

# User Guide

## Database of Nonstop Mutations in Cancer

(<http://NonStopDb.dkfz.de>)

The database of Nonstop Mutations in Cancer (NonStopDb) is a curated database of nonstop extension and readthrough mutations in human cancer. NonStopDb allows biologists to easily extract and download nonstop mutations in cancer using multiple search options. Search criteria available in NonStopDb include the amino acid extension length, recurrence of mutation, site of mutation and cancer type. We describe each search facet below.

### Search by Gene

This feature allows the user to search for nonstop mutations present in a gene of interest (Figure 1) using one of the following nomenclature:

1. Gene Name
2. Ensembl IDs.

Synonymous or alias gene names (e.g. p53 for TP53) are allowed and the search is case-insensitive.

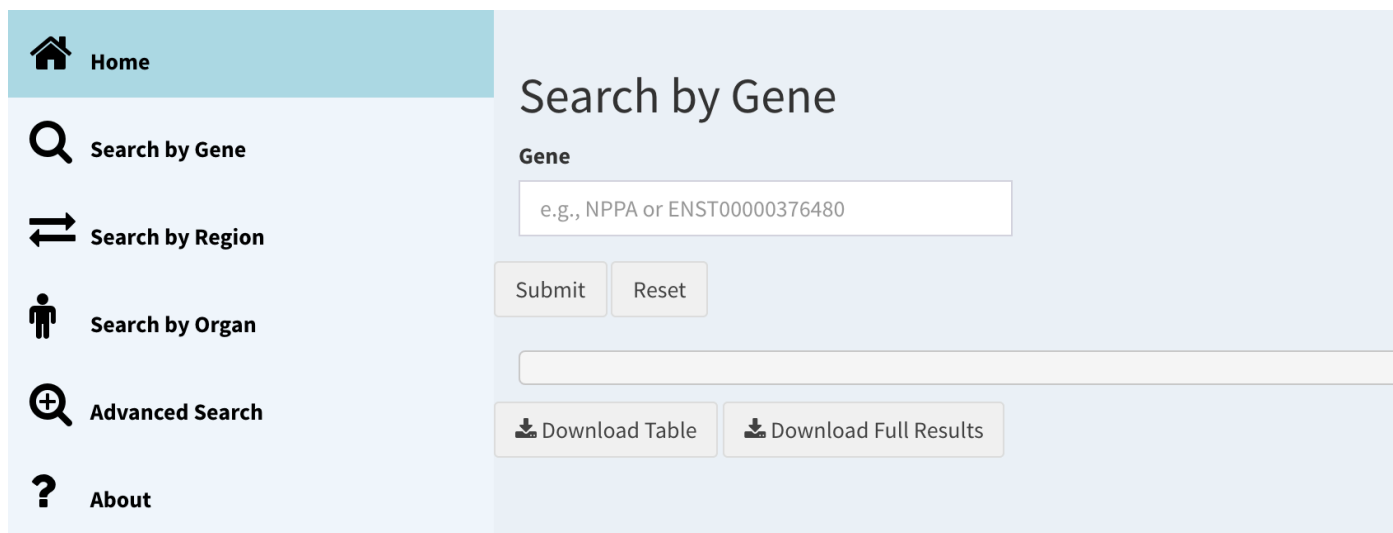


Figure 1: Search by Gene.

For example, Figure 2 shows the results page for the gene *SMAD4*. The summary information in Cancer Gene Census<sup>1</sup> for the gene is shown. The link to Genecards<sup>2</sup> for the gene is also shown.

The user can download the results using one of the two options:

- **Download Table:** This button allows the user to download the displayed results as a csv file.
- **Download Full Results:** This button allows the user to download results for all mutations along with detailed sample information plus additional information not displayed in the NonStopDb table including e.g. FATHMM pathogenicity scores and the mutation burden of the respective tumor sample (the total number of mutations annotated in COSMIC for the tumor if whole genome sequenced).

[Home](#)
[Search by Gene](#)
[Search by Region](#)
[Search by Organ](#)
[Advanced Search](#)
[About](#)

### Search by Gene

Gene

Submit Reset

Gene: SMAD4

[Link to Genecard](#)

Cancer Gene Census summary

Gene Symbol: SMAD4  
Name: SMAD family member 4  
Entrez Gene ID: 4089  
Genome Location: 18:51047047-51078467  
Chr. Band: 18q21.1  
Synonyms: Q13405, JIP, DPC4, MADH4, ENSG00000141646, 4089, SMAD4

Show 10 entries														Search:	
Gene name	Mutation CDS	Mutation AA	Mutation ID	Mutation ID Count	Mutation Class	Extension Length (AA)	Sample name	Organ System	Site	Histology	Mutation genome position	SNP	CGC		
SMAD4	c.1657T>C	p.*553R	COSM1168854	1	Extension	40	8031704	Digestive System	Pancreas	Ductal Adenocarcinoma	chr18:51078465-51078465	no	yes		
SMAD4	c.1658G>C	p.*553S	COSM4388062	1	Extension	40	8068044	Digestive System	Pancreas	Ductal Adenocarcinoma	chr18:51078466-51078466	no	yes		
SMAD4	c.1657T>G	p.*553G	COSM4389249	1	Extension	40	8067507	Digestive System	Pancreas	Ductal Adenocarcinoma	chr18:51078465-51078465	no	yes		
SMAD4	c.1659A>G	p.*553W	COSM308304	2	Extension	40	T_CCA_TH_1	Digestive System	Bile Duct	Carcinoma	chr18:51078467-51078467	no	yes		
SMAD4	c.1659A>G	p.*553W	COSM308304	2	Extension	40	A035	Digestive System	Bile Duct	Carcinoma	chr18:51078467-51078467	no	yes		
SMAD4	c.1659A>C	p.*553C	COSM4728234	2	Extension	40	T3202	Digestive System	Large Intestine	Adenocarcinoma	chr18:51078467-51078467	no	yes		
SMAD4	c.1659A>C	p.*553C	COSM4728234	2	Extension	40	T3202	Digestive System	Large Intestine	Adenocarcinoma	chr18:51078467-51078467	no	yes		
SMAD4	c.1658G>T	p.*553L	COSM1150713	4	Extension	40	T2929	Digestive System	Large Intestine	Adenocarcinoma	chr18:51078466-51078466	no	yes		
SMAD4	c.1658G>T	p.*553L	COSM1150713	4	Extension	40	T2929	Digestive System	Large Intestine	Adenocarcinoma	chr18:51078466-51078466	no	yes		
SMAD4	c.1658G>T	p.*553L	COSM1150713	4	Extension	40	126	Digestive System	Large Intestine	Adenocarcinoma	chr18:51078466-51078466	no	yes		

Showing 1 to 10 of 11 entries

Download Table Download Full Results

Previous 1 2 Next

Figure 2: Results for search by gene in NonStopDb.

The result columns provide following information:

- **Gene name:** Name of the gene.
- **Mutation CDS:** Mutation position in the coding sequence and nucleotide change (For e.g., c.1659A>G indicates a change in the coding sequence at nucleotide 1659 mutated from A to G).
- **Mutation AA:** Mutation position listed as amino acid from stop codon (\*) to an amino acid in the one-letter-code.
- **Mutation ID:** Unique identifier of each mutation (as present in COSMIC database).

<sup>1</sup> <http://cancer.sanger.ac.uk/census/>

<sup>2</sup> <http://www.genecards.org/>

- **Mutation ID Count:** This column shows the recurrence level of each mutation. The number in this column represents the number of times that the respective mutation was found in different tumor samples.
- **Mutation Class:** This column shows recurrence level of each mutation. The number in this column represents number of times that the respective mutation was found in different tumor samples.
- **Extension length (AA):** Number of amino acids appended due to a nonstop mutation including the amino acid encoded by the mutated stop codon.
- **Sample name:** Name of the tumor sample.
- **Organ System:** Organ system for the sample.
- **Site:** Organ or tissue of the sample with the respective mutation.
- **Histology:** Histology of the sample.
- **Mutation genome position:** Genomic coordinates of each respective mutation in human genome assembly GRCh38.
- **SNP COSMIC:** Information whether the mutation is also listed as SNP in the COSMIC database (yes/no).
- **CGC:** Information whether the mutated gene is also listed in the Cancer Gene Census (CGC) (yes/no).

Detailed information for each sample can be viewed by clicking on the ⊕ icon.

Figure 3 shows an example of sample information for mutation ID COSM308304.

- **Chrom:** Chromosome information.
- **Start:** Start position of the mutation.
- **End:** End position of the mutation.
- **Strand:** Strand information.
- **Transcript ID Update:** Update of the transcript ID.
- **Transcript ID updated – Mutation CDS:** Mutation position in the updated Transcript ID coding sequence and nucleotide change.
- **EXT NT Sequence incl. mut STOP:** Nucleotide sequence of the extension including the mutated position in the STOP codon (in the example below, the STOP codon TAG changed to TGG) as well as the second STOP codon (TGA in the example).
- **EXT AA Sequence incl. mut STOP:** Amino acids sequence of the extension including the mutated STOP codon (in the example below, the mutated STOP codon produced a tryptophan – first “W” in the sequence).

Gene: SMAD4

[Link to Genecard](#)

Cancer Gene Census summary

Gene.Symbol: SMAD4

Name: SMAD family member 4

Entrez.GeneId: 4089

Genome.Location: 18:51047047-51078467

Chr.Band: 18q21.1

Synonyms: Q13485, JIP, DPC4, MADH4, ENSG00000141646, 4089, SMAD4

Show 10 entries

Search:

Gene name	Mutation CDS	Mutation AA	Mutation ID	Mutation ID Count	Mutation Class	Extension Length (AA)	Sample name	Organ System	Site	Histology	Mutation genome position	SNP	CGC
SMAD4	c.1657T>C	p.*553R	COSM1168854	1	Extension	40	8031704	Digestive System	Pancreas	Ductal Adenocarcinoma	chr18:51078465-51078465	no	yes

Chr

chr18

Start

51078465

End

51078465

Strand

+

Transcript ID Update

ENST00000342988.7

Transcript ID updated - Mutation CDS

c.1657T>C

EXT NT Sequence incl. mut STOP

CGA GGT CTT TTA CCG TTG GGG CCC TTA ACC TTA TCA GGA TGG TGG ACT ACA AAA TAC AAT CCT GTT TAT AAT CTG AAG ATA TAT TTC ACT TTT GTT CTG CTT TAT CTT TTC ATA AAG GGT TGA

EXT AA Sequence incl. mut STOP

RGLLPLGLTLGGWTTKYNPVYNLKIFYTFVLYLFIKG\*

Figure 3: Detailed sample information for each sample having the mutation ID COSM308304 in the table of results.

# Search by Region

This option allows the user to search for mutations present within the given genomic coordinates of human genome assembly GRCh38 / hg38. For example, Figure 4 shows the list of mutations present in chromosome 2 region 2000000-12000000.

Home

Search by Gene

Search by Region

Search by Organ

Advanced Search

About

Search by Region

Chromosome

chr2

Region (Genome assembly version: GRCh38)

2000000-12000000

Submit

Reset

Show 10 entries

Search:

Gene name	Mutation CDS	Mutation AA	Mutation ID	Mutation ID Count	Mutation Class	Extension Length (AA)	Sample name	Organ System	Site	Histology	Mutation genome position	SNP	CGC
ADAM17	c.2474A>T	p.*825L	COSM4096320	1	Extension	7	TCGA-BR-8690-01	Digestive System	Stomach	Adenocarcinoma	chr2:9490178-9490178	no	no
ALLC	c.1175A>T	p.*392L	COSM3581448	1	Extension	17	TCGA-EE-A29E-06	Integumentary System	Skin	Malignant Melanoma	chr2:3702562-3702562	no	no
C2orf50	c.489G>T	p.*163Y	COSM5033173	1	Extension	66	LP6005500-DNA_B03	Digestive System	Esophagus	Adenocarcinoma	chr2:11144111-11144111	no	no
CPSF3	c.2053T>C	p.*685R	COSM6634849	1	Extension	6	T261	Digestive System	Large Intestine	Adenocarcinoma	chr2:9473015-9473015	no	no
KCNF1	c.1483T>C	p.*495R	COSM5040911	1	Extension	53	3101B7_032_T	Digestive System	Liver	Hepatocellular Carcinoma	chr2:10913909-10913909	no	no
KIDINS220	c.5316A>G	p.*1772W	COSM6202927	1	Extension	14	CHC2697T	Digestive System	Liver	-	chr2:8730720-8730720	no	no
RNF144A	c.878A>G	p.*293W	COSM3737628	1	Extension	71	SJRH8036A	Musculoskeletal System	Striated Muscle	Rhabdomyosarcoma	chr2:7039759-7039759	no	no

Showing 1 to 7 of 7 entries

Download Table

Download Full Results

Previous

1

Next

Figure 4: Search by Region (genome coordinates).

# Search by Organ

This option allows the search for nonstop mutations in a specific tumor entity on the basis of their site of origin in a hierarchical manner. The user first selects an organ system followed by site (tumor / organ) and histology of interest. The organ systems are divided into nine classes (shown in Figure 5): Cardiovascular System, Digestive System, Endocrine System, Genitourinary System, Integumentary System, Lymphatic System, Musculoskeletal System, Nervous System and Respiratory System.

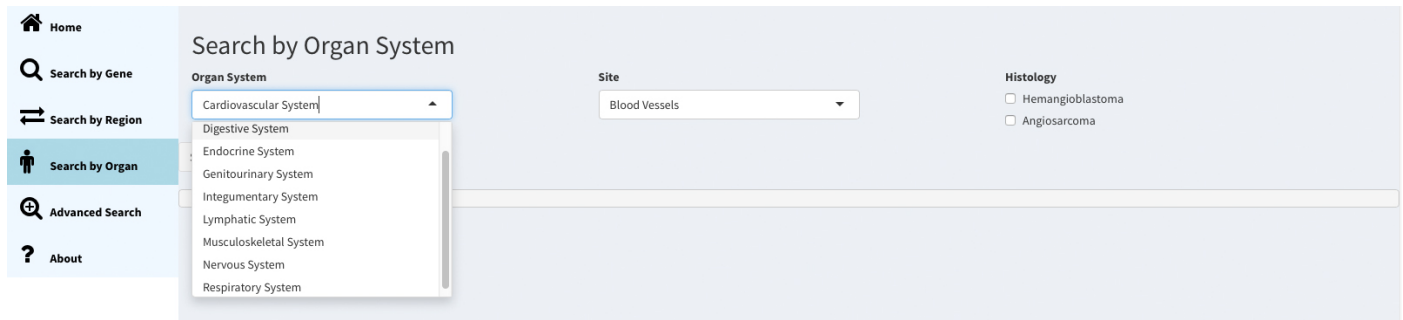


Figure 5: Search by organ - selection of organ system.

After selection of the organ system, the user selects an organ of interest / primary site and histology (Figure 6). The following example shows a list of nonstop mutations present in the digestive system followed by the choice of site and histology (Figure 7).

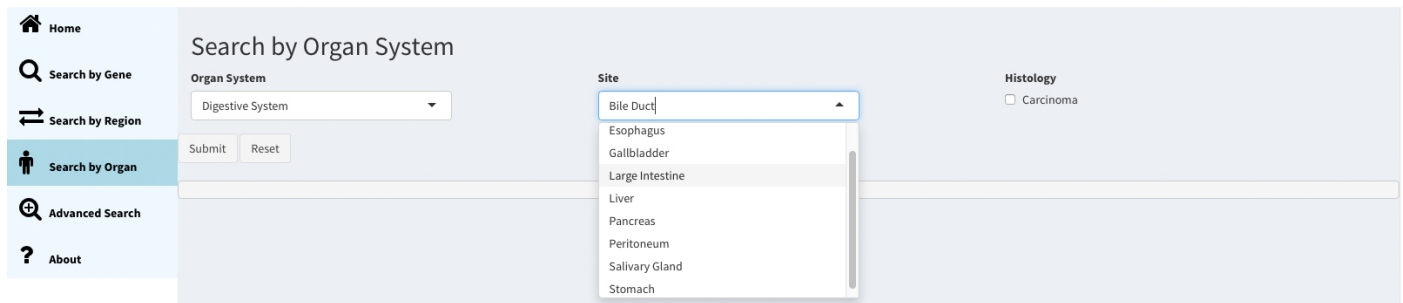


Figure 6: Search by organ - selection of primary site.

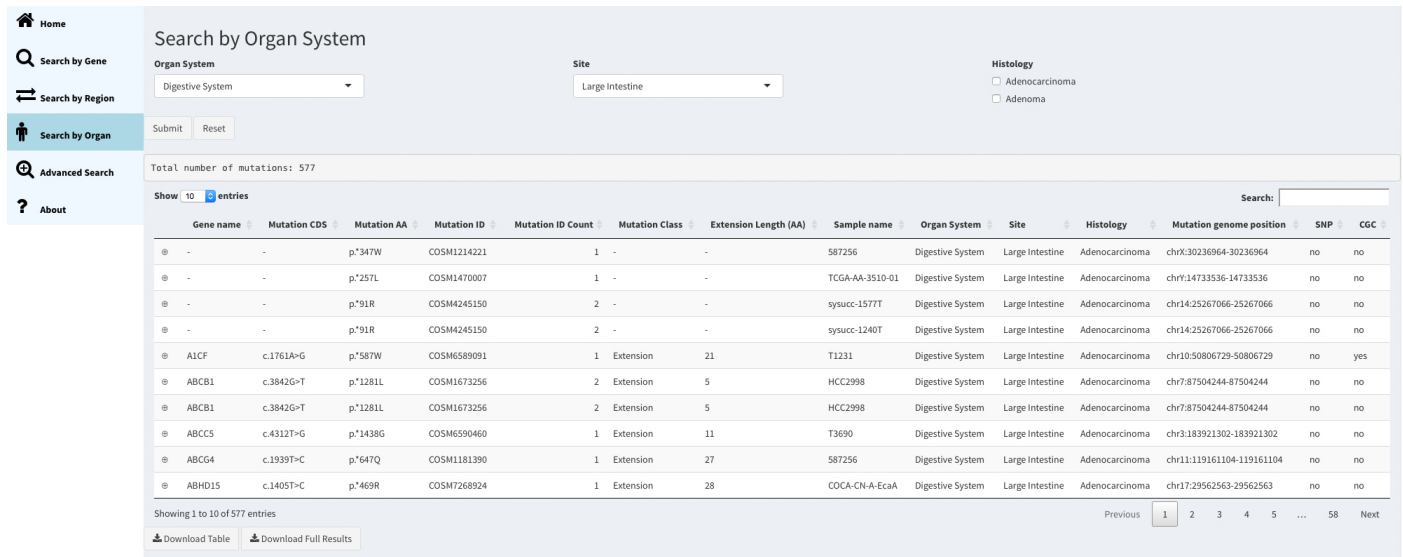


Figure 7: Search by organ - selection of primary histology and results.

# Advanced search

This search option has multiple search parameters including extension length, site and histology of nonstop mutations, as well as the mutation type (Figure 8). Here, users can also perform batch searches by providing a list of up to 100 genes. Additionally, all or only selected genes listed as cancer-related genes in the Cancer Gene Census (v89) can be included.

Home

Search by Gene

Search by Region

Search by Organ

Advanced Search

About

Advanced search

Gene list

Paste list of genes (max 100), one gene per row.

Extension Length

0

679

☐ Limit to CGC genes

Organ System

Digestive System

Site

Large Intestine

Nonstop Mutation Type

All Nonstop Mutations

Histology

☐ Adenocarcinoma

☐ Adenoma

Submit

Reset

Download Table

Download Full Results

Figure 8: Advanced search.

Below is an example of a search for nonstop mutations that are found in cancer of the large intestine in in all known cancer genes (Figure 9).

Home

Search by Gene

Search by Region

Search by Organ

Advanced Search

About

Advanced search

Gene list

Paste list of genes (max 100), one gene per row.

Extension Length

0

679

☒ Limit to CGC genes

Organ System

Digestive System

Site

Large Intestine

Nonstop Mutation Type

All Nonstop Mutations

Histology

☐ Adenocarcinoma

☐ Adenoma

Submit

Reset

Total number of mutations: 26

Showing 10 entries

Search:

Gene name	Mutation CDS	Mutation AA	Mutation ID	Mutation ID Count	Mutation Class	Extension Length (AA)	Sample name	Organ System	Site	Histology	Mutation genome position	SNP	CGC
ATM	c.9171A>C	p.*3057C	COSM6045693	1	Extension	29	HUB-02-B2-042	Digestive System	Large Intestine	Adenocarcinoma	chr11:108365508-108365508	no	yes
B2M	c.358T>C	p.*120Q	COSM3690641	2	Extension	65	TCGA-CA-6717-01	Digestive System	Large Intestine	Adenocarcinoma	chr15:44716340-44716340	no	yes
CHCHD7	c.205T>C	p.*69R	COSM1673922	1	Extension	24	HCT-116	Digestive System	Large Intestine	Carcinoma	chr8:56216503-56216503	no	yes
CIITA	c.3392G>C	p.*1131S	COSM5072257	1	Extension	35	TCGA-AG-A00Y-01	Digestive System	Large Intestine	Adenocarcinoma	chr16:10923302-10923302	no	yes
FBXW7	c.2124A>G	p.*708W	COSM6662607	1	Extension	53	T605	Digestive System	Large Intestine	Adenocarcinoma	chr4:152322881-152322881	no	yes
HEY1	c.915A>C	p.*305Y	COSM290805	1	Extension	6	TCGA-AG-A036-01	Digestive System	Large Intestine	Adenocarcinoma	chr8:79765188-79765188	no	yes
LRP1B	c.13798T>C	p.*4600Q	COSM4698820	2	Extension	12	T578	Digestive System	Large Intestine	Adenocarcinoma	chr2:140233188-140233188	no	yes
LRP1B	c.13798T>C	p.*4600Q	COSM4698820	2	Extension	12	T578	Digestive System	Large Intestine	Adenocarcinoma	chr2:140233188-140233188	no	yes
MAP2K1	c.1180T>C	p.*394Q	COSM1214372	1	Extension	36	587392	Digestive System	Large Intestine	Adenocarcinoma	chr15:66490613-66490613	no	yes
MAX	c.291A>C	p.*97Y	COSM6708941	1	Extension	23	T3611	Digestive System	Large Intestine	Adenocarcinoma	chr14:65084180-65084180	no	yes

Showing 1 to 10 of 26 entries

Download Table

Download Full Results

Previous123Next

Figure 9: Advanced search - results showing only mutations within CGC genes occurring in cancer of the large intestine.