

Cellosaurus newsletter 10 of January 2023

1) Summary of changes from release 39 to release 44

Since our last newsletter in October 2021, we have deployed five Cellosaurus releases; the latest one is release 44 of December 12, 2022 which contains information on 144,568 cell lines from 805 species. During this 15 months period we added about 13,600 new entries, but also updated the information contained in more than 50% of the existing entries.

Among the entries added, 3,000 are hybridomas, thus the number of this category of cell lines increased from 7,000 to slightly above 10,000. Hybridomas are the “dark matter” of cell lines: many are developed by pharmaceutical companies and are not described in any publication (only a subset is described in patents), similarly biotech companies that sell monoclonal antibodies rarely describe or distribute the parent hybridomas. Thus only a small subset of all the developed hybridomas can become Cellosaurus entries, but we are committed to cover these important cell lines.

We have continued retrofitting information in existing entries with an emphasis in recovering as much as possible information on the site of sampling (organ or tissue) and the cell type. We plan, in 2023, to complete this particular retrofitting activity and to map all terms used for anatomy and cell type to the UBERON and CL ontologies. To prepare this important change, Paula Duek who works part time on the Cellosaurus has been very active in creating new terms for concepts which were missing in these two powerful ontologies.

The rapid expansion of the number of references cited in Cellosaurus entries is continuing. We have added 3,100 references in these 15 months to reach a grand total of 25,963 distinct references (23,882 journal publications, 1,687 patents, 255 book chapters and 139 theses).

In term of STR profile information we have now almost 8,500 entries that contain such information, an increase of about 500 entries since release 39.

The Cellosaurus is now cross-referenced to 102 distinct resources as we have added cross-references to 10 additional resources: Abcam, CancerTools, dbGAP, DSMZCellDive, EGA, ICLDB, Innoprot, NRFC, PubChem cell line pages and Rockland while cross-references to 3 resources were removed (CCLE, eagle-I and TCB).

We have recently created three new comment (CC) topics to specifically store some of the information that was previously stored either in the characteristics or miscellaneous comment topics:

- Various information regarding the donor of the cell line which was formerly found in the "Miscellaneous" comment is now stored in a new CC line "Donor information". Example: *Donor information: At sampling donor was not affected with Alzheimer disease but has a 25% risk.*
- Information regarding finite cell line senescence which was formerly found in the "Characteristics" comment is now stored in a new CC line "Senescence". Example: *Senescence: Senescens at 18 PDL (PubMed=12077343).*
- Information regarding the susceptibility of a cell line to viral infection, the presence of integrated viruses or any other virology-related information which were all formerly found in the "Characteristics" comment are now stored in a new CC line "Virology". Example: *Virology:*

Susceptible to infection and early propagation by guinea pig adenovirus (GPA_{AdV}) (PubMed=32226606).

2) The Cellosaurus on ExPASy

The traffic toward the Cellosaurus on ExPASy is continuing to increase. Since it was made available in May 2015, it has been visited 4.2 million times by almost 2.0 million distinct users that have browsed 12.7 million pages.

In August 2022, the Cellosaurus, while continuing to be hosted on the SIB ExPASy server, acquired its own domain name: “cellosaurus.org”. Thus its home page is now <https://www.cellosaurus.org/>. All existing external links are functional as they are automatically redirected to the new ones.

3) New educational resources for the Cellosaurus

Thanks to Monique Zahn from the SIB Training group there is now an online introductory course on the Cellosaurus that takes about an hour to complete: <https://edu.sib.swiss/course/view.php?id=585>

Camille Mary who is working part time for the Cellosaurus has created two short educational video that were posted on the SIB YouTube channel, one is to explain the process behind the creation of a cell line Research Resource Identifier (RRID) (see <https://www.youtube.com/watch?v=Cz64B4ShS64>) while the second one is an introductory video which is targeted towards lab scientists that have never heard of the Cellosaurus. We thus encourage you to share this second video with as many scientists as possible: <https://www.youtube.com/watch?v=xKA2Alele0g>

4) The Cellosaurus Application Programming Interface (API)

Pierre-André Michel who is in charge of Cellosaurus software development recently introduced an API for the Cellosaurus (<https://api.cellosaurus.org/>). It is a very powerful API that allows to query and retrieval of Cellosaurus data at a very fine-grained level (58 distinct fields are available). The API response can be in four different formats: Json, XML, TSV and text. Extensive help is available from the API home page and we also provide a streamline interface for users to test their queries (<https://api.cellosaurus.org/search/form.html>)

5) The Cellosaurus scientific advisory board (SAB)

In early 2022 we established a Scientific Advisory Board (SAB) for the Cellosaurus. The SAB is composed of seven external experts widely recognized for their contribution in the fields of activity of the Cellosaurus. The SAB acts as a consultative body, its principal responsibilities being to advise on scientific objectives and priorities, to evaluate the outcomes and to propose means of improving outcomes and visibility. More information on the members of the Cellosaurus SAB is available at: <https://www.cellosaurus.org/SAB.html>

PS: Do not forget to subscribe to our Twitter page (<https://twitter.com/Cellosaurus>) for tweets about new developments regarding the Cellosaurus and the universe of cell lines. Alternatively if, like us, you are a bit fed up with the vagaries of this social platform since it was acquired by Elon Musk we now also have a Mastodon account: <https://fediscience.org/@cellosaurus>