Guide to Human Cell Line Authentication

This protocol sets out the essential steps to follow when authenticating human cell lines.

SUMMARY

1. Send the frozen cells or their isolated DNA to a genotyping laboratory that works according to the global standard for the authentication of human cell lines [1].
2. Compare the short tandem repeat (STR) profile obtained to the STR reference profile from leading cell banks via integrated WEB search engines or Cellosaurus [2-4]. A minimum of 13 STR loci must be analysed [1].
3. If the cell line name is identical and both STR profiles show matches of more than 80% when considering the STR data for at least 13 loci, the sample is authenticated to be from the original donor. However, this result does not mean that all samples of this cell line will be authentic.
4. If you do not receive the name of the cell line under investigation and STR matches are below 80% to any STR data set, the sample STR profile is unique and unknown to leading cell banks (rare case).
5. If you obtain the name of another cell line and the STR profile is the same or matches over 80%, the cell line is probably false or cross-contaminated.
6. Known false cell lines are listed in the ICLAC Register of Misidentified Cell Lines [5]. The committee provides detailed information in case of difficult authentication issues.

The right start of cell line authentication

As a rule, cell lines should be obtained from reliable and confidential sources such as public cell banks. If cell lines are available in a running laboratory, their identity should be checked before starting a new project. Always check the name of the cell line for existing problems before you start experiments and before you test it. This can be done using Cellosaurus, which provides a broad range of cell line information [3-4], and the ICLAC Register of Misidentified Cell Lines, which focuses specifically on false cell lines [5].

The latest versions of Cellosaurus and the ICLAC Register are available at:

https://www.cellosaurus.org

http://iclac.org/databases/cross-contaminations/

This step will alert you to possible problems with the cell line and may save you considerable time and effort. However, it should be noted that if a cell line is reported to be cross-contaminated, it does not necessarily mean that all samples are affected. There may be multiple stocks of cell lines with the same name from many sources, only some of which may be cross-contaminated. New cell lines that do not enter a laboratory through public cell banks should always be validated using STR genotyping.

A Starting material for STR profiling

The starting material -

1. **If establishing a cell line:**
   In the case of a *de novo* cell line establishment, freeze biomaterial from the donor in liquid nitrogen (part of the biopsy, blood cells, or cheek swab). The DNA from this sample can later be used to unequivocally prove that the new cell line is from the specified donor.
2. **For new scientific projects:**
   It is suggested that authentication tests be performed at the beginning and end of experimental work when working with cell lines already in the laboratory. In-lab cryopreservations should be subjected to STR profile testing after multiple vials from a large batch of a cell line have been

frozen in parallel. Continuous culture for more than 4 months or multiple freeze/thaw cycles without STR testing is considered poor tissue culture practice.

B. Generation of STR profiles

STR profiling should only be performed in-house if the testing is routinely performed and meets the criteria of the Human Cell Line Authentication Standard [1].

For external providers, it is also important to deliver fast and reliable service with appropriate accreditation and quality assurance. The supplier should provide an STR profile and STR reference data that comply with the standard [1].

Some STR genotyping providers require genomic DNA, while others accept either cryopreserved cells or genomic DNA. Always confirm requirements with the provider before preparing samples.

C. What to do with the result when it arrives: STR profile comparison

An STR profile is obtained by PCR amplification of a set of STR loci. For each STR locus, the profile consists of a series of numbers that correspond to the alleles found in the sample. For example, 13 STR loci give a maximum of 26 numbers (corresponding to 26 alleles) if all are heterozygous, in a "normal" cell line (non-cancer). For the authentication of cell lines, at least 13 nuclear STR loci are required plus amelogenin to determine the sex (XX or XY). Unfortunately, male-derived cells often lose the Y chromosome in culture [6], so sex determination can only be considered indicative.

The next step is to compare the STR profile with a reference sample: for newly produced cell lines this should be the donor's DNA, whereas for existing cell lines it is the STR reference data from the leading cell banks, where cell lines were often deposited by the laboratories that established them. Guidelines for interpreting the results are given in the Standard for Authentication of Human Cell Lines [1]. A worksheet is also available from ICLAC that summarises the essential concepts to consider when comparing a test sample to a second reference sample (see References and Resources section).

For a reliable result, a comparison should be made with all existing STR profiles available. For this purpose, there are widely used interactive online databases of cell banks with verified reference data or extensive independent platforms. Among the best known are:

CLASTR: https://www.cellosaurus.org/str-search/

ATCC: https://www.atcc.org/search-str-database

DSMZ: https://celldive.dsmz.de/str

Remarks: STR profiling does not allow you to differentiate between cell lines that have come from the same donor or cell lines that have been subcloned apart. Basically, there may be slight variations in the STR profiles from different cultures derived from the same individual. This can be caused by genetic drift, i.e., microsatellite instability or loss of heterozygosity events in the cell. It is important to use the compliance criterion (80% compliance threshold for results using at least 13 STR loci) as recommended by the standard [1] to account for a small variation in some cultured samples.

Using an STR profile database:

Enter your STR profile data as indicated in the search form for the different loci (e.g., D5S818, TH01, Amelogenin). The online STR profile databases use a minimum set of loci that must be present; additional loci increase the discriminatory power of the comparison and should be entered if available. The samples are compared using a matching algorithm. The closest matches are usually listed first in search results. To interpret your results, it is helpful to look at the percentage agreement or EV value. Authentic samples should always give a result in the range of agreement of 80 to 100% (EV 0.8 to 1.0). The Cellosaurus database (https://www.cellosaurus.org) is extremely helpful for the origin of a cell line, providing lists of

relevant publications, cell line relationships, and other data (subclones, parental / children cell lines, chromosomal translocations, specific genome mutations) [3].

D. What to do with misidentified cell lines

If your tests indicate that a cell line has been identified as incorrect, the holdings should be discarded and all colleagues who have other holdings of that cell line should be informed.

If you should discover cross-contamination, you can generally publish your finding that the cell line is incorrectly identified. If the cell line is widely used and you believe there are no authentic stocks, publishing it is an important step in alerting others in the scientific community.

Please also contact ICLAC at info@iclac.org so that the cell line can be checked for inclusion in the ICLAC Register of incorrectly identified cell lines.

References and Resources

References cited in this document:

[1] Korch CT et al. Human Cell Line Authentication: Standardization of Short Tandem Repeat (STR) Profiling, ASN-0002-2022. ATCC Standards Development Organization ASN-0002-2022. Manassas, VA, USA. Available at: https://webstore.ansi.org/standards/atcc/ansiatccasn00022022

[2] Dirks WG et al. Cell line cross-contamination initiative: an interactive reference database of STR profiles covering common cancer cell lines. Int J Cancer 2010; 126(1): 303-304 doi: 10.1002/ijc.24999.

[3] Bairoch A. The Cellosaurus, a cell-line knowledge resource. *J Biomol Tech* 2018; 29(2): 25-38. doi: 10.7171/jbt.18-2902-002. Cellosaurus available at: https://cellosaurus.org

[4] Robin T et al. CLASTR: The Cellosaurus STR similarity search tool - A precious help for cell line authentication. *Int J Cancer* 2020; 146(5): 1299-1306. doi: 10.1002/ijc.32639. CLASTR available at: https://www.cellosaurus.org/str-search/

[5] Capes-Davis A et al. Check your cultures! A list of cross-contaminated or misidentified cell lines. *Int J Cancer* 2010; 127(1): 1-8. doi: 10.1002/ijc.25242.

[6] Yu M et al. A resource for cell line authentication, annotation and quality control. Nature 2015; 520(7547): 307-311 doi: 10.1038/nature14397.

For additional ICLAC resources relating to contaminated cell lines and authentication testing, see:

- ICLAC Register of Misidentified Cell Lines
  http://iclac.org/databases/cross-contaminations/
- Worksheet: Match Criteria for Human Cell Line Authentication
  http://iclac.org/resources/match-criteria-worksheet
- Advice to Scientists: Making Cell Line Authentication Part of Everyday Culture Practice
  https://iclac.org/resources/advice-scientists/