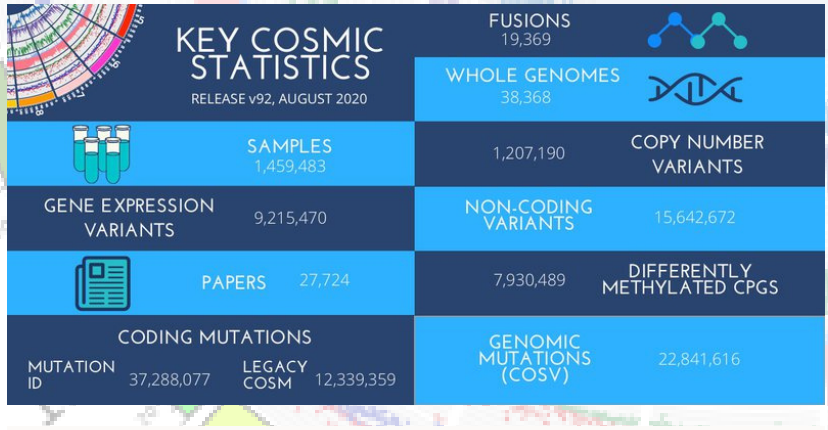
Discover what COSMIC has to offer

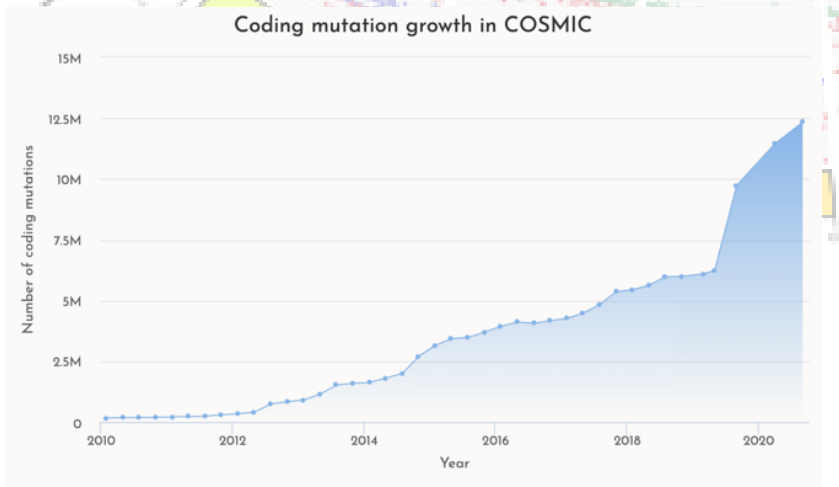
COSMIC – the Catalogue of Somatic Mutations in Cancer – is the world's largest source of expert manually curated somatic mutation information relating to human cancers.

COSMIC provides a huge breadth and depth of high quality somatic data, which has been carefully curated, combined and standardised, across 16 years of expert manual curation.

To date, we list over 37 million coding mutations plus substantial coverage of all other oncogenic mutation mechanisms (COSMIC v92, August 2020) for exploration, available on both GRCh 37 (hg19) and GRCh 38 (hg38) with consequence annotations on both parallel transcriptomes and proteomes.

We pride ourselves on being the gold standard in somatic databases and boast well over 10,000 citations. Our user base consists globally in excess of 20,000 registered scientists, bioinformaticians and clinicians, including multiple field leaders and prominent clinical and pharmaceutical companies.





- COSMIC
- Cell Lines Project (CLP)
- COSMIC 3D
- Cancer Gene Census (CGC)
- Cancer Mutation Census (CMC)
- Mutational Signatures
- Mutation Actionability in Precision Oncology

Be sure to follow us on social media and sign up to our release announcements to keep abreast of our latest developments.

cosmic-sanger cosmic_sanger COSMIC blog

COSMIC Catalogue Of Somatic Mutations In Cancer

wellcome sanger institute

Version 2.2 March 2021

Discover what COSMIC has to offer

COSMIC

COSMIC – the Catalogue of Somatic Mutations in Cancer – is the world's largest source of expert manually curated somatic mutation information relating to human cancers. COSMIC provides a huge breadth and depth of high quality somatic data, which has been carefully curated, combined and standardised, across 16 years of expert manual curation.

The COSMIC suite includes:

Cell Lines Project (CLP)

The Cell lines site provides an interface for the Cell Lines Project, based at the Wellcome Sanger Institute, (UK). COSMIC cell lines coverage provides high quality, consistent annotations for coding mutations, expression variants and copy number variants across 1,020 cell lines commonly used as laboratory models. For more information, click here.

COSMIC 3D

A platform for exploring cancer mutations and how they cluster, in the context of 3D protein structures.

Cancer Gene Census (CGC)

The Cancer Gene Census (CGC) is an ongoing effort to catalogue those genes which contain mutations that have been causally implicated in cancer and explain how dysfunction of these genes drives cancer. The content, the structure, and the curation process of the Cancer Gene Census was described and published in Nature Reviews Cancer.

Hallmarks of Cancer

Based on the concept defined by D. Hanahan and R. A. Weinberg, COSMIC, in collaboration with Open Targets, integrates functional descriptions focused on Hallmarks of Cancer into the Cancer Gene Census. Hallmark pages visually explain the role of a gene in cancer by highlighting, which of the classic behaviours are displayed by the gene product and whether they are promoted or suppressed.

Cancer Mutation Census (CMC)

The Cancer Mutation Census (CMC) is a new system characterising the impact of every mutation in cancer. It integrates transparent metrics on multiple biological and population properties, with observations in clinical databases to rank and score which mutations are most important in every form of human cancer.

Mutational Signatures

Different mutational processes generate unique combinations of mutation types, termed "Mutational Signatures" which characterises the biological processes driving these genetic changes. Part of the Mutographs, CRUK Grand Challenge.

Mutation Actionability in Precision Oncology

The aim of COSMIC Actionability is to comprehensively curate the current state of precision oncology, describing in substantial detail which drugs target specific mutations in a range of human cancers, from marketed drugs through clinical development phases.

Be sure to follow us on social media and sign up to our release announcements to keep abreast of our latest developments.

cosmic-sanger cosmic_sanger COSMIC blog

COSMIC Catalogue Of Somatic Mutations In Cancer

wellcome sanger institute

Version 2.2 March 2021

Manual curation

Our curation team of postdoctoral scientists is responsible for the manual curation of the key cancer driving genes listed in the Cancer Gene Census. Our manual approach to curation allows us to capture great depth, across a wide variety of high resolution phenotypes, capturing not only the phenotypes but the mutations and extensive annotations such as the patient characteristics and journey, drug resistance information, environmental exposures and therapeutic response.

Genomic curation

Our genomic curation captures a broad perspective of somatic data, providing mutation annotations on every human gene on every human regulatory element. Our genomic screen data includes peer reviewed large scale genome screening data from over 500 independent cancer genome studies plus data from International Cancer Genome Consortium (ICGC), The Cancer Genome Atlas (TCGA) and Pediatric Cancer Genome Project (PCGP). We have a semi-automated process to capture genomic screens in which a software pipeline interprets the genomic coordinates, calculates protein consequences and follows the latest industry standards for genomic and proteomic data including following HGVS standards for mutation annotations, as a product of the latest Ensembl release version.

A very explorable dataset of over 1,500 cancer types

By combining the broad genomic data together with a deep genic data and standardising it in one place, COSMIC provides a very explorable dataset of over 1,500 cancer types. During the lifespan of COSMIC, we have created a unique classification system presenting four levels of cancer site and four levels cancer histology for tested mutations, which is updated at every release. We loosely follow the World Health Organisation (WHO) classification system for tumours with greater detail, often preceding the approvals to WHO.

We always strive to capture and combine some of the least accessible cancer mutation data from a wide variety of sources, to create a broad genetic resource across all forms of human cancer. We ensure COSMIC is at the forefront of fully standardised and FAIR bioinformatic resources. Current standards of nomenclature, annotation and description are adopted to enhance COSMIC's support for FAIR principles, increasing its accessibility and interoperability.

What does COSMIC include?

High quality manual curation and genome-wide screen data

- Coding & non-coding
- Copy number variations
- Structural rearrangements
- Tissue distribution
- Drug resistance data
- Gene expression and methylation data
- Gene fusions
- Feature data

Website tools to explore the data

- Gene and Gene Fusion pages presenting comprehensive genotypic and phenotypic data
- Cancer Browser offering a disease focused view
- COSMIC 3D to explore protein structure with COSMIC mutations imposed on them
- CONAN — a tool to indicate the copy number of selected genes
- GA4GH Beacon Query — a GA4GH compliant method to query variants across a single nucleotide
- COSMIC in BigQuery — search COSMIC via the ISB Cancer Genomics Cloud

COSMIC
Catalogue Of Somatic Mutations In Cancer

wellcome sanger institute

Version 2.2 March 2021

Access to download COSMIC data in descriptive and industry standard forms

The COSMIC data is available via the website, for all users and the download files for registered users. You may need a license to access the download files.

Available from the website as:

- Download data files (file description available for each of files)
- Filtered download files
- Scripted downloads

For more information please access our Licensing page.

Download taster files of COSMIC data

Want to have a look at the data yourself? We have made the first 100 lines of each download file freely available so you can try out the data. Please see the COSMIC about page for further details and to download samples.

Regular data releases

New and potentially significant data are continually captured and made available through three significant updates to COSMIC each year. Further details can be found in our Release Notes and on our blog.

Licensing the COSMIC database

We have a range of licensing options to suit your needs. The fee structure is intended to be fair and was designed in consultation with industry. We have various commercial licensing models to fit various levels of data distribution and several tiers of pricing based on the company size. We want to encourage and support scientific entrepreneurship wherever possible, thus the use of COSMIC data by startups is discounted, upon qualifying conditions.

A commercial license will allow you to:

- Download the entire COSMIC database in easily integrable files
- Distribute part of or the full COSMIC database (depending on license type)
- Full access to the helpdesk for support

If you require a COSMIC licence and would like to find out more about our licensing models and pricing, please contact our bioinformaticssales@qiagen.com, our commercial distribution partner.

Additional information about our licensing can be found on our licensing page.

Contact details
Kate Noble
COSMIC Business Manager
+44 (0)1223 499 943
cosmic-translation@sanger.ac.uk

Wellcome Sanger Institute
Wellcome Genome Campus
Hinxton
Cambridgeshire, CB10 1SA
United Kingdom

COSMIC
Catalogue Of Somatic Mutations In Cancer

wellcome sanger institute

Version 2.2 March 2021