# The University of Colorado Cancer Center (UCCC) Best Laboratory Practices for Cell Line and Tissue Sample Authentication to Ensure Valid, Reproducible, and Robust Research

Christopher T Korch[1] and Rebecca E Schweppe[2]

[1]Division of Medical Oncology
[2]Division of Endocrinology, Metabolism, and Diabetes
School of Medicine
University of Colorado - Anschutz Medical Campus
12801 East 17th Avenue, RC-1 South
Aurora, Colorado 80045, USA

# Table of Contents

## Executive Summary

### Rationale & Goals

Begley and others have reported finding that a majority (50% - 90%) of important published biomedical research is irreproducible [11]. The contamination or misidentification of cell cultures (cell lines, stem cells, and xenografts, both of tissue samples and cell suspensions) can seriously compromise research using these cells to model diseases, biological phenomena, and drug testing in pre-clinical trials. The biomedical research community has a major problem with using imposter cell lines and incorrect tissue samples. The world-wide average incidence of using misidentified cell lines is 22% [38]; whereas, the incidence of such cell lines and tissue samples at the University of Colorado School of Medicine is 29%. This problem is partially addressed by granting agencies and journals requiring some type of cell line authentication. The NIH has issued three notices, effective January 25, 2016, that "key biological and/or chemical resources" need to be regularly authenticated (NIH NOT-OD-15-103, -16-011, -16-012), and guidelines for reporting preclinical research (http://www.nih.gov/about/reporting-preclinical-research.htm). However, the best approach to this problem is to tackle it at its source; specifically in the research laboratories of UCCC members.

To meet this need, we have designed a set of Best Laboratory Practices (BLPs) for the authentication of cell lines and tissue samples by four methods. Briefly, the four methods for authentication of cell lines and tissue samples include:

1. Verification of cell line identity via the Cellosaurus and ICLAC websites,
2. Genotyping by STR genotyping,
3. Identification of any non-human species present in human cultures, and
4. Mycoplasma testing of all cell cultures.

### Critical Validation Steps and Methods for Cell Lines and Tissues

These BLPs apply to all laboratory members of CU research teams using cell lines, xenografts, and tissue samples in their research where the identity of the samples at the species and tissue level are specified and critical for the research. To utilize these BLPs, we have developed detailed recommendations of how they can be implemented in the research laboratories of members of the CU Cancer Center summarized in the following critical validation steps and methods one should use before, during, and upon completion of a project.

1. Before starting a project, all cell lines intended to be used should be checked that they are neither listed in Cellosaurus as being problematic cell lines nor included in the ICLAC list of misidentified cell lines: http://iclac.org/databases/cross-contaminations.
2. All cell lines for a project should be obtained from a validated source (if possible, from e.g., the ATCC, DSMZ, ECACC, JCRB, RIKEN) that provides both STR genotyping and mycoplasma testing data for the samples. The UCCC Cell Technology Shared Resource (CTSR) sells authenticated cell lines that are certified free of mycoplasma.
3. Annual analysis by STR / DNA genotyping of human cell lines or tissue cultures is strongly recommended to confirm the authenticity of their genetic identity. Methods include STR

profiling, which is available through the Barbara Davis Center BioResources Core Facility Molecular Biology Unit at the University of Colorado Anschutz Medical Campus.

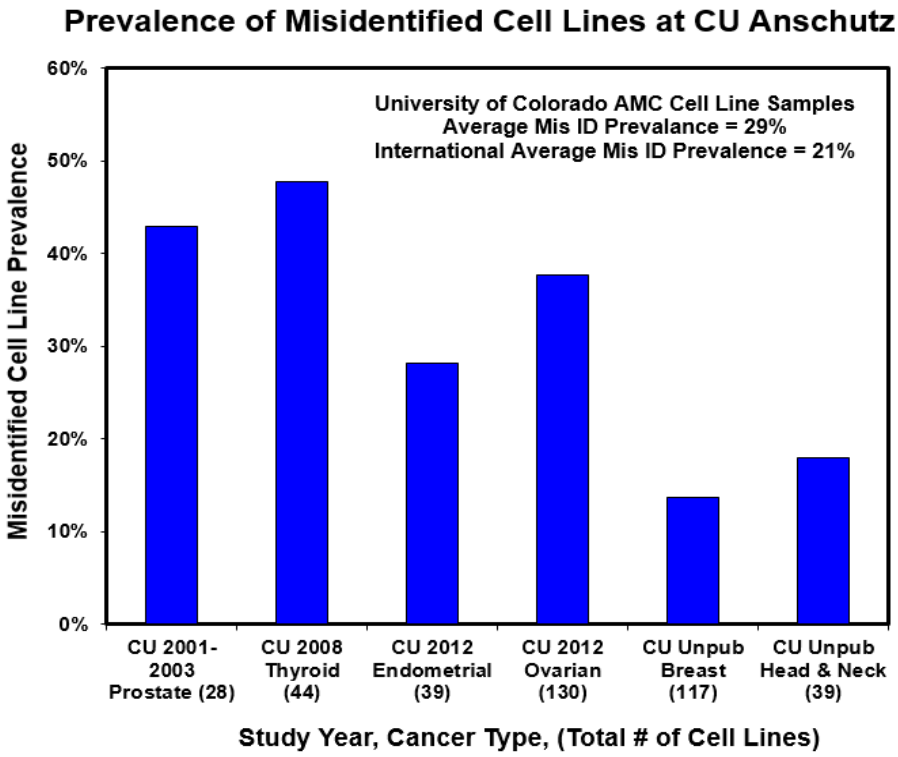
4. New human cell lines or tissue cultures received by a laboratory should be authenticated by STR / DNA profiling and tested for mycoplasma contamination as soon as possible and before being used in the laboratory. Cells should be quarantined (kept in a separate incubator) before use or before being put into an incubator with non-contaminated cell lines. The only exception is if the cell line has been received directly from a repository that authenticates cell lines, and provides both an STR profile made from the same batch of cells, and a certification that they are mycoplasma-free.
5. Newly established human cell lines or tissue cultures need to be authenticated by STR / DNA genotyping of both the cell line and the tissue from which they were derived.
6. Where available, mouse, rat, and dog cell lines and strains may be authenticated by either STR or SNP genotype analysis.
7. A large batch of the cell line should be grown up and aliquoted into standard freezer vials. Before use, one aliquot of the batch should be STR genotyped to confirm the authenticity of the batch of cells. The remainder vials of this seed stock should serve as a reserve for a complete set of experiments and any aliquot should not be passaged more than 20 times.
8. In the case of primary tumor or explants (e.g., PDX models), multiple aliquots of the original tissue sample cell suspension should be stored and one aliquot genotyped to verify its identity matches that of the original patient sample.
9. Cell lines or tissue cultures that have been subjected to biological selections (e.g., after transfection, selection for drug resistance, passaging through animals as xenografts, etc.) should be authenticated by STR / DNA genotyping before further use in the laboratory. For instance, a novel phenotype might be expected, but could be due to contamination.
10. Whenever the cell line appears to respond abnormally or shows unexpected behavior or abnormal morphology during its use in a project, it should be authenticated by STR genotyping before further use, and/or the experiments should be repeated with a new aliquot of the authenticated seed stock.
11. If the cell lines or tissue cultures are part of a grant application or manuscript submission, they must be authenticated before submission of a grant/manuscript (currently a requirement of the NIH and to various extents by multiple journals).
12. Cell lines or tissue cultures distributed to other labs within and outside the University of Colorado should be accompanied with copies of the cell line's STR genotyping data and mycoplasma test results. The recipient should also confirm the STR profile and test for the presence of mycoplasma.
13. All cell lines should be tested for mycoplasma on a regular basis, preferably by PCR-based kits, which are commercially available. This service is also offered through the Barbara Davis Center's BioResources Core Facility- Molecular Biology Unit

In summary, these recommendations for evaluating cell cultures and tissue samples prior to their use, include developing appropriate laboratory protocols to minimize the risk of working with misidentified cell samples, and confirming the authenticity of the cell lines and tissues used in

the experiments before submitting grants, manuscripts, and sharing samples with fellow researchers. The goal of these Best Laboratory Practices is to protect the integrity, validity, and reproducibility of biomedical research produced by members of the UCCC.

## Introduction

Begley [11] reported that a majority (50%-90%) of published biomedical research is not reproducible and Freedman et al. [25] estimated that up to US $28 billion is spent each year on irreproducible biomedical research. A major contributor to the irreproducibility of biomedical research is the infrequent authentication of established cell lines and tissue samples and the frequent use of misidentified cell lines [26, 27, 33, 38, 39, 53, 54, 69]. At the University of Colorado, the UCCC DNA & Sequencing Shared Resource was, in 2001, one of the first core labs to offer cell line authentication in the US and internationally. Between 2001 and 2014, of the samples submitted for analysis, we found the incidence of misidentified human cell lines being used in research ranged between 14% and 48%, with an average of 29% (see adjacent figure; Korch, unpublished data; no data post-closure of this core by UCCC). A similar prevalence of misidentification was seen among PDX samples submitted to this core (Korch, unpublished data, 2019).




Studies of data from labs around the world show the incidence of misidentified cell lines ranges from 10% to 100%, with an average of 22% or 2 of 9 cell lines being incorrect [38]; i.e. lower than the UCCC average. Many cell lines that were shown by Stanley Gartler to be misidentified as early as 1967 [30] are still being used under their false identities [32, 33, 39] to model the incorrectly identified tissues. Vaughan et al. reported that out of 574 articles between 2000 and 2014 using the cell line KB established in 1955, only 57 described the cell line correctly as actually being an imposter derived from HeLa cells [68]. Korch and Capes-Davis reported that the HeLa-derived cell lines Hep-2 and Intestine 407 (originally claimed to be from liver and normal intestinal cells) were used in 8,497 and 1,397 articles, respectively, under their false identities and this usage continues to this day [39]. Horbach and Halffman [33] found that 251 of over 500 known false cell lines with no known authentic samples were used in 32,755 publications and these articles were cited conservatively one-half million times in the scientific

literature. Additionally, 7% or more of human genomic data are sequences derived from mycoplasma, a bacterial genus that frequently contaminates cell line cultures [41] and which can alter metabolic responses and impact the analysis of contaminated cell lines [1, 9, 14, 22-24].

In 2018, the Biocompare company [12, 55] produced an excellent documentary that is worthwhile viewing, in which we and eleven other prominent scientists, who have been combating this cell line authenticity problem, discussed many aspects of this issue. In 2022, the *International Journal of Cancer* (IJC) reported that 22.9% of submitted manuscripts which used human cell lines that had different levels of problems with the identifies of the cell lines used for research (i.e., 5.4% of the manuscripts required minor edits of incomplete cell line identity information; 9.8% had moderate problem due to using, e.g., a false cell line which had to be removed because it was misidentified but which was not a major basis of the study; and 7.6% which used two or more false cell lines and/or falsified data and documentation) [63]. Of all the manuscripts, 4.7% were rejected because of "severe, unaddressed cell line problems," and yet were found subsequently to have been published in other journals without addressing the cell line issues.

Clearly, this major problem continues to contaminate the biomedical literature with unreliable research [32-34, 36-39, 42, 63, 68, 71, 72]. Many have proposed that journals and funding agencies, as gate keepers of the biomedical literature, should require authentication of cell lines for acceptance of manuscripts and grant applications [29, 43, 48, 49, 52, 69]. Some journals and some funders, such as the NIH, have begun to require some level of sample authentication, but most requirements imposed by these gate keepers are not very stringent [63] allowing for continued publication of research based on false cell lines [25, 39, 68].

One approach for dealing with published reports based on using misidentified cell lines is that the articles be retracted. This appears to be impractical and an avoided option. Thousands of publications have used false cell lines [32, 33, 39, 68], but only about 100 articles have been retracted or issued notices of concern or corrigenda (see Appendices 2 and 3 in [38] and search Retraction Watch http://retractiondatabase.org/RetractionSearch.aspx for articles retracted because of contaminated cell line and tissue samples). Retractions of research articles are costly, as illustrated by the work of Linger et al. [44-47], infrequent [38], and can be detrimental for the reputations of researchers and their institutions [64]. Stern et al. estimated the average financial cost of retracting an article is between US$300,000 and US$400,000, not to mention its costs in decreased reputation of the research group and institution and their ability to obtain grants in the future [64]. Potentially, many billions of dollars have been used to support research based on false cell lines [25, 39].

The cross-contamination or misidentification of cell cultures, cell lines, stem cells, and xenografts (both of tissue samples and cell suspensions) can have seriously compromised research using these cells to model diseases, biological phenomena, and drug testing in pre-clinical trials, development and improvement of therapies for metabolic diseases and cancers of specific organs. This ultimately results in the waste of limited research resources [33, 38, 39, 61, 63]. Clearly, researchers need to beware of basing their research on publications using false/imposter and misidentified cell lines and citing publications based on such cell lines.

Six valuable resources have been developed to address this enormous problem: The International Cell Line Authentication Committee's website (ICLAC) [35], Cellosaurus – a Cell Line Knowledge Resource website [8]; and four guidance manuals for the authentication of human and mouse cell lines by analysis of STR genotypes, namely two ANSI-ATCC standards [5, 40] and two Assay Guidance manuals available on the NCBI website [3, 59]. To address to this alarming problem, the NIH issued three notices that "key biological and/or chemical resources" need to be regularly authenticated (NOT-OD-15-103, NOT-OD-16-011, NOT-OD-16-012, effective January 25, 2016). Several granting agencies now require some degree of authentication of key biologic reagents, including cell lines and tissue samples. Increasingly, a number of journals are starting to require that authors provide some kind of evidence that the cell lines that they used were authenticated prior to manuscript acceptance. The IJC has an exemplary cell line policy when considering manuscripts for publication (see [63] and Supplementary Information).

It is imperative that members of the biomedical research community alter their research practices in a manner that recognizes the need for these BLPs so that this important problem can be addressed. Several publications describe "best practices" for the culturing, handling, and authentication of cell lines and tissue samples [4, 5, 10, 28, 31, 50, 58-61, 63]. Building on the *Cell Line Authentication Policy of the MD Anderson Cancer Center* [67] (which is the only publicly available cell line policy we could find), the information on the websites of the International Cell Line Authentication Committee (ICLAC) and Cellosaurus, and our own experience (in both our own research and running the UCCC DNA Sequencing & Analysis core facility), we propose the following cell line and tissue sample authentication Best Laboratory Practices (BLPs) for University of Colorado biomedical research.

## Best Practices for the Authentication Cell Line and Tissue Samples used in Biomedical Research

The UCCC Best Laboratory Practices for Cell Line Authentication recommend that all cell cultures and tissue samples used by any member of a research team should be validated by all four of the following authentication methods. These methods should be applied at the following stages of a project: a) at the beginning of a project before using samples for experiments, b) regularly during handling of samples, c) whenever a novel phenotype is noted, d) after any phenotypic selection, e) at the end of a project, f) before submitting grants and manuscripts, and g) prior to sharing samples with fellow researchers. In addition, Microsatellite Stable (MSS) cell cultures should not be passaged more than 10-20 times and not used when unusual phenotypes are noticed. In such cases, the experiments should be started anew with a fresh aliquot of the authenticated batch of cells. MSI-Unstable cell cultures should not be passaged more than 5-10 times because of their inherently rapid genetic drift.

**Method 1** - Confirming that the cell lines proposed to be used for research are known to be authentic and not misidentified/cross-contaminated lines by verifying their identity

on the Cellosaurus (https://www.cellosaurus.org/) and ICLAC (https://iclac.org/) websites;

**Method 2 -** Genotyping by STR / DNA profile analysis of human cell samples (i.e., cell lines, patient-derived explants (PDX), xenograft tissue) and by STR (mouse, dog, rat) or SNP genotyping of other non-human cell lines to authenticate their genetic identity;

**Method 3 -** Identification of any non-human species present in human culture samples by species-specific PCR analyses (e.g. [19]) or comparable methods to ensure the samples are authentic and not contaminated with cells from inappropriate animal species [6, 19, 34]; and

**Method 4 -** Analysis of all cell culture samples to ensure they are not contaminated with mycoplasma or other microbes, which can affect cell line phenotypes.

All laboratory research team members working with cell lines, xenografts, and tissue samples are encouraged to implement these four methods in their laboratories to ensure that they publish valid and reproducible research. The laboratories may need to identify a service facility that can perform at least Method 2, while the other three methods can be performed in most laboratories. The researchers will need learn how to interpret the data from the four methods as is described in the ANSI-ATCC ASN-00002-2022 Standard [40], the NCBI Assay Guidance Manual [3], and the Match Criteria explanation on the ICLAC website (https://iclac.org/resources/match-criteria-worksheet/).

## Implementation of Guidelines and Methods to Ensure the Authenticity of Cell Lines and Tissue Samples and Adherence to these Best Laboratory Practices

We have developed the following guidelines and methods to encourage implementation and adherence with the BLPs based on the above four authentication methods, These suggestions are derived, in part, from those presented in Appendix G of the 2022 revision of the ANSI-ATCC Standard ASN-0002 describing Human Cell Line Authentication by STR DNA profiling [40] and from other resources. These steps consist of (I) determining which cell lines and tissues are appropriate for the planned experiments, (II) implementing appropriate laboratory protocols to minimize the risk of working with misidentified cell samples, (III) being aware of any warning signs that the cell line and tissue sample are no longer correct, (IV) confirming the authenticity of the cell lines and tissues used in the experiments before submitting grants, manuscripts, and sharing samples with fellow researchers, and (V) considering patient confidentiality when publishing human STR data.

### Disclaimer

Mention of specific commercial equipment, instruments, or materials (or suppliers, software, etc.) in this document is only to foster understanding. It does not imply that we

recommend or endorse any of the materials or equipment identified or imply that these are necessarily the best available for the purpose.

## I. Eight Steps for Evaluating Cell Lines and Tissue Samples Prior to their Use in Research

1. All laboratories should establish mandatory training in the Best Laboratory Practices and safe tissue culturing techniques referenced above to prevent contaminating the established cell lines with cells from foreign cell lines or microbes including mycoplasma. These precautions are described on the ICLAC website and by Capes-Davis and Freshney [15], Freshney [28], Geraghty et al. [31], Korch et al. [40], the 2018 OECD Guidance Document on Good In Vitro Method Practices (GIVIMP) and others [10, 50, 58-60].

   To implement these techniques, each laboratory should develop a set of documents detailing these standard best laboratory practices for the consistent and required uniform training of all laboratory personnel as described in the ANSI-ATCC Standard ASN-0002-2022 [40]. The goal is to minimize inconsistent and irreproducible results arising through variable verbal transmission of laboratory practices. See the attached Supplementary Information document for additional details.

   Laboratories are encouraged to (a) provide new lab personnel with detailed mandatory training on the handling of cell lines and tissue samples and (b) offer all lab personnel annual refresher training on these techniques to reduce the likelihood of sample mix-ups, etc.

2. Before obtaining cell lines for a research project, the names of these cell lines should be checked for what is known about them on the websites of ICLAC [17], Cellosaurus [8], the Cell Line Project at the COSMIC Catalogue Of Somatic Mutations In Cancer [66], and the Biosample database of the NCBI (https://www.ncbi.nlm.nih.gov/biosample, which has STR electropherographic data for some of the included cell line data).

   - Are any of the cell lines known to be misidentified?
   - Is there a reference STR or SNP genotype or other information about the cell lines available for comparison?

   Compile the collected information in a searchable database for future use, including in a spreadsheet (e.g., Excel or .CSV file available from the CLASTR-Cellosaurus website) which would allow searching for potentially matching STR profiles.

3. It is best that established cell lines are obtained only from sources that can demonstrate the authenticity of the material. Commercial sources, such as the ATCC, RIKEN, JCRB, or DSMZ, do this routinely. Most colleagues are not likely or able to provide authenticated cell lines, which were not developed in their labs.

4. Prior to their use, all cell cultures brought into a lab should be incubated in a separate Quarantine Incubator until the cultures are shown to be both free of microbial contamination and not cross-contaminated with other cell lines, especially rapid growers (for example as, but not limited to, HeLa, M14, MCF-7). An exception can be if the cell line was received directly from a reliable repository, such as one of those described above (item 3), with documentation confirming its identity and purity.
5. When establishing new cell lines or xenografts, it is critical to determine the STR genotype of the original patient tissue sample to serve as a reference genotype for all future work. Preserve original tissue samples and, if possible, a normal sample of tissue or blood for any other future molecular characterization of the total (genomic and mitochondrial) DNA.
6. These BLP guidelines should apply also for the use of cell lines from other organisms, i.e., check for interspecies contamination when working with cells from different species, such as human xenografts in mice. When possible, authenticate cell lines and tissues from mouse, rat, dog, and other species using STR or SNP genotyping [2, 6, 8, 13, 19, 21, 51, 56, 65]. New databases and technologies are under continuous development and should be incorporated into the laboratory's best research practices.
7. Cell lines should also be considered potential sources of infectious agents such as pathogenic viruses. This could be due to the original patient sample being infected or could have occurred sometime between the time that the cells were first established and when received by the researcher's laboratory. Therefore, these cultures should be handled with appropriate *Universal Precautions*. The CDC and WHO have published online protocols for the safe handling of human cell cultures [18, 70].
8. Newly established human and mouse cell lines and preclinical models (xenografts, 3D cultures, stem cells, etc.) generated by a laboratory should be subjected to short tandem repeat (STR) genotype analysis and other appropriate molecular techniques, including species confirmation [6, 19] and phenotype, to establish baselines for future authentication efforts.

## II. Most Common Causes of Cell Line Contamination/Misidentification and Recommended Cell Culture Handling Protocols to Avoid these Pitfalls

A. Most Common Causes of Cell Line Contamination/Misidentification:

- Handling multiple lines at the same time in a tissue culture hood since some cells, like those of HeLa, have been shown to "fly" by survival and dispersal in aerosols [20, 55].
- Having two persons working together in the same hood at the same time.
- Using the same bottle of growth medium for multiple cell lines.
- Using the same pipette to apply medium or wash solutions to several different cultures.
- Mislabeling a tissue culture flask or putting the wrong cell suspension in a flask.

- Working with a more aggressive, faster growing cell culture before working with a less aggressive slower growing culture instead of vice versa.
- Not properly cleaning the working surface in the hood between cell lines.
- Storing flasks and equipment in the tissue culture hood and thus preventing complete sanitation of the hood's surface.
  The hood should be cleaned by wiping down with a 10% bleach solution followed by 70% ethanol, and then irradiating it with UV light both before and after a day's work. Note that the UV lights in hoods may not be effective as the light tubes lose their intensity with use and most labs never monitor their intensity over time.
- Not cleaning the tissue culture incubator regularly and not using clean water in the heater jacket of the incubator.

These potential pitfalls should be avoided stringently in order to minimize the chances of cell line mix-ups and cross-contaminations and to preserve the authenticity of the culture samples in one's research. Below are ten recommendations for the handling of cell lines and tissue cultures that should be implemented in all laboratories to avoid using the wrong samples.

## B. Ten Protocols Recommended for Handling Cell and Tissue Cultures

1. Dedicate individual bottles of media to each cell lines to avoid accidental cross-contamination of cell lines.
2. Label media bottles, culture flasks, and storage vials with clear labels or with printed labels before starting to use them to avoid misidentification of samples.
3. Work with only one cell line at any time to avoid mix-ups and cross-contaminations.
4. When handling cell lines, work with the faster growing cultures and new, untested (quarantined) cells last to minimize the likelihood of rapid growers and potentially contaminated cell lines taking over slower growing cultures.
5. Initially, a large batch of a chosen cell line should be generated from which multiple aliquots are preserved in liquid nitrogen. Test an aliquot of this batch to determine its STR or SNP genotype and whether it is contaminated with mycoplasma before the cells are used for research.
   - Does the STR genotype match the published data or the STR genotype of the original donor?
6. Parson et al. [57] and Korch et al. ([37, 38], unpublished observations) have described how the STR genotypes of MSI-Unstable cell lines can change more rapidly than those of MSI-Stable cell lines with passaging. Cell lines can be checked for microsatellite instability (MSI) using tests such as described by Bacher et al. [7], the ANSI-ATCC ASN-0002-2022 Standard [40], or the Promega kit (MD1641, see https://www.promega.com/products/molecular-diagnostics/amplification/microsatellite-instability-msi-analysis/?catNum=MD1641). MSI and genetic drift due to excessive

passaging may provide an explanation for experimental STR genotypes not matching exactly those expected for the cell line(s) and their phenotypes changing.

7. Cell lines routinely used by a laboratory should be authenticated at regular intervals to confirm their identity; a minimum of semi-annual testing is recommended. Minimally. cell lines should be authenticated at the beginning and conclusion of a study so as not to jeopardize the submission of manuscripts and grants and the reproducibility of the findings. The best testing regime would be at the following time points:

   a) The start of the project before using cell lines for experiments,
   b) Regularly during a project when handling samples / culturing cells,
   c) Whenever a novel phenotypic behavior is noticed (see item 1 in section III),
   d) After phenotypic selection (e.g., drug resistance, growth as xenografts; see item 2 in section III),
   e) At the end of a project,
   f) Before submission of grant applications,
   g) Before manuscript submissions, and
   h) Prior to sharing of samples with fellow researchers.

   It can be useful to frequently save aliquots of the residual cell suspensions after passaging cells and to store them frozen at -20°C (viability is not necessary). If future analysis detects cross-contamination or sample mix-ups at one of the above steps a through g, DNA from these frozen samples can be extracted and analyzed to ascertain when cross-contaminations or mix-ups may have occurred. This could indicate from which stage the experiments should be repeated.

8. Generally, MSI-Stable cell lines should not be passaged more than 10-20 times and MSI-Unstable cell lines should not be passaged more than 5-10 times to minimize the effects of genetic drift of these cultures.

9. The level of confluence of cells can affect gene expression, protein expression and phosphorylation, and the localization of proteins in cells. Therefore, one should harvest cells at consistent levels of confluence to maintain reproducibility.

10. Media components are variable between suppliers and between batches from the same supplier [62]. Therefore, each batch of culture media and their components (e.g., serum) should be recorded and checked, if necessary, for suitability before being used so that results are reproducible.

## III. Three Possible Warning Signs of Undesirable Changes in Cell Line Cultures

Below are three common signs that a cell line culture has been infected or replaced by a population of undesired intra- or inter-species cells.

1. Sudden changes in cell line behavior (e.g., grows faster, change of morphology, or other phenotype) often indicates contamination, an outgrowth of a variant subline, or even the presence of another species. Authenticate such cell lines to verify their identity and purity.

2. The identity of cell lines subjected to selection after modification (e.g., stable transfection and drug selection) should be verified as this process can lead to outgrowth of variants or cross-contaminating "imposter" or hardier cells.
3. Langdon [41] showed that about 7% of DNA sequences from the Human 1000-Genomes Project are from mycoplasma, common human parasitic or commensal bacteria, which commonly contaminate cell cultures. Mycoplasma are known to alter metabolic responses and impact DNA and RNA analysis of contaminated cell line cultures [1, 9, 14, 22-24]. Therefore, all cell lines should be routinely checked for mycoplasma infection and treated accordingly when necessary.

   Cultures can be tested for the presence of mycoplasma with a PCR assay (e.g., Bulldog-Bio eMyco Plus kit) or a luciferase luminescence-based assay (e.g., MycoAlert™ PLUS Mycoplasma Detection Kit Catalog # LT07-703).

## IV. Analytical Steps in the Cell Line and Tissue Authentication Process

During the use of cell lines and prior to publishing research results, the frequency of STR genotyping and percent match to the reference sample should be reported using the "Tanabe" algorithm [16] as explained elsewhere [3, 40]. Investigators should understand that cell lines or tissue samples having identical or nearly identical STR- or SNP-based genotypes does not prove that the samples are genetically identical. It only shows that they were originally derived from the same donor.

Any investigator sharing cell lines with fellow researchers, has the ethical responsibility to provide only authenticated cell lines. Results of STR genomic profiling and mycoplasma testing should accompany the cell line transferred to the new investigator.

All laboratories should ensure they comply with the cell line authentication policies of the granting agencies that support their research. If policies from multiple agencies apply, the most stringent regulations should be followed.

## V. Considerations of Patient Confidentiality when Publishing Human STR Data

When establishing institutional cell line and tissue authentication guidelines as described herein and deciding to publish human STR genotyping data, patient confidentiality and genetic privacy may need to be considered. The human STR genotyping assays were developed for identification of forensic samples, with newer kits detecting the alleles in 25 or more STR loci. These loci are in noncoding chromosomal regions that are of minimal prognostic or diagnostic value. This expansion of available genetic data increases the potential for linking cell lines with deleterious germline mutations to specific patients and familial relatives, but only if the donor samples are not anonymous or as in forensic cases where individuals have their STR genotypes in a forensic database.

To address this concern, the M.D. Anderson Cancer Center **(**see Supplemental Information**)** and the Japanese Protection of Personal Information Act require that the alleles at only eight STR loci be published as opposed to alleles at thirteen STR loci recommended by the 2022 ANSI-ATCC standard. Genetic privacy and patient confidentiality regulations in the USA, Europe, Australia seem not to specify whether STR genotypes may be published. Journals do not specify a maximum limit of the number STR loci data. The International Journal of Cancer (IJC) requests data for a minimum of 8 STR loci. However, there are already 8,327 human STR genotypes in Cellosaurus (as of November 2022), which were obtained from published and unpublished sources of cell lines from anonymous donors and can include data for up to 31 STR loci, with most profiles containing data for between 8 and 17 loci. Souren et al. [63] argue for one limitation on STR data and that is that the actual electropherograms should not be published, but only be provided to the editor and reviewers for their use. This is because the IJC received several manuscripts in which the authors had copy-and-pasted electropherograms from articles published by other authors. Greater details of this issue are presented in the Supplementary Information.

Therefore, we recommend all scientists and research entities consider the above when establishing their own guidelines regulating the release of human STR results and comply with their local and national regulations that govern the publication of these data.

### VI. Supplementary Guidelines, Recommendations, and Useful Information

A printout of the ***Cell Line Authentication Flow Chart*** below can be laminated and posted in tissue culture laboratories as a reminder to all of the critical steps for ensuring authenticated, valid, and reproducible results. Additional sets of recommendations and useful information are described in the attached Supplementary Information file.

## Definitions

The following definitions of three frequently used terms when discussing the authenticity cell lines are taken from the ICLAC website.

### Authentication

The aim of authentication is to confirm or verify the identity of a cell line, demonstrating that it is derived from the correct species and donor. Testing involves comparison of a test sample to other reference samples from that donor, or to a database of reference samples if donor material is not available, to see whether their genotypes correspond. Ideally, the test method should distinguish between different species and different individuals within that species, although this will depend on the technology available to the field of authentication testing.

Not all currently used test methods have the power of discrimination of STR profiling or SNP testing; therefore, authentication may not in all cases lead to unambiguous identification of

cells to a specific donor. Where unambiguous identification is not possible, species verification using methods such as mitochondrial CO1 barcoding [6] or species-specific PCR of mitochondrial cytochrome B [19], is used as the best alternative currently available.

### Cross-contamination

The term contamination refers to introduction of foreign material into a cell culture. Cross-contamination occurs when that foreign material consists of cells from another culture, either human cells or non-human cells arising from species such as mouse or rat. Cross-contamination initially results in a mixed culture, containing cells from the authentic culture and the contaminant. If the contaminant has a survival advantage – for example, if it proliferates more rapidly – it will overgrow and replace the authentic cells within the culture. A contaminant usually comes from a different donor or species and so can be detected by authentication testing.

### Misidentification

A misidentified cell line no longer corresponds to the donor or species from which it was originally established. Misidentification may arise due to cross-contamination. It may also arise from a variety of errors, including mislabeling of samples. If it happens early – for example, during cell line establishment – there will be no authentic material retained, and the cell line is considered to be a falsely identified or misidentified cell line. If misidentification happens late – for example, after the cell line is established and distributed to other locations – then authentic material may still exist and only some stocks may be false.

Misidentification does not refer to problems with the technical procedure of authenticating cell lines. It also does not typically extend to other characteristics such as tissue type, cell type, or disease state. If the tissue type, cell type, or disease state of a cell line is incorrectly attributed, the cell line is considered to be misclassified.

## Contact Information

The authors can be contacted at the following addresses:
Christopher.Korch@CUAnschutz.edu and Rebecca.Schweppe@CUAnschutz.edu.

## References

1. **Aldecoa-Otalora E, Langdon W, Cunningham P, Arno MJ.** 2009. Unexpected presence of mycoplasma probes on human microarrays. Biotechniques **47:**1013-1015. PMID: 20047202
2. **Almeida JL, Hill CR, Cole KD.** 2014. Mouse cell line authentication. Cytotechnology **66:**133-147. PMID: 23430347
3. **Almeida JL, Korch CT.** 2022. Authentication of Human and Mouse Cell Lines by Short Tandem Repeat (STR) DNA Genotype Analysis. *In* Sittampalam GS GA, Brimacombe K, et al. (ed), Assay Guidance Manual [Internet]. National Center for Advancing Translational Sciences; 2004-. Bethesda, MD.

4. **American Type Culture Collection Standards Development Organization Workgroup ASN-0002.** 2010. Cell line misidentification: the beginning of the end. Nat Rev Cancer **10:**441-448. PMID: 20448633
5. **ATCC SDO.** 2011. ASN-0002. Authentication of Human Cell Lines: Standardization of STR Profiling. ANSI eStandards Store, *on* ATCC-Standards Development Organization (SDO). https://webstore.ansi.org/standards/atcc/ansiatccasn00022011. Accessed September, 2021.
6. **ATCC SDO.** 2015. ASN-0003 Species-level identification of animal cells through mitochondrial cytochrome c oxidase subunit 1 (CO1) DNA barcodes, *on* ATCC-Standards Development Organization (SDO). https://webstore.ansi.org/Standards/ATCC/ANSIATCCASN00032015. Accessed
7. **Bacher JW, Sievers CK, Albrecht DM, Grimes IC, Weiss JM, Matkowskyj KA, Agni RM, Vyazunova I, Clipson L, Storts DR, Thliveris AT, Halberg RB.** 2015. Improved detection of microsatellite instability in early colorectal lesions. PLoS One **10:**e0132727. PMID: 26252492
8. **Bairoch A.** 2018. The Cellosaurus, a cell-line knowledge resource. J Biomol Tech **29:**25-38. PMID: 29805321
9. **Barile MF.** 1968. Mycoplasma and cell cultures. Natl Cancer Inst Monogr **29:**201-204. PMID: 4976863
10. **Baust JM, Buehring GC, Campbell L, Elmore E, Harbell JW, Nims RW, Price P, Reid YA, Simione F.** 2017. Best practices in cell culture: an overview. In Vitro Cell Dev Biol Anim **53:**669-672. PMID: 28808859
11. **Begley CG.** 2013. Six red flags for suspect work. Nature **497:**433-434. PMID: 23698428
12. **Biocompare.** 2018. Reproducibility issues in life science research: The importance of cell line authentication, First ed, vol 2022, p Video interview of 13 scientists working on the issue of cell line authentication., San Francisco, CA 94080, USA. PMID:
13. **Callari M, Batra AS, Batra RN, Sammut SJ, Greenwood W, Clifford H, Hercus C, Chin SF, Bruna A, Rueda OM, Caldas C.** 2018. Computational approach to discriminate human and mouse sequences in patient-derived tumour xenografts. BMC Genomics **19:**19. PMID: 29304755
14. **Callaway E.** 2014. Contamination hits cell work. Nature **511:**518. PMID: 25079533
15. **Capes-Davis A, Freshney RI.** 2021. Freshney's Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications, 8th Edition, Eighth ed. Wiley-Blackwell, Hoboken, NJ, USA.
16. **Capes-Davis A, Reid YA, Kline MC, Storts DR, Strauss E, Dirks WG, Drexler HG, MacLeod RA, Sykes G, Kohara A, Nakamura Y, Elmore E, Nims RW, Alston-Roberts C, Barallon R, Los GV, Nardone RM, Price PJ, Steuer A, Thomson J, Masters JR, Kerrigan L.** 2013. Match criteria for human cell line authentication: where do we draw the line? Int J Cancer **132:**2510-2519. PMID: 23136038
17. **Capes-Davis A, Theodosopoulos G, Atkin I, Drexler HG, Kohara A, MacLeod RA, Masters JR, Nakamura Y, Reid YA, Reddel RR, Freshney RI.** 2010. Check your cultures! A list of cross-contaminated or misidentified cell lines. Int J Cancer **127:**1-8. PMID: 20143388
18. **Center for Disease Control and Prevention C.** 2009. CDC Biosafety in Microbiological and Biomedical Laboratories (BMBL), 5th edition *on* USA Department

of Human Health and Services. https://www.cdc.gov/labs/BMBL.html). Accessed May 2019.
19. **Cooper JK, Sykes G, King S, Cottrill K, Ivanova NV, Hanner R, Ikonomi P.** 2007. Species identification in cell culture: a two-pronged molecular approach. In Vitro Cell Dev Biol Anim **43:**344-351. PMID: 17934781
20. **Coriell LL.** 1962. Detection and elimination of contaminating organisms. Natl Cancer Inst Monogr **7:**33-53. PMID: 13881339
21. **Didion JP, Buus RJ, Naghashfar Z, Threadgill DW, Morse HC, 3rd, de Villena FP.** 2014. SNP array profiling of mouse cell lines identifies their strains of origin and reveals cross-contamination and widespread aneuploidy. BMC Genomics **15:**847. PMID: 25277546
22. **Drexler HG, Dirks WG, MacLeod RA, Uphoff CC.** 2017. False and mycoplasma-contaminated leukemia-lymphoma cell lines: time for a reappraisal. Int J Cancer **140:**1209-1214. PMID: 27870004
23. **Drexler HG, Uphoff CC.** 2002. Mycoplasma contamination of cell cultures: Incidence, sources, effects, detection, elimination, prevention. Cytotechnology **39:**75-90. PMID: 19003295
24. **Drexler HG, Uphoff CC, Dirks WG, MacLeod RA.** 2002. Mix-ups and mycoplasma: the enemies within. Leuk Res **26:**329-333. PMID: 11839374
25. **Freedman LP, Cockburn IM, Simcoe TS.** 2015. The Economics of Reproducibility in Preclinical Research. PLoS Biol **13:**e1002165. PMID: 26057340
26. **Freedman LP, Gibson MC, Ethier SP, Soule HR, Neve RM, Reid YA.** 2015. Reproducibility: changing the policies and culture of cell line authentication. Nat Methods **12:**493-497. PMID: 26020501
27. **Freedman LP, Gibson MC, Wisman R, Ethier SP, Soule HR, Reid YA, Neve RM.** 2015. The culture of cell culture practices and authentication--Results from a 2015 Survey. Biotechniques **59:**189-190, 192. PMID: 26458546
28. **Freshney RI.** 2016. Culture of animal cells: a manual of basic technique and specialized applications, 7th ed. Wiley-Blackwell, Hoboken, New Jersey, U. S. A.
29. **Fusenig NE, Capes-Davis A, Bianchini F, Sundell S, Lichter P.** 2017. The need for a worldwide consensus for cell line authentication: Experience implementing a mandatory requirement at the International Journal of Cancer. PLoS Biol **15:**e2001438. PMID: 28414712
30. **Gartler SM.** 1967. Genetic markers as tracers in cell culture. Natl Cancer Inst Monogr **26:**167-195. PMID: 4864103
31. **Geraghty RJ, Capes-Davis A, Davis JM, Downward J, Freshney RI, Knezevic I, Lovell-Badge R, Masters JR, Meredith J, Stacey GN, Thraves P, Vias M.** 2014. Guidelines for the use of cell lines in biomedical research. Br J Cancer **111:**1021-1046. PMID: 25117809
32. **Gorphe P.** 2019. A comprehensive review of Hep-2 cell line in translational research for laryngeal cancer. Am J Cancer Res **9:**644-649. PMID: 31105993
33. **Horbach SPKM, Halffman W.** 2017. The ghosts of HeLa: How cell line misidentification contaminates the scientific literature. PLoS One **12:**e0186281. PMID: 29023500

34. **Huang Y, Liu Y, Zheng C, Shen C.** 2017. Investigation of Cross-Contamination and Misidentification of 278 Widely Used Tumor Cell Lines. PLoS One **12:**e0170384. PMID: 28107433
35. **International Cell Line Authentication Committee (ICLAC).** 14 October 2018 2018. Database of cross-contaminated or misidentified cell lines. http://iclac.org/databases/cross-contaminations/. Accessed 09 December 2018.
36. **Korch C, Hall EM, Dirks WG, Ewing M, Faries M, Varella-Garcia M, Robinson S, Storts D, Turner JA, Wang Y, Burnett EC, Healy L, Kniss D, Neve RM, Nims RW, Reid YA, Robinson WA, Capes-Davis A.** 2018. Authentication of M14 melanoma cell line proves misidentification of MDA-MB-435 breast cancer cell line. Int J Cancer **142:**561-572. PMID: 28940260
37. **Korch C, Spillman MA, Jackson TA, Jacobsen BM, Murphy SK, Lessey BA, Jordan VC, Bradford AP.** 2012. DNA profiling analysis of endometrial and ovarian cell lines reveals misidentification, redundancy and contamination. Gynecol Oncol **127:**241-248. PMID: 22710073
38. **Korch C, Varella-Garcia M.** 2018. Tackling the Human Cell Line and Tissue Misidentification Problem Is Needed for Reproducible Biomedical Research. Adv Mol Pathol **1:**209-228.e236. PMID: https://doi.org/10.1016/j.yamp.2018.07.003
39. **Korch CT, Capes-Davis A.** 2021. The Extensive and Expensive Impacts of HEp-2 [HeLa], Intestine 407 [HeLa], and Other False Cell Lines in Journal Publications. SLAS Discov **26:**1268-1279. PMID: 34697958
40. **Korch CT, Hall EM, Dirks WG, Sykes GR, Capes-Davis A, Butler JM, Neve RM, Nims RW, Storts DR, Tian F, Nardone RM.** 2022. Human Cell Line Authentication. Standardization of Short Tandem Repeat (STR) Profiling. ASN-0002 Revised 2022, November 2022 ed. American National Standards Institute (ANSI) - American Type Culture Collection (ATCC) Standards Development Organization, Manassas, Virginia, United States.
41. **Langdon WB.** 2014. Mycoplasma contamination in the 1000 Genomes Project. BioData Min **7:**3. PMID: 24872843
42. **Liang-Chu MM, Yu M, Haverty PM, Koeman J, Ziegle J, Lee M, Bourgon R, Neve RM.** 2015. Human biosample authentication using the high-throughput, cost-effective SNPtrace(TM) system. PLoS One **10:**e0116218. PMID: 25714623
43. **Lichter P, Allgayer H, Bartsch H, Fusenig N, Hemminki K, von Knebel Doeberitz M, Kyewski B, Miller AB, zur Hausen H.** 2010. Obligation for cell line authentication: appeal for concerted action. Int J Cancer **126:**1. PMID: 19882693
44. **Linger RM, DeRyckere D, Brandao L, Sawczyn KK, Jacobsen KM, Liang X, Keating AK, Graham DK.** 2009. Mer receptor tyrosine kinase is a novel therapeutic target in pediatric B-cell acute lymphoblastic leukemia. Blood **114:**2678-2687. PMID: 19643988
45. **Linger RM, Graham DK.** 2014. A Common Case of Cell Line Misidentification. https://thewinnower.com/discussions/9-a-common-case-of-cell-line-misidentification. Accessed
46. **Linger RM, Lee-Sherick AB, DeRyckere D, Cohen RA, Jacobsen KM, McGranahan A, Brandao LN, Winges A, Sawczyn KK, Liang X, Keating AK, Tan AC, Earp HS, Graham DK.** 2013. Mer receptor tyrosine kinase is a therapeutic target in pre-B-cell acute lymphoblastic leukemia. Blood **122:**1599-1609. PMID: 23861246

47. **Linger RMA, DeRyckere D, Brandao L, Sawczyn KK, Jacobsen KM, Liang X, Keating AK, Graham DK.** 2012. Retraction. Linger RMA, DeRyckere D, Brandao L, Sawczyn KK, Jacobsen KM, Liang X, Keating AK, Graham DK. Mer receptor tyrosine kinase is a novel therapeutic target in pediatric B-cell acute lymphoblastic leukemia. Blood. 2009;114(13):2678-2687. Blood **120:**1533. PMID: 22753864
48. **Masters JR.** 2002. False cell lines: The problem and a solution. Cytotechnology **39:**69-74. PMID: 19003294
49. **Masters JR.** 2002. HeLa cells 50 years on: the good, the bad and the ugly. Nat Rev Cancer **2:**315-319. PMID: 12001993
50. **McLaren RS, Reid Y, Storts DR.** 2013. Human cell line authentication: the critical first step in any project using human cell lines. Methods Mol Biol **963:**341-353. PMID: 23296621
51. **Morgan AP, Fu CP, Kao CY, Welsh CE, Didion JP, Yadgary L, Hyacinth L, Ferris MT, Bell TA, Miller DR, Giusti-Rodriguez P, Nonneman RJ, Cook KD, Whitmire JK, Gralinski LE, Keller M, Attie AD, Churchill GA, Petkov P, Sullivan PF, Brennan JR, McMillan L, Pardo-Manuel de Villena F.** 2015. The Mouse Universal Genotyping Array: From Substrains to Subspecies. G3 (Bethesda) **6:**263-279. PMID: 26684931
52. **Nardone RM.** 2007. Eradication of cross-contaminated cell lines: a call for action. Cell Biol Toxicol **23:**367-372. PMID: 17522957
53. **Neimark J.** 2014. Trial and Error - The Dirty Little Secret of Cancer Research. Discover Magazine**:**44-51. PMID:
54. **Neimark J.** 2015. Line of attack. Science **347:**938-940. PMID: 25722392
55. **Oliver T.** June 11, 2018 2018. Combating Misidentified and Contaminated Cell Lines, *on* Biocompare. https://www.biocompare.com/Editorial-Articles/348977-Authenticate-Combating-Misidentified-and-Contaminated-Cell-Lines/. Accessed September 23, 2022.
56. **Ono K, Satoh M, Yoshida T, Ozawa Y, Kohara A, Takeuchi M, Mizusawa H, Sawada H.** 2007. Species identification of animal cells by nested PCR targeted to mitochondrial DNA. In Vitro Cell Dev Biol Anim **43:**168-175. PMID: 17516125
57. **Parson W, Kirchebner R, Muhlmann R, Renner K, Kofler A, Schmidt S, Kofler R.** 2005. Cancer cell line identification by short tandem repeat profiling: power and limitations. FASEB J **19:**434-436. PMID: 15637111
58. **Price PJ.** 2017. Best practices for media selection for mammalian cells. In Vitro Cell Dev Biol Anim **53:**673-681. PMID: 28726187
59. **Reid Y, Storts D, Riss T, Minor L.** 2013. Authentication of Human Cell Lines by STR DNA Profiling Analysis. https://www.ncbi.nlm.nih.gov/books/NBK144066/. Accessed
60. **Reid YA.** 2017. Best practices for naming, receiving, and managing cells in culture. In Vitro Cell Dev Biol Anim **53:**761-774. PMID: 28986713
61. **Schweppe RE, Korch C.** 2018. Challenges and advances in the development of cell lines and xenografts. Adv Mol Pathol **1:**239-251. PMID:
62. **Sikora MJ, Johnson MD, Lee AV, Oesterreich S.** 2016. Endocrine response phenotypes are altered by charcoal-stripped serum variability. Endocrinology **157:**3760-3766. PMID: 27459541
63. **Souren NY, Fusenig NE, Heck S, Dirks WG, Capes-Davis A, Bianchini F, Plass C.** 2022. Cell line authentication: a necessity for reproducible biomedical research. Embo j **41:**e111307. PMID: 35758134

64. **Stern AM, Casadevall A, Steen RG, Fang FC.** 2014. Financial costs and personal consequences of research misconduct resulting in retracted publications. Elife **3:**e02956. PMID: 25124673
65. **Steube KG, Koelz AL, Drexler HG.** 2008. Identification and verification of rodent cell lines by polymerase chain reaction. Cytotechnology **56:**49-56. PMID: 19002841
66. **Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, Fish P, Harsha B, Hathaway C, Jupe SC, Kok CY, Noble K, Ponting L, Ramshaw CC, Rye CE, Speedy HE, Stefancsik R, Thompson SL, Wang S, Ward S, Campbell PJ, Forbes SA.** 2019. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res **47:**D941-d947. PMID: 30371878
67. **University of Texas MACC.** 2013. Cell Line Authentication Policy UTMDACC Institutional Policy # ACA1040. https://www.mdanderson.org/documents/core-facilities/Characterized%20Cell%20Line%20Core%20Facility/CCLC_%20Policy_ACA1044.pdf. Accessed
68. **Vaughan L, Glanzel W, Korch C, Capes-Davis A.** 2017. Widespread use of misidentified cell line KB (HeLa): Incorrect attribution and its impact revealed through mining the scientific literature. Cancer Res **77:**2784-2788. PMID: 28455420
69. **Vaught J, Korch CT.** 2021. The Continuing Saga of Cell Line Misidentification. Biopreserv Biobank **19:**357-358. PMID: 34637628
70. **World Health Organization WHO.** 2004. Laboratory Biosafety Manual, 3rd edition, *on* World Health Organization (WHO). https://www.who.int/csr/resources/publications/biosafety/WHO_CDS_CSR_LYO_2004_11/en/. Accessed May 2019.
71. **Ye F, Chen C, Qin J, Liu J, Zheng C.** 2015. Genetic profiling reveals an alarming rate of cross-contamination among human cell lines used in China. FASEB J **29:**4268-4272. PMID: 26116706
72. **Yu M, Selvaraj SK, Liang-Chu MM, Aghajani S, Busse M, Yuan J, Lee G, Peale F, Klijn C, Bourgon R, Kaminker JS, Neve RM.** 2015. A resource for cell line authentication, annotation and quality control. Nature **520:**307-311. PMID: 25877200

# Cell Line Authentication Flow Chart