



# dbVar: A Genomic Structural Variation Database

A collection of human genomic structural variation

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

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

## Scope and Access

The NCBI dbVar Structural Variation database houses human genomic structural variants (SV) greater than 50 base pairs in length. From the dbVar homepage (A, [www.ncbi.nlm.nih.gov/dbvar/](http://www.ncbi.nlm.nih.gov/dbvar/)), you can search (B), browse (C), view and download variant data from over 150 studies, such as 1000 Genomes Phase 3 ([estd219](#)), Genome in a Bottle ([nstd175](#)), Clinical Structural Variants ([nstd102](#)), gnomAD ([nstd166](#)), etc. You can access the variants using the Study Browser or the graphical Genome Browser. Individual Study and Variant Pages include links to the raw data as well as to related information at other NCBI and external resources. Bulk data downloads are available by FTP at <ftp.ncbi.nlm.nih.gov/pub/dbVar/data> (D).



In 2018, dbVar introduced a new comprehensive set of non-redundant structural variants (NR set) consisting of unique insertions, duplications, and deletions. These compact files are suitable for use as references in the analysis of human structural variation. For example, you can use the dbVar NR set to filter and annotate other datasets in a broad range of applications including variant discovery and identifying rare and/or clinical variants. The dbVar NR set currently includes more than 2.5 million deletions, 1.3 million insertions, and 400 thousand duplications. The NR set is updated monthly as new variants are added to dbVar. You can find more information about the NR set, brief tutorials, and ways to access NR set FTP files from github at [github.com/ncbi/dbvar/tree/master/Structural\\_Variant\\_Sets](https://github.com/ncbi/dbvar/tree/master/Structural_Variant_Sets) (E).

## Searching in dbVar

To search dbVar, type terms in the search box and click the "Search" button (F). For more refined results, use field-limited terms connected with Boolean operators. For example, querying with [17\[chromosome\] AND deletion\[phenotype\]](#), retrieves all deletion structural variants on human chromosome 17 (G).

The search result page displays the variant ID, the type of variation and other key features in a table. Click the coordinates on different genome assemblies in the Location column (H) to open a graphical presentation of the variants in the dbVar **Genome Browser**. Click the "... more genes" link in the "Genes in region" column (I) to change display to the Summary format so records for affected genes are directly accessible. A set of filters (J) in the left-hand column allows quick refining of the variant list according to your interest.

The screenshot shows the dbVar homepage. Callout A points to the 'dbVar' logo. Callout B points to the search box. Callout C points to the 'Study Browser' link. Callout D points to the 'FTP Data Download' link. Callout E points to the 'Clinical Structural Variation Walkthrough' section. Callout F points to the 'Search' button.

The screenshot shows the search results page. Callout G points to the search URL <http://bit.ly/2cSSByR>. Callout H points to the 'Location' column in the table. Callout I points to the 'Genes in region' column in the table.

Variant Region ID	Type	Number of Variant Calls	Study ID	Organism	Clinical Assertion	Location	Genes in region
4349853	copy number variation	1	nstd102	human	Pathogenic	GRCh37.(hg19) chr17: 34,819,191-36,194,230; GRCh38.p12 chr17: 36,463,349-37,834,613; GRCh38.p12 chr17:INT.187614.1: 698,253-2,973,295; NCBI36.(hg18) chr17: 31,893,304-33,268,343	TADA2A, ZNHIT3, 35 more genes
nsV3971867	copy number variation	2	nstd102	human	Pathogenic	GRCh38.(hg38) chr17: 37,710,654-37,710,654; GRCh37.p13 chr17: 36,070,662-36,070,662; NCBI36.(hg18) chr17: 33,144,775-33,144,775	HNF1B, LOC105371754
nsV3877390	copy number variation	1	nstd102	human	Pathogenic	GRCh37.(hg19) chr17: 34,815,072-36,192,492; GRCh38.p12 chr17: 36,459,259-37,832,872; GRCh38.p12 chr17:INT.187614.1: 694,163-2,971,557; NCBI36.(hg18) chr17: 31,889,185-33,268,305	LHX1, HNF1B, 35 more genes

## The Advanced Search Builder

The "Advanced" link (A) under the dbVar search box links to the Advanced Search builder, which helps you customize and refine your searches through combining search terms and setting field limit.

In the Advanced Search Builder page, you can access indexing fields (B) and terms indexed in each field (C) through the "Show index list" link (D), as well as your search history (E). The query builder function allows you to combine indexing terms and entries in the current search history with Boolean operators (F) for more specific retrieval. Click the "Edit" link (G) to manually edit search terms.

- Chromosome Outer Start
- Chromosome Start
- ClinVar Accession
- Detection Method
- Filter**
- Gender
- Gene Full Name
- Gene Name
- Library Abbreviation
- MIM ID
- MeSH Terms
- MeSH Unique ID
- Method Platform

dbVar Advanced Search Builder

"clin pathogenic"[Filter] <https://www.ncbi.nlm.nih.gov/dbvar/advanced> [Clear](#)

[Edit](#) [Search](#) or [Add to history](#)

Builder

Filter  [Hide index list](#)

- clin none (4470200)
- clin not provided (2296)
- clin pathogenic (5638)**
- clin risk factor (2)
- clin uncertain significance (4795)
- dbvar all (4481898)
- dbvar bioproject (617317)

[Previous 200](#) [Next 200](#) [Refresh index](#) [Show index list](#)

or [Add to history](#)

History [Download history](#) [Clear history](#)

Search	Add to builder	Query	Items found	Time
#23	<a href="#">Add</a>	Search "diabetes"[Phenotype] AND #21	9	16:16:49
#22	<a href="#">Add</a>	Search "diabetes"[Phenotype]	192	16:15:27
#21	<a href="#">Add</a>	Search "clin pathogenic"[Filter]	5638	16:14:02
#17	<a href="#">Add</a>	Search human[orgn] AND 17[chromosome] AND deletion[phenotype]	4	16:02:16

## dbVar Study Browser

The dbVar homepage provides links to the Study Browser page, where you can browse by study IDs. The Study Browser table (below) sums up the available studies in a table. Clicking a column header, such as the "Variant Region Count" (G), sorts the list according to the value in that column so you can locate studies with certain characteristics. Identifiers for individual studies in the Study column (H) links to detailed display of specific studies, while the citations in the Publication column (I) link to relevant articles in PubMed. You can use criteria filters in the right-hand column (J) to narrow the list of studies displayed.

dbVar: Study Browser

Date	Publication	Study	Organism	Variant Region Count	Variant Call Count
2018/09	<a href="#">Szafranski et al. 2018</a>	<a href="#">nstd153</a>	Human	8	8
2018/04	<a href="#">Kucukkilic et al. 2018</a>	<a href="#">nstd159</a>	Human	2	350
2018/02	<a href="#">Fu et al. 2018</a>	<a href="#">nstd156</a>	Human	963	11,796
2018/02	<a href="#">Adewoye et al. 2018</a>	<a href="#">nstd155</a>	Human	1	2,502
2018/01	<a href="#">Rambo-Martin et al. 2018</a>	<a href="#">nstd141</a>	Human	123	274
2017/12	<a href="#">Möller et al. 2017</a>	<a href="#">nstd154</a>	Human	1	7
2017/09	<a href="#">Gambin et al. 2017</a>	<a href="#">nstd149</a>	Human	36	41
2017/08	<a href="#">Gardner et al. 2017</a>	<a href="#">nstd144</a>	Chimpanzee, Human	37,798	38,422
2017/07	<a href="#">Luo et al. 2017b</a>	<a href="#">estd233</a>	Human	1,026	1,026
2017/07	<a href="#">Lu et al. 2017</a>	<a href="#">nstd145</a>	Human	8,237	26,917
2017/03	dbSNP <a href="#">rated variants</a>	<a href="#">nstd90</a>	Human	4,715	4,722
2017/02	<a href="#">Blanco-Corral et al. 2017</a>	<a href="#">estd232</a>	Human	8	8
2017/02	<a href="#">Shang et al. 2017</a>	<a href="#">nstd143</a>	Human	18	18
2017/02	<a href="#">Rahbari et al. 2017</a>	<a href="#">nstd142</a>	Human	3	7
2017/02	<a href="#">Walker et al. 2017</a>	<a href="#">nstd132</a>	Human	6,173	15,888
2017/01	<a href="#">Fan et al. 2017</a>	<a href="#">nstd140</a>	Human	22,644	22,644

Filter by Study Type

- Control Set (87)
- Case-Set (35)
- Case-Control (17)
- Collection (11)
- Curated Collection (5)
- [See more...](#)
- Source: NCBI

Filter by Method

- Sequencing (43)
- SNP array (31)
- Oligo aCGH (17)
- BAC aCGH (6)
- Curated (5)
- [See more...](#)
- Source: NCBI

Filter by Variant

- >=10



### dbVar Genome Browser

**Ideogram View**

Unplaced/unlocalized scaffolds: 168  
Alt loci/patches: 401

**Pick Assembly**

GCF\_000001405.38 (GRCh38.p12)

**Search**

Search: TERT

Name	Location
TERT	Chr5: 1,295,047 - 1,295,047
LOC110806263	Chr5: 1,298,988 - 1,298,988
MCPH1	Chr8: 6,406,592 - 6,648,505
MCPH1	NW_018654717.1: 8.4K - 438
TP53	Chr17: 7,668,402 - 7,687,550
CLPTM1L	Chr5: 1,317,744 - 1,345,070

**User Data and Track Hubs**

**Region Content**

Data in view

Study ID	Variant Calls
estd3	1
estd20	7
estd22	2
estd59	1
estd180	2

Data available for region

Study ID	Variant Calls
estd3	1
estd20	7
estd22	2
estd59	1
estd180	2

## dbVar Genome Browser

The Genome Browser (A) provides a way to locate reported variants for genomic regions of interest. You can specify a genomic region by using either coordinates or annotated features known to occur in these regions, such as gene names (B). Click a retrieved entry from the list to zoom to the feature (C), click the (+) sign (D) to the right of studies to add tracks for mapped variants. Hover a variant to see its detail in a popup (E).

Homo sapiens: GRCh38.p12 (GCF\_000001405.38) Chr 5 (NC\_000005.10): 1,248,979 - 1,299,235

Reset All Share this page Help Version 2.8 Setup Page

NC\_000005.10

Tools Tracks

Genes, NCBI Homo sapiens Annotation Release 109...

esv26971

Variation ID: dbVar: esv26971  
Supporting Variant Calls: 1  
Variant Region Type: Unknown  
Validation Status: Not Tested  
Total Length: 1,326

Links & Tools  
BLAST Genome-specific: NC\_000005.10 (1,272,523..1,273,848)  
BLAST Genomic: NC\_000005.10 (1,272,523..1,273,848)  
FASTA View: NC\_000005.10 (1,272,523..1,273,848)  
GenBank View: NC\_000005.10 (1,272,523..1,273,848)  
dbVar: esv26971

## dbVar Study Pages

The details of a specific study are shown in the Study Page (F). The default "Variant Summary" tab (G) lists variants mapped to different chromosomes. Variants identified by the study are available through the "Variant in this study" link (H).

**Variant Summary**

Number of Samplesets: 12

Sampleset ID: 1  
Sampleset Type: Case  
Description: Children referred to Signature Genomics with a general diagnosis of intellectual disability (ID) and/or developmental delay (DD).  
Size: 15,767  
Organisms: Homo sapiens  
Sampleset Phenotype(s): Developmental Disabilities

Samples for sampleset 1 (displaying 100 of the 15767 samples)

Sample ID	Subject ID	Subject Phenotype
9881096	9881096	Developmental Disabilities Global developmental delay
9881101	9881101	Autism Developmental Disabilities Global developmental delay

**nstd54 (Cooper et al. 2011)**

Organism: Human  
Study Type: Case-Control  
Submitter: Evan Eichler  
Description: Copy Number Variants from 15,767 cases of Developmental Delay and Intellectual Disability from Signature Genomics, and 8329 Control Samples. This study contains samples in common with [Coe et al. 2014](#). Due to analysis differences (see manuscripts) please use the case samples (Sampleset 1) from only one of these submissions. Control sample sets do not overlap and may be combined.  
Publication(s): Cooper et al. 2011

Detailed Information: [Download 81345 Variant Regions](#), [Download 468909 Variant Calls](#), [Download Both](#), [FTP](#)

Variant Summary Samplesets Experimental Details Validations

Assembly used for analysis:  
Remapped: GRCh38.p12 (hg38)  
Remapped: GRCh37.p13 (hg19)  
Submitted: NCBI36 (hg18)

Variant Summary for: GRCh38.p12 (hg38)

Sequence ID	Chr	Number of Variant Regions	Number of Variant Calls	Placement type	Link to graphical display
NC_000001.11	Chr1	5,660	34,564	Remapped	<a href="#">NC_000001.11</a>
NC_000002.12	Chr2	5,525	35,212	Remapped	<a href="#">NC_000002.12</a>

The "Samplesets" tab (I) displays samples used in the study. Information provided includes the IDs of the participating subject and their demonstrated phenotypes (J). The "Experimental Details" tab (K, details not shown) provides methodologies and experimental platforms used when such information is made available by the submitters.

## dbVar Variant Record View

In the tabular display of a search result (A), clicking the variant ID (B) opens a detailed variant view display of that record.

The top section of this variant view (C), from left to right, provides a summary of the variant, overview of its genomic placement, and links to relevant records in other NCBI databases. The variant's chromosomal mapping is indicated by green arrow.

### nsv4436665

Organism: [Homo sapiens](#)

Study: [nstd102 \(Clinical Structural Variants\)](#)

Variant Calls: 1

Variant Type: copy number variation

Validation: Not tested

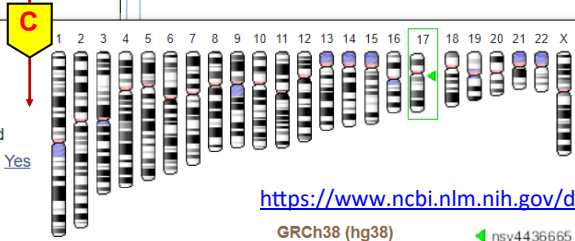
Method Type: Multiple

Clinical Assertions: [Yes](#)

Submitted on: GRCh38 (hg38)

Region Size: 1

Description: NM\_000267.3:c.2385delA AND Neurofibroma



#### Links to Other Resources

ClinVar: [RCV000787329.1](#)

Overlapping Genes

Source: NCBI

<https://www.ncbi.nlm.nih.gov/dbvar/variants/nsv4436665/>

GRCh38 (hg38)

nsv4436665

#### Search results

Items: 1 to 200 of 236

<< First < Prev Page 1 of 2 Next > Last >>

Number of Variants: 200

<https://go.usa.gov/xV6sH>

Variant Region ID	Type	Number of Variant Calls	Study ID	Organism	Clinical Assertion	Location	Genes in region
nsv4436665	copy number variation	1	nstd102	human	Pathogenic	GRCh38 (hg38) chr17: 31,227,585-31,227,585 ; GRCh37.p13 chr17: 29,554,603-29,554,603 ; NCBI36 (hg18) chr17: 26,578,729-26,578,729	NF1

Genome View **D** Variant Region Details and Evidence Validation Information Clinical Assertions Genotype Information

Genome View

Select assembly: GRCh38 (hg38): Chr17

Overlapping variant regions from other studies: 169 SVs from 29 studies. See in: [genome view](#)

Submitted genomic 31,227,585 - 31,227,585

SV-Anon4 - Segment Map

Genes

dbVar Clinical Structural Variants (nstd102)

nsv3909684 [+1]

nsv3905245 [+1]

nsv972367 [+1]

nsv4351367 [+1]

nsv4354864 [+1]

nsv4349684 [+1]

nsv3907838 [+1]

nsv3922941 [+1]

nsv3906384 [+1]

nsv3906384 [+1]

nsv3913906 [+1]

nsv3888575 [+1]

Links & Tools

dbVar: [nsv3888575](#)

Variation ID: dbVar: nsv3888575

Supporting Variant Calls: 1

Variant Region Type: Copy number variation

Validation Status: Not Tested

Total Length: 1

Tabs underneath the summary provide other relevant details. Specifically, the "Genome View" (D) presents the variant graphically in its genomic context. Hovering the mouse pointer over a variant brings out its summary in a popup (E). The

Genome View Variant Region Details and Evidence **F** Validation Information Clinical Assertions Genotype Information

Variant Region Placement Information

Variant Region ID	Placement Type	Score	Assembly	Assembly Unit	Reciprocity	Sequence ID	Chr	Start	Stop
nsv4436665	Submitted genomic		GRCh38 (hg38)	Primary Assembly		NC_000017.11	Chr17	31,227,585	31,227,585
nsv4436665	Remapped	Perfect	GRCh37.p13	Primary Assembly	First Pass	NC_000017.10	Chr17	29,554,603	29,554,603

Variant Call Information

Variant Call ID	Type	Method	Analysis	Subject Phenotype	Clinical Interpretation	Source of Interpretation	ClinVar ID
nssv15755125	deletion	Multiple	Multiple	See cases	Pathogenic	ClinVar	<a href="#">RCV000787329.1</a>

Variant Call Placement Information

Variant Call ID	Placement Type	Score	HGVS	Assembly	Reciprocity	Sequence ID	Chr	Start	Stop
nssv15755125	Submitted genomic								
nssv15755125	Remapped								

"Variant Region Details and Evidence" (F) provides detailed genomic coordinates on different genome builds. The "Clinical Assertions" tab (G) provides information on the phenotype and clinical interpretation when data is available.

## References

- Lappalainen I, et. al. (2013) dbVar and DGVA: public archives for genomic structural variation. *Nucleic Acids Res.* 41 (D1):D936-41. [PMID: 23193291](#).
- Sneddon TP and Church DM. (2012) Online resources for genomic structural variation. *Methods Mol Biol.* 838:273-89. [PMID: 22228017](#).
- dbVar help documentation: [www.ncbi.nlm.nih.gov/dbvar/content/help/](http://www.ncbi.nlm.nih.gov/dbvar/content/help/)