



INCREASING SCIENTIFIC REPRODUCIBILITY THROUGH CELL LINE AUTHENTICATION

WHAT IS THE ISSUE?

Cell lines provide an invaluable foundation for biomedical research and are selected because of their cell/tissue type or disease state specificity. Use of cell lines enables researchers to (1) interrogate mechanistic pathways based on cell type specificity, (2) investigate biological processes relevant to cell type or disease state, and (3) develop disease relevant models essential to drug discovery and development. Thus, it is imperative that these studies provide consistent, reproducible datasets that show that the cells used are what they are thought to be. When cell lines fail to behave as expected or yield data that is irreproducible, misidentification and contamination are often likely causes.

BACKGROUND

Science has long depended on the ability of peer researchers to independently reproduce experimental findings. Indeed, the retention of experimental data is codified in federal and institutional policies, e.g., the National Institutes of Health (NIH) generally requires that data be retained for three years following the submission of the final report post funding. Some government programs and several scientific journals require retention for up to seven years.

Cell line misidentification and cross-contamination have long been recognized as significant issues in biomedical research [1-3]. Originally identified as a significant problem in 1967 when Dr. Stanley Gartler discovered cross-contamination and misidentification of several cancer cell lines with HeLa by isoenzyme profiling [4], cell line identity is now coming to the attention of funding agencies and publishers. Misidentification of cell lines negatively impacts scientific progress in many ways, from rendering subsequent research disputable to wasteful consumption of already finite research resources [5, 6]. Inter- and intra-species cross-contaminations have been documented in 18% to 36% of cell lines stored in various biobanks worldwide [7, 8]. HeLa cells account for much of the cross-contamination that occurs because the aggressive proliferative capability of these cells allows them to overtake and replace a cell line

they have contaminated. The International Cell Line Authentication Committee (ICLAC), which was formed to highlight the issue and promote adoption of authentication in biomedical research practice, cites that 23% of the cell lines listed on their website (<http://iclac.org/>) are contaminated by HeLa cells. One high profile example of cell line cross-contamination involves the multidrug-resistant MCF-7 breast adenocarcinoma line (originally named MCF-7/AdrR and later re-designated NCI/ADR-RES), which served as an important and widely used research tool during the last two decades. The real identity of these cells has been in doubt since 1998. The origin of NCI/ADR-RES cells has now been revealed by single nucleotide polymorphisms (SNP) and karyotype analyses, carried out at the Sanger Institute and the National Cancer Institute (NCI), respectively. The results of these published analyses demonstrate that NCI/ADR-RES cells are in fact derived from OVCAR-8 ovarian adenocarcinoma cells [5]. The case of NCI/ADR-RES cells highlights the widespread problem of cell line cross-contamination and misidentification. The use of this misidentified line has resulted in countless hours of wasted research time, approximately \$100 million in wasted research funding and the generation of irrelevant and irreproducible data. Fortunately, this is a tractable problem that can be avoided by scrupulous authentication of cell stocks and adoption of a few simple rules in cell culture practice.

A well-known study conducted DNA fingerprinting using standard DNA Short Tandem Repeat (STR) profiling and identified contaminated cells from the renowned NCI-60 panel, which was used for over 25 years in various drug-screening studies [6]. Several lines from this panel that were previously assumed to be disparate lines were found to share common origins. As a result, this panel was retired and will be replaced with cancer cell lines that are better characterized.

In 2007, Dr. Roland Nardone, professor emeritus at The Catholic University of America,



INCREASING SCIENTIFIC REPRODUCIBILITY THROUGH CELL LINE AUTHENTICATION

supported by a group of cell culture experts, published an open letter to Secretary of the United States Department of Health and Human Services as a call for action at the national level to combat the use of misidentified and cross-contaminated cell lines in biomedical research.

In response to Dr. Nardone's letter, the NIH released a notice regarding cell line authentication, which urged manuscript and grant reviewers to scrutinize cell line authenticity [9]. In June 2015, the NIH released a notice regarding authentication of key biological resources required for grant applications [10]. Many publications also now require evidence of cell line authentication. For example, the journal *Nature* requires that authors provide cell line authenticity information for lines used within a submitted manuscript which includes: the method used, the results of authentication testing, and the results of mycoplasma contamination testing. As of May 2015, these requirements are particularly emphasized for cancer research where the issue of cell line misidentification has been well documented; however, authors in all disciplines are strongly encouraged to comply with these reporting criteria [11].

Cell line authentication

Correct cell line identification and authentication can be a relatively simple step in one's research protocol. Authentication can be achieved by determining the genetic signature of a cell line by STR profiling or DNA fingerprinting and comparing the results with established databases such as the BioSample (NCBI), ATCC, or ICLAC STR databases to confirm identity. STRs are tracts of repetitive DNA ranging from 3-7 base pairs in length and are repeated 5 to 50 times in the genome. These repeats are well distributed throughout the human genome and are a rich source of highly polymorphic markers, which can be detected using the polymerase chain reaction (PCR). Alleles of the STR loci are distinguished by the length of a PCR product across the region which demonstrates the number of copies of the repeat sequence. These products are size fractionated

and are further distinguished from one another using fluorescence detection. The biorepositories at the Coriell Institute for Medical Research (Coriell) in Camden, New Jersey, have been using an assay based on this principle for almost 20 years, using 6 STR loci to establish the identity between blood samples or skin fibroblast cells and DNA, lymphoblastoid cell lines (LCLs) and induced pluripotent stem cells (iPSCs) derived from these primary samples. There are now several commercially available kits which use STRs that overlap with the newly published standards from the American National Standards Institute (ANSI) [12]. These kits allow users to search for an identity match against STR profiles that have been deposited in various databases. Coriell has adopted use of these ANSI standards for STR-based cell ID assays. It is important to note that not all STR profiles are deposited into open access databases, as doing so could potentially result in the public release of personally identifying genetic information, which samples may not have been specifically consented to allow.

Single nucleotide polymorphisms (SNPs) are genetic variations between members of the same species. SNPs at multiple loci can also be used as a genetic test for identity. SNP typing assays have been published for forensic and cell line authentication applications. A 52-plex SNP assay has the same rate of discrimination as a 16-plex STR profile. Although there are commercially available kits, widespread adoption of SNP assays for cell line authentication is not yet prevalent due to lack of a centralized online reference database. These assays have potential to become mainstream in the future.

Microbial contamination detection

Another prevalent problem for cultured cell lines is microbial contamination with mold, yeast, bacteria and/or mycoplasma. While contaminations with mold, yeast and bacteria are easily detectable, both macro- and microscopically, mycoplasma contamination is not. Mycoplasma are the smallest bacteria yet



INCREASING SCIENTIFIC REPRODUCIBILITY THROUGH CELL LINE AUTHENTICATION

discovered and lack an outer cell wall making them visually undetectable and impervious to many antibiotics. These small microbes can also be easily aerosolized and can survive for long periods on many laboratory surfaces waiting to infect the next open vessel. As a result, mycoplasma contamination rates in the research community remain high. Although they cannot be seen, mycoplasma can cause significant changes to a host cell functions, affecting growth, morphology, metabolism, and/or expression profiles [13, 14]. Rates of mycoplasma contamination in continuously cultured cell lines within the U.S. have been estimated at 11-15% and have been found to be even greater in Europe (25%) and up to 80% in Japan [15, 16]. Traditional culture/growth assays for mycoplasma fail to support growth of all known mycoplasma species and hence the recently available commercial assays based upon detection using real-time PCR are extremely sensitive and reliable. For example, the MycoAlert™ kit from Lonza can detect up to 95 different mycoplasma species, as low as 1 copy per genome.

THE CORIELL SOLUTION

Coriell is committed to providing the scientific community with the highest quality biospecimens, which are necessary for successful research endeavors. For more than 60 years, Coriell has been a trusted source of cell lines and DNA for global biomedical researchers. Coriell is uniquely qualified to lead the field in addressing the scientific community's need to authenticate cell lines and ensure cell line integrity. Coriell offers a range of research services, performed under strict quality systems and to the ISO 9001:2008 standard, that have long supported national and international science.

DNA fingerprinting

Initially developed to provide identification of DNA generated from blood samples or cell culture throughout the process of purification or amplification, Coriell's STR assay has been in place for almost 20 years and has been performed on tens of thousands of samples.

This assay involves the PCR amplification of 6 highly variable STR regions in the human genome, plus the amelogenin alleles on the X and the Y chromosomes for sex determination. This assay is performed when a blood sample or cell line is first received, and again when DNA purification is complete or when a cell line is established or expanded. The size of the PCR product determines the allele for each of these STR regions and allows a unique identification for each blood sample and cell line. Coriell also uses the more recently released AmpFLSTR® Identifiler® cell line authentication systems, which utilizes 16 STR markers and has a random probability match of 1 in 3.65×10^{17} , for further discrimination and is incorporating this as standard for all new submissions. The use of both systems in parallel allows Coriell to refer to our in-house database established with the 6-plex assay, while building a new database with ANSI standard STRs which can be used to query external databases and provide customers with much needed cell line authentication services.

Additional quality control

In addition to cell line authentication, cell cultures at Coriell are grown and frozen in antibiotic-free media to aid in the detection and prevention of contamination. Cells and culture medium are tested both at the freeze pool and upon recovery from freeze for contamination with trypticase soy broth, Sabouraud's dextrose broth, tryptose phosphate broth, and on blood agar plates, and incubated at 30°C (for mold) and 37°C (for bacterial contaminants) for a minimum of 2 weeks. Externally submitted cultures are handled in an isolated biosafety cabinet and tested for contamination upon receipt/recovery. Coriell also performs an additional, highly sensitive, real-time quantitative PCR for bacterial sequences, using Zymo Research's Femto Bacterial DNA Quantification Kit, which can detect as little as 20 femtograms of bacterial DNA. Further, cell lines are tested for mycoplasma contamination at freeze recovery via a sensitive real-time quantitative PCR assay. Externally submitted cultures and cultures established from tissue biopsy material



INCREASING SCIENTIFIC REPRODUCIBILITY THROUGH CELL LINE AUTHENTICATION

are tested for mycoplasma at the time of freeze by PCR and Hoechst staining and/or microbiological methods.

Bloodborne pathogen testing

Coriell historically has tested all newly established cell cultures for the presence of HIV. A TaqMan real-time PCR assay is used for detecting HIV-1 provirus in DNA samples. Coriell currently offers evaluation of multiple bloodborne pathogens, including HIV-1, HIV-2 and HCV from cell line RNA, and HBV from cell line DNA utilizing real-time PCR kits from Primerdesign™ and genesig®, each of which is designed to detect a broad range of clinically relevant subtypes. These assays offer quantitative detection of bloodborne pathogens and provide copy number readouts for each sample tested.

CONCLUSION

Cell line integrity has come to the forefront of biomedical research as decades worth of research performed with misidentified or contaminated cell lines has been revealed. The critical need to authenticate cell lines has been highlighted and is now being implemented as a requirement for new funding and publications. As always, when using cell lines, best practices begin with controlling the cell culture environment and using good aseptic technique. Accidental cross-contamination can be better controlled by not sharing media between cell lines, processing only one cell line at a time in the cell culture hood, and by changing gloves frequently. Sterility can be monitored and better controlled with routine sterility checks of media and equipment, elimination of antibiotic use, and routine mycoplasma testing. Additionally, as cell lines are being propagated, morphology and growth characteristics should be monitored to track phenotypic drift of a cell population. These are best practices established and utilized by Coriell's biorepositories.

As the biomedical research community strives to decrease the use of misidentified or contaminated cell lines and to improve reproducibility of results, Coriell is excited to offer our gold standard suite of Cell Line Integrity

services to ensure cell line authenticity. Our Cell Line Integrity services address cell line identity through the use of STR profiling and sterility with mycoplasma, bacteria and bloodborne pathogen testing. These services provide researchers with solutions to ensure the highest quality results from their research.

For more information about Coriell's Cell Line Integrity services please visit <https://catalog.coriell.org/celllineintegrity> or contact Coriell's Customer Support Department:

Customer Support
Coriell Institute for Medical Research
403 Haddon Avenue, Camden, NJ 08103
[800-752-3805](tel:800-752-3805) | customerservice@coriell.org



INCREASING SCIENTIFIC REPRODUCIBILITY THROUGH CELL LINE AUTHENTICATION

REFERENCES

1. Yu, M, *et al.* **2015**. A resource for cell line authentication, annotation and quality control. *Nature*. 520: 307-311.
2. Freedman, LP, *et al.* **2015**. Reproducibility: Changing the policies and culture of cell line authentication. *Nature Methods*. 12: 493-497.
3. Liang-Chu, MM, *et al.* **2015**. Human biosample authentication using the high-throughput, cost-effective snprtrace(tm) system. *PLoS One*. 10: e0116218.
4. Gartler, SM. **1967**. Genetic markers as tracers in cell culture. *National Cancer Institute Monograph*. 26: 167-195.
5. Liscovitch, M and Ravid, D. **2007**. A case study in misidentification of cancer cell lines: Mcf-7/adrr cells (re-designated nci/adrr-res) are derived from ovc8 human ovarian carcinoma cells. *Cancer Letters*. 245: 350-352.
6. Lorenzi, PL, *et al.* **2009**. DNA fingerprinting of the nci-60 cell line panel. *Molecular Cancer Therapeutics*. 8: 713-724.
7. Chatterjee, R. **2007**. Cell biology. Cases of mistaken identity. *Science*. 315: 928-31.
8. Hughes, P, *et al.* **2007**. The costs of using unauthenticated, over-passaged cell lines: how much more data do we need? *Biotechniques*. 43: 575-86.
9. NIH. Notice number: NOT-OD-08-017. Re: Authentication of Cultured Cell Lines.
<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-017.html>
10. NIH. Notice number: NOT-OD-15-103. Re: Enhancing Reproducibility through Rigor and Transparency.
<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-103.html>
11. Nature Publishing Group. **2015**. Availability of data, material and methods.
<http://www.nature.com/authors/policies/availability.html>
12. Alston-Roberts, C, *et al.* **2011**. Authentication of Human Cell Lines: Standardization of STR Profiling. 499-501.
13. Olarerin-George, AO and Hogenesch, J B. **2015**. Assessing the prevalence of mycoplasma contamination in cell culture via a survey of NCBI's RNA-seq archive. *Nucleic Acids Research*. 43(5): 2535-2542.
14. Uphoff, CC and Drexler, HG. **2002**. Detection of mycoplasma in leukemialymphoma cell lines using polymerase chain reaction. *Leukemia*. 16: 289-293.
15. Rottem, S and Barile, MF. **1993**. Beware of mycoplasmas. *Trends in biotechnology*. 11(4): 143-151.
16. Ryan, JA. **1994**. *Understanding and managing cell culture contamination*. Corning Technical Bulletin.
<http://catalog2.corning.com/Lifesciences/media/pdf/cccontamination.pdf>