

NGS Catalog Manual

8/14/2012

1. Data Source

Species	NGS Technologies	No. Of Records	No. of Journal	Data collected from
Human	Exome Sequencing, Whole Genome Sequencing, Transcriptome sequencing (RNA-Seq)	285	64	Literature

2. Description of Column Headings for NGS Catalog

We have collected the following information about NGS studies,

Paper Information:

Author: Last name of first author

Date: Publication data (online (electronic publication) data if available)

Journal: Full journal name

Title: Title of paper (linked to PubMed abstract)

Study Information:

Sequencing Technology: Next-generation sequencing method used in the study, include exome sequencing, RNA sequencing and whole genome sequencing

Platform: Sequencing platform used in study

Disease/Trait: Disease or trait examined in study

Mutation type: Somatic mutation or germline mutation examined in study

Population: Population of the samples in study

Samples: Sample constitution in study

Max Read Length (bp): Maximum read length in study

Average Fold Coverage: The average coverage of sequencing data in study

Mapping Tool(s): Alignment tools used in study

Variant Calling Tool(s): Variant calling tools used in study

Reference Genome: Reference genome used in study

Public SNPs Database: Known SNPs used in study

Results Summary:

SNVs: Number of total and novel SNVs detected

Short Indels: Number of total and novel indels detected

Copy Number Variations: Number of total and novel copy number variations

Large Structural Variants: Number of total and novel structural variations

Reported gene(s): Gene(s) reported by author

Identified gene fusion(s): Identified gene fusion by author

Identified mutation(s):

3. Web interface

Users can browse or search the data at different levels.

Browse:

- 1) browse by sequencing technology;
- 2) browse by platform;
- 3) browse by disease;
- 4) browse by publication date (beginning date and end date);
- 5) browse by journal.

Search:

1) Quick search can be performed by choose any from “*Sequencing technology*”, “*Platform*”, “*Disease/Traits*”, “*Journal*” and “*Date of journal published*” (begining date and end date) or performed by entering value in the search box of **Software names** (e.g., BWA), **Author name** (i.e last name of the first author, like Puente), and **Gene names** (e.g., MYD88).

2) If searching without any input, it will display all the results from our database.

Searching based on browse options:



Home | [Documents](#) | [Software](#) | [Pipelines](#) | [Upload](#) | [Feedback](#) | [Citation](#) | [Contact](#)

Disease / Trait:	<input type="text"/>	Sequencing technology:	<input type="text" value="Exome Sequencing"/>
or:	<input type="text"/>	Platform:	<input type="text"/>
Gene:	<input type="text"/>	Software:	<input type="text"/>
Mutation type:	<input type="text"/>	First author (last name):	<input type="text"/>
Journal:	<input type="text"/>	From:	<input type="text"/> <input type="text"/> <input type="text"/>
or:	<input type="text"/>	To:	<input type="text"/> <input type="text"/> <input type="text"/>

157 records were found in the NGS catalog database.

Searching based on multiple selections:

The screenshot shows the NGS Catalog search interface. The search form is highlighted with red circles, indicating the following selections:

- Disease / Trait: Chronic lymphocytic leukemia
- Sequencing technology: Exome Sequencing
- Gene: SF3B1

Below the search form, a search result is displayed with a table of details:

Author / Date / Journal / Title	Mutation Type	Disease / Trait	Total SNVs	Novel SNVs	Short Indels	Copy Number Variations	Large Structural Variants	Reported Gene(s)	Identified Gene Fusion(s)
Quesada et al. 2011-12-11 Nature Genetics Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia	Somatic mutation	Chronic lymphocytic leukemia						NOTCH1 SF3B1 POT1 CHD2 LRP1B	

User can get more detailed information like the sample information, information about the tools used (mapping and variant calling tools), PubMed link to the paper and NCBI gene link to the genes etc. by clicking on the particular paper title in the results table. In addition, we have provided the mutations information by clicking on mutation type (somatic mutation or germline mutation)

4. Data Download:

Users can download the results option using the **SAVE** option above the results tables. You can either download the file in CSV format.