.Asp179Valfs*23	g.46660532G>A
+/.	7	c.636C>T		r.(=)	p.(=)	g.46660532G>A
+/.	7	c.643C>T		r.535_652del	p.Asp179Valfs*23	g.46660525G>A
+/+?	7	c.643C>T		r.535_652del	p.Asp179Valfs*23	g.46660525G>A
+?/+?	7i	c.652+1G>A	-	r.spl	p.?	g.46660515C>T
+/.	7i	c.652+1G>A	->	r.spl?	p.?	g.46660515C>T







In vitro record



BRCA1 variants tested in in vitro functional assays



SUMMARY record

### 11	LO	$ \sqrt[3]{\mathbf{D}^3} $	LOVD 3.0 shared installation			rent LOVD ubmitter				
### 11	Leiden Open Variat	ion Database								
## 11 c.1001C>7 c.(?) p.(Glu337Argfs*9) o encode carrly onset)	+/+	11	c.984_985insC	r.(?)		-				
### 1	+/+	11	c.984_988del	r.(?)	p.(Cys328*)					
#/# 11 c1012A>T c(7) p.(Lys338*) o encode	-/-	11	c.1001C>T	r.(?)	p.(Pro334Leu)	-				
+/+ 11 c.1008dup r.(?) p.(Glu337Argfs*9) o encode protein prot							acron in	eractive		
### 11 c.1016del r.(?) p.(\(\frac{1}{2}\) p.(\(\frac{1}2\) p.(\(\frac{1}{2}\) p.(\(\frac{1}2\) p.(\(\f	+/+	11	c.1008dup	r.(?)	p.(Glu337Argfs*9)	The second secon		-	SUMMARY record	
-/+ 11 c.1016dup r.(?) p.(\(\)\(\)\(\)\(\)\(\)\(\)\(\)\(\)\(\)\(+/+	11	c.1012A>T	r.(?)	p.(Lys338*)		ENICMA classification criteria	_	SLIMMARY record	
+/+ 11 c.1016_1017insC r.(?) p.(Lys339Asnfs*7) terior teri	+/+	11	c.1016del	r.(?)	p.(Lys339Argfs*2)		ENTOPIA Classification criteria	-	SOMMAN TECOID	
+/+ 11 c.1016_1017/insC r.(?) p.(Lys339Asnfs*7) terior torial (ENIGMA:BRCA[1] [c.1001C%3ET]}, PubMed: Lindor class as per Plon 2008 (PMID:18951446). Class 1 based on posterior probability = 0.000000271 Variant allele predicted to encode truncated non-functional protein Variant	+/+	11	c.1016dup	r.(?)	p.(Val340Glyfs*6)		ENIGMA classification criteria	-	SUMMARY record	
holds for class as per Plon 2008 (PMID:18951446). Class 1 based on posterior probability = 0.00000271 Variant allele predicted to encode truncated non-functional protein Variant allele predicted to encode trun	+/+	11	c.1016_1017insC	r.(?)	p.(Lys339Asnfs*7)	1,50	ENIGMA classification criteria,	-	SUMMARY record	
class as per Plon 2008 (PMID:18951446). Class 1 based on posterior probability = 0.000000271 Variant allele predicted to encode truncated non-functional protein	+/+	11	c.1017_1018insA	r.(?)	p.(Val340Serfs*6)	-				
truncated non-functional protein Variant allele predicted to encode ENIGMA classification criteria SUMMARY record					(PMID:18951446 on posterior prol	2008 5). Class 1 based				
Variant allele predicted to encode truncated non-functional protein							ENIGMA classification criteria	3	SUMMARY record	
truncated non-functional protein Variant allele predicted to encode truncated non-functional protein Variant allele predicted to encode Variant allele predicted to encode ENIGMA classification criteria SUMMARY record SUMMARY record					Variant allele pre	edicted to encode	ENIGMA classification criteria	-	SUMMARY record	
truncated non-functional protein Variant allele predicted to encode					Variant allele pre	edicted to encode	ENIGMA classification criteria	ä	SUMMARY record	
***************************************							ENIGMA classification criteria	-	SUMMARY record	
							ENIGMA classification criteria	-	SUMMARY record	

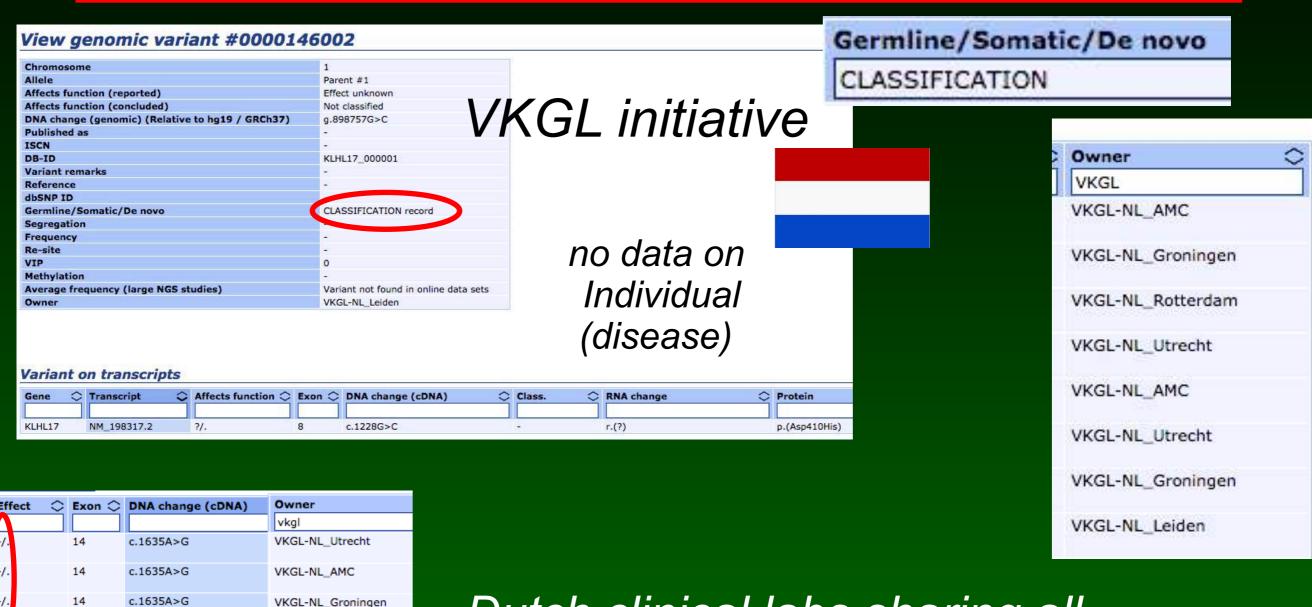
ENIGMA variant interpretation committee

(BRCA associated or not, 5 classes)





CLASSIFICATION record

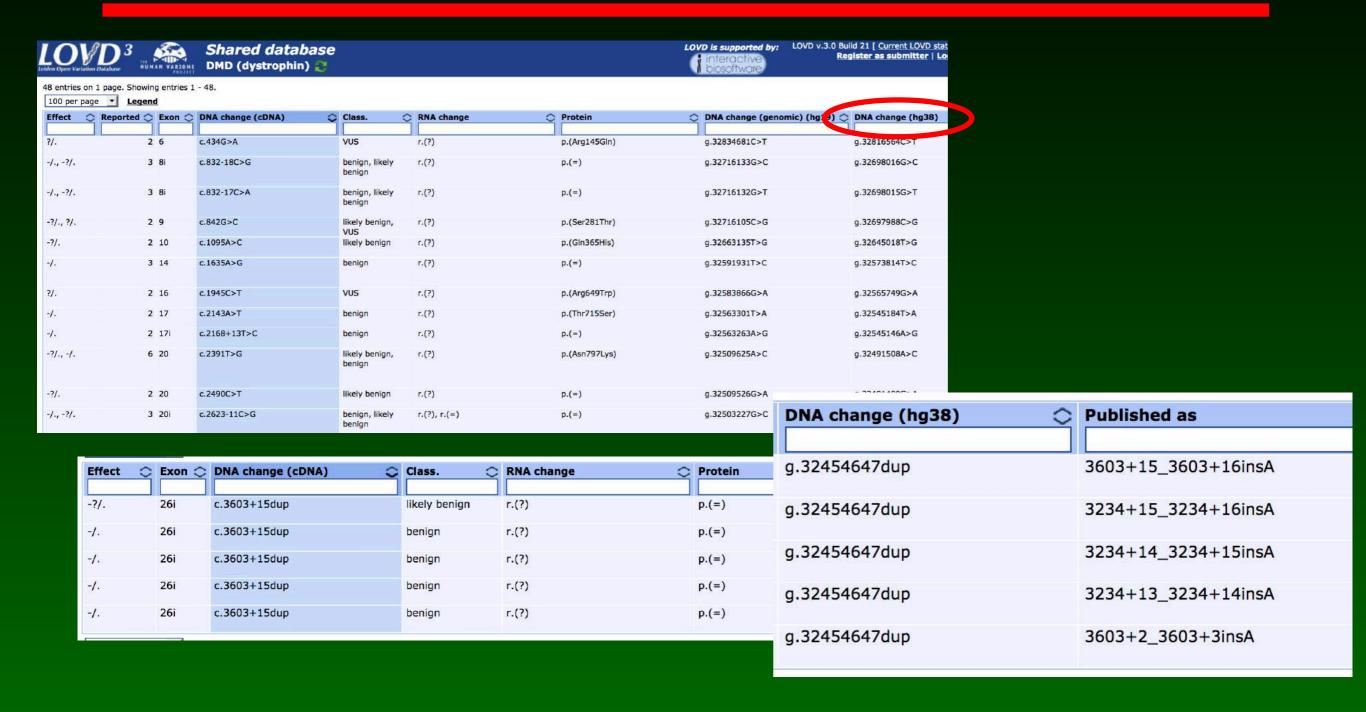


Dutch clinical labs sharing all variants and their classification





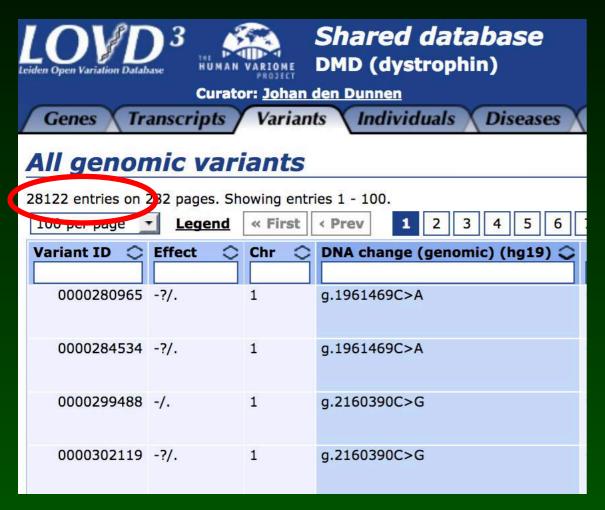
VKGL initiative







VKGL initiative



Shared database **DMD** (dystrophin) Individuals Diseases Genes **Transcripts Variants** All genomic variants 105780 entries or 1058 pages. Showing entries 1 - 100. 100 per page ▼ « First | < Prev Legend Variant ID 🔷 Effect 🔷 Chr 🔷 DNA change (genomic) (hg19) 0000245460 -/. g.69270A>G 0000343193 ?/. g.879375C>T 0000304227 -/. g.881627G>A 0000320401 ?/. g.898757G>C

public: 28,122 total: 105,780

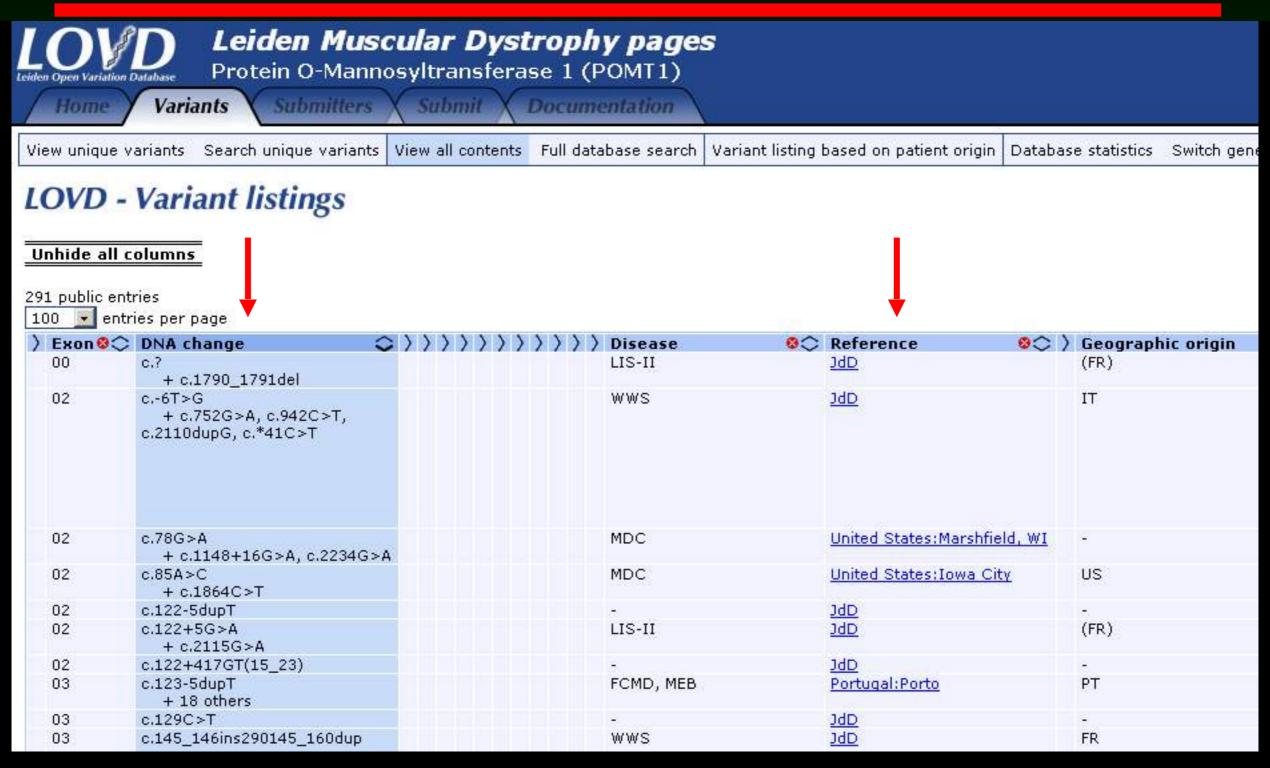
Dutch clinical labs sharing all variants and their consensus classification

1139 variants no consensus 29 variants opposite classification





Minimal reporting



variant + contact details

How improve sharing?

• ...make variant submission obligatory

for publication for approved grants

for clinical lab as part of QC to get accreditation

...to get a sample sequenced

+ QC standards for LSDBs

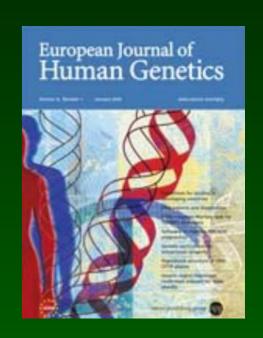


Journals

luman Mutation



- Human Mutation first to demand using HGVS nomenclature first to demand database submission before accepting a paper for publication
- European Journal of Human Genetics demanding HGVS nomenclature & DB submission ...and checking every paper !!







Adopt a gene!

CV

Curator for the ... gene variant database.





many orphan genes





become a foster parent database curator



