

## Misidentified and non-verifiable cell lines

*Last updated: 5 February 2025*

### Cell lines

A cell line is a population of cells that can be grown in a laboratory culture indefinitely. Cell lines are an essential tool for biomedical research because they allow biological experiments to be performed in vitro (i.e. outside of a living organism). For instance, researchers developing new cancer drugs will certainly test the effect of their drug candidates on many different cancer cell lines before ever considering testing the drug candidate on live animals, let alone human patients.

Because cell lines can grow indefinitely, one research laboratory or laboratory supplier can take a few cells from their cell line stock and give them to another laboratory for that laboratory to begin culturing their own stock. By far the most popular cell line is HeLa, and the many thousands of [HeLa cell stocks](#) used in biomedical research around the world can all be traced back to the original stock of "immortal" cervical cancer cells taken from [Henrietta Lacks in 1951](#).

### Misidentified/contaminated cell lines

One complication of relying on cell lines for biomedical research is that stocks will often become cross-contaminated by other cell lines. For instance, HeLa is extremely aggressive as cancer cell lines go and HeLa cells will easily overtake other cell line stocks that are stored nearby. The popular gastric cancer cell line [BGC-823](#) is one such victim of HeLa's aggression; there are no remaining uncontaminated stocks of BGC-823 and thus it is no longer considered an appropriate experimental model for gastric cancer.

Cellosaurus [BGC-823 \(CVCL\\_3360\)](#)

[\[Text version\]](#)

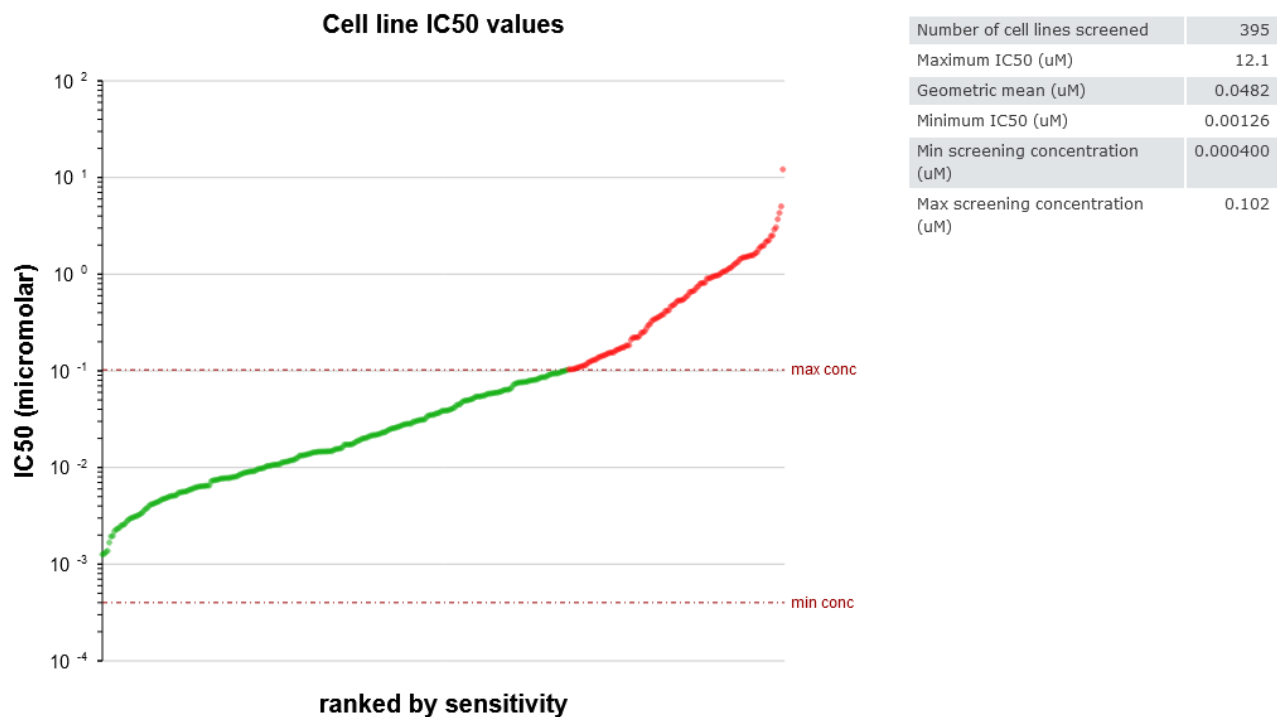
Cell line name	BGC-823
Synonyms	BGC823
Accession	CVCL_3360
Resource Identification Initiative	To cite this cell line use: BGC-823 (RRID:CVCL_3360)
Comments	<p><b>Problematic cell line: Contaminated. Shown to be a HeLa derivative (CCRID; PubMed=26116706; PubMed=28851942). Originally thought to be a gastric adenocarcinoma.</b></p> <p>Part of: Cancer Dependency Map project (DepMap) (includes Cancer Cell Line Encyclopedia - CCLE).</p> <p>Registration: International Cell Line Authentication Committee, Register of Misidentified Cell Lines; ICLAC-00570.</p> <p>Population: African American.</p> <p>Transformant: NCBI_TaxID: <a href="#">333761</a>; Human papillomavirus type 18 (HPV18).</p> <p>Omics: Proteome analysis by 2D-DE.</p> <p>Omics: Transcriptome analysis by microarray.</p> <p>Miscellaneous: Formerly the CCRID database had 2 entries describing this cell line (3111C0001CCC000062, 3131C0001000700011). They were one of the sources for the STR profile of this entry.</p> <p>Misspelling: BSG823; Note=Occasionally.</p> <p>Derived from site: In situ; Uterus, cervix; UBERON=<a href="#">UBERON_0000002</a>.</p>

*The Cellosaurus entry for [BGC-823](#) now warns about the cell line being contaminated by HeLa.*

The International Cell Line Authentication Committee maintains [a register of misidentified cell lines](#) that is currently 593 entries long. A good number of these cell lines were contaminated by cells from a completely different organism, such as human salivary gland cell line [CAC2](#), which is actually made up of unknown cells from a rat.

[Cellosaurus](#) is an encyclopedia of thousands of cell lines and will link to studies showing that a cell line is contaminated or otherwise misidentified.

Because different cell lines can have a wide variety of responses to the same treatment, it is essential that researchers avoid misidentified cell lines and only work with authenticated cell lines that actually represent their system of interest. Consider the fact that cancer cell lines can have wildly different sensitivity to the chemotherapy drug [paclitaxel](#).



*Half-maximal inhibitory concentration (IC50), in this context, is a measure that describes what concentration of a drug is needed to inhibit the growth of a cell line by 50%. The [Genomics of Drug Sensitivity in Cancer \(GDSC\) database](#) estimated these values for 395 cell lines for the drug paclitaxel. Note that paclitaxel can be more than a thousand times as potent against some cell lines than others, even for cell lines from the same type of cancer.*

## Cell line verification/authentication

Researchers should always verify that their cell lines stocks are what they believe them to be. There are a number of ways researchers can verify the integrity of their cell line stocks (many of which are [detailed by the American Type Culture Collection](#)). The most common authentication technique is [short tandem repeat \(STR\) profiling](#), which identifies specific molecular signatures in a cell line's genome.

## Non-verifiable cell lines

The names of cell lines are usually just a jumble of letters and numbers and thus can often be easily confused. For instance, the name of [BGC-823](#) is very similar to that of [MGC-803](#), another misidentified gastric cancer cell line. One might easily misspell these cell lines as "BGC-803" or "MGC-823". However, it was recently reported by [Oste et al. \(2024\)](#) that many publications will use these misspelled identifiers to refer to another cell line entirely distinct from these existing cell lines. For instance, [Zhong et al. \(2021\)](#) report experiments in cell lines "BGC-803" and "BSG-823" in addition to experiments in the contaminated cell lines BGC-823 and MGC-803.

Hundreds of articles have referred to experiments in these cell lines despite there being no indication that these cell lines actually exist; there are no entries for these cell lines in any cell line indices, they cannot be found in any supplier catalogs, there are no articles describing how these cell lines were established and no one seems to have produced any genetic profiles of these cell lines to confirm their identities. Oste et al. identified eight such "miscellings": BGC-803, BSG-803, BSG-823, GSE-1, HGC-7901, HGC-803, MGC-823 and TIE-3, although there are certainly many more that have not been studied in detail.

### Example 1: Contaminated cell lines

[Liu et al. \(2017\)](#) report experiments in the cell lines [MGC-803](#), [L02](#) and [SMMC-7721](#). However, each of these cell lines are contaminated by HeLa and thus are no longer considered suitable models for their respective cancers.

### Example 2: Non-verifiable and contaminated cell lines

[Yang et al. \(2018\)](#) mention and provide experimental results in 10 different cell lines, of which 4 are problematic, shown below in bold:

- **SUN-216**: mentioned once in text of paper and again in Figure 1B, both spelled as SUN-216. SNU-216 is an existing cell line. A possible non-verifiable cell line identifier, but not one that has yet been studied in depth.
- **BGC-823**: Contaminated cell line.
- AGS
- **BGC-803**: Mentioned once in text and again in Figure 1B. One of the eight non-verifiable cell line identifiers studied by Oste et al. Likely derived from a typo that confused the cell lines MGC-803 and BGC-823, both contaminated.
- NUGC4
- MKN74
- MKN45
- **SGC-7901**: Contaminated cell line.
- HGC-27
- GES-1

## **Additional resources**

- [ATCC Cell Line Authentication Test Recommendations](#)
- [Cellosaurus](#)
- [International Cell Line Authentication Committee \(ICLAC\)](#)
- [ICLAC-curated reviews on cell line misidentification](#)
- ["Misspellings or 'miscellings'—Non-verifiable and unknown cell lines in cancer research publications" \(2024\)](#)