Doing Good Science: Authenticating Cell Line Identity

Jill Harley Dunham[1], Pam Guthmiller[2], [1]Ph.D. Candidate Department of Pharmacology Emory University, Atlanta, GA, [2]Promega Corporation

INTRODUCTION

The ability to repeat previously published results in scientific research is important for confirming discoveries. If data is not repeatable, it is doubtful that anyone will continue to pursue that work. Much of biomedical research is done with cultured cells obtained from repositories such as American Type Culture Collection (ATCC) or from fellow researchers. An estimated 15–20% of the time, cells used in experiments have been misidentified or cross-contaminated with another cell line (1–3). ATCC, along with the Coriell Institute for Medical Research, European Collection of Cell Cultures (ECACC), Deutsche Sammlung von Mikrorganismen and Zellkulturen (DSMZ), and the Japanese Collection of Research Resources, all have received cell line submissions that, upon authentication, were determined to have been misidentified by the depositor (4,5). In addition, stem cell lines can become contaminated by the mouse cells making up the feeder cell layers on which stem cells are propagated.

Clearly contamination and misidentification can pose a huge threat to the quality of publications and legitimacy of research findings produced from any of these cell cultures. For this reason, many repositories now authenticate cell line submissions and monitor cross-contamination.

HISTORICAL PERSPECTIVE ON THE PROBLEM

The first human cancer cell line, HeLa, was developed in 1952, and for the next 15 years, many more human cell lines from different tissues were developed (6). In 1968, researchers discovered that many cultured cells exhibited characteristics that did not match the characteristics of the original source. This was the first evidence that particular methods of culturing cells could produce unpredictable changes to the cells. While the development of better techniques helped decrease cross-contamination, few tests were available to determine which cells were already affected. Since then, however, standardized methods have been developed that are both quick

Like this cartoon?
Check out Tech 'Toons
www.promega.com/pnotes/techtoons/

Everything was going along fine until they discovered their HeLa cell line expressed Y chromosome markers.

The Importance of cell line authentication has been recognized already. Recently, *Nature* mandated STR fingerprint data for papers reporting new human embryonic stem cell lines but not other lines.

and inexpensive to perform. While these improvements should have eliminated much cellular cross-contamination, it remains a prominent issue. ATCC and other similar repositories now monitor cross-contamination and authenticate all cell lines that they distribute, but most individual investigators do not adhere to the same meticulous authentication processes. In fact, a 2004 survey of approximately 500 biologists by Gertrude Buehring of the University of California, Berkeley, and her colleagues showed that less than half of all researchers regularly verify the identities of their cell lines using any standard techniques, such as DNA fingerprinting by short tandem repeat (STR) analysis (6). Without requiring that all cell lines be authenticated, misidentification will remain a significant problem.

AUTHENTICATION SAVES TIME AND MONEY

Aside from the issue of inconsistent or questionable data, cross-contamination also wastes time and money. For instance, Mordechai Liscovitch, a cancer researcher in Israel, says that he and his lab spent three years working on two breast cancer cell lines (MCF-7 and MCF-7/AdrR, now renamed NCI/ADR-RES) that they believed were related, only to discover later that the cell lines were actually unrelated. Although some researchers suspected the misidentification of these cell lines as early as 1998, the actual identity of these lines was not confirmed until recently (4,7). NCI/ADR-RES is not derived from the breast cancer cell line MCF-7 as originally thought but rather from the ovarian cancer cell line OVCAR-8 (7). All three of these cell lines are part of the NCI-60 panel of cells that are routinely used in drug-screening applications (4).

The Liscovitch lab cancelled the publication of a manuscript that contained erroneous conclusions based on the mistaken cell line identity; however, an unknown number of studies have been published containing conclusions based on misidentified cell lines. Charles Patrick Reynolds of the University of Southern California and the Children's Hospital Los Angeles' Institute for Pediatric Clinical Research estimates, up to 35–40% of previously published cell biology papers would need to be retracted due to invalid data. These estimates have caused Roland Nardone at The Catholic University of America to call for authentication both as a condition for receipt of grant funds from major agencies such as the National Institutes of Health and American Cancer Society as well as for publication of cell culture-based research in leading journals. He has also requested education for technicians and scientists about prevention and detection of cross-contamination. He has worked to gain professional societies' endorsements of these proposed policies and sponsored conferences and workshops to facilitate adoption of the standards. Dr. Nardone's belief that these changes are vital to scientific research led him to cocreate the Cell Line Authentication Global Awareness Month [May 2008], started by "an ad hoc group of scientists because of the chaos and waste caused by rampant misidentification and cross-contamination of cell lines."

The problems of cross-contamination have been recognized by organizations, such as the ATCC (7) and FDA, which require that in-process materials, such as cell lines, that are used to produce pharmaceuticals be tested for identity and purity (8). Similarly, *Nature* recently mandated STR fingerprint data for papers reporting new human embryonic stem cell lines but not other lines (4). Although journals and funding agencies recognize the problems of cell line misidentification, they are uncertain how best to address them. When would researchers be asked to confirm cell line identity, before or after peer review? Where would the resources come from to confirm authors' assertions of cell line identity?

In addition to wasted time and money, inconsistent or nonreproducible findings, and retraction of publications, a potential for health consequences also exists in misidentified cell lines. Drugs, vaccines and other biomedicines all are created based on findings in the lab, often times via cell culture. Products made using misleading or false data can cause major delays in the production and availability of treatments for a variety of diseases. The longer it takes for a treatment to be developed, the more people these diseases affect.

STR-BASED METHODS PROVIDE EASY, QUICK CELL LINE IDENTIFICATION

All of these issues can be prevented with inexpensive and now standard procedures used to authenticate cell lines, but the procedures must be performed. Roderick MacLeod and his colleagues at DSMZ, German Collection of Microorganisms and Cell Cultures, have found that about 90% of scientists ignore or refuse a cell bank's request for new lines, preventing the establishment of cell line DNA fingerprints for future attempts at verification (5,7). Researchers need to be educated on how to detect intra- and inter-species cross-contamination, as well as why it is so important to do so.

Many methods, such as isoenzyme analysis, karyotyping, human lymphocyte antigen (HLA) typing, and characterization of amplified fragment length polymorphisms

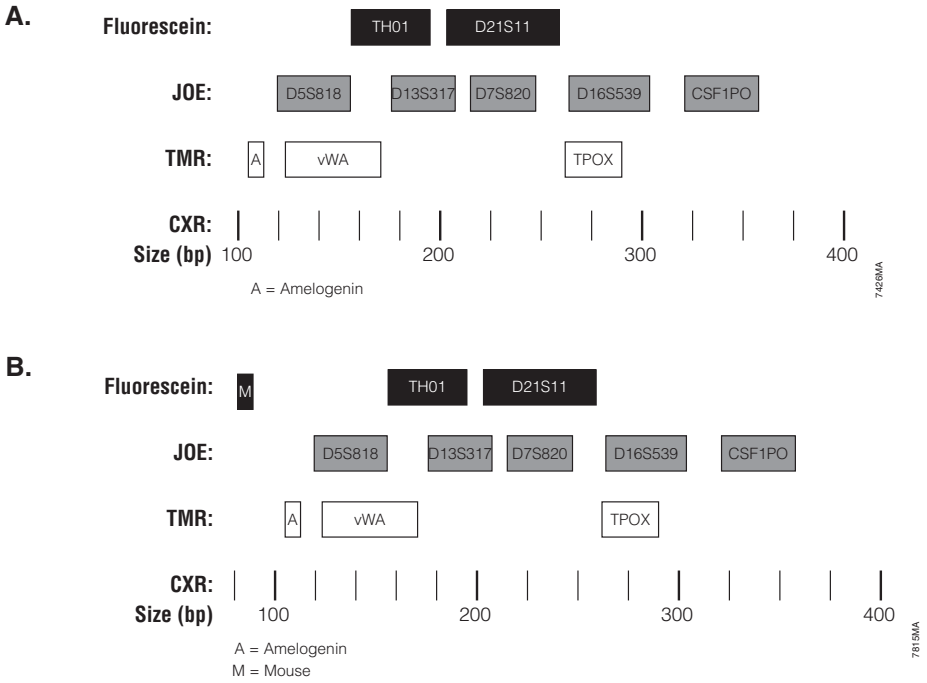

Figure 1. Allele ranges for the StemElite™ ID and Cell ID™ Systems. Panel A. STR fragments amplified by the Cell ID™ System. **Panel B.** Fragments amplified by the StemElite™ ID System. In both systems, fragments are labeled with different dyes and are separated by capillary electrophoresis based on size. A size standard is included in each sample to determine the size of the fragments. Fluorescein-labeled loci are shown in black. JOE-labeled loci are shown in gray. TMR-labeled loci are shown in white. The CXR-labeled Internal Lane Standard 600 fragments are represented by black bars.

Authentication with confidence

Cell ID™ System uses nine STR loci and Amelogenin for gender identification, resulting in a genetic profile with a random match probability of 1 in 2.92×10^9.

StemElite™ ID System uses nine STR loci, Amelogenin for gender identification and one mouse locus, resulting in a random match probability of 1 in 2.92×10^9 and simultaneously detecting as little as 1% mouse contamination in a human cell line.

(AFLP), have been used to identify cross-contamination in cell culture. However, a superior method is STR-profiling, well established in the field of DNA-based forensic identification. At the ATCC, STR analysis is performed using multiplex PCR (Promega PowerPlex® 1.2 System) to simultaneously amplify eight STR loci plus Amelogenin for gender determination (9,10). A unique pattern of repeating DNA is generated for each human cell line analyzed, so the DNA profile of each new stock is verified by comparing to the baseline profile. STR profiling already has prevented the further distribution of six different cell lines at ATCC after it revealed the presence of Y chromosome-specific amplification products in cell lines that are derived from females. The research community hopes that STR profiling will provide a global reference technique to detect and eliminate cell line contamination.

STEMELITE™ ID AND CELL ID™ SYSTEMS ALLOW STR-BASED AUTHENTICATION OF CELL LINES

Promega is a leader in providing STR-profiling systems for forensic and paternity applications, and the PowerPlex® 1.2 STR analysis system has become the "gold standard" tool used by cell culture facilities to authenticate cell lines. The StemElite™ ID System[a–d] and Cell ID™ System[a–d] offer simpler methods for authenticating cell lines. These improved systems include the reagents required to successfully and simply identify and authenticate human cell lines as well as detect intra-species cell line cross-contamination.

The Cell ID™ System uses STR analysis of specific, highly polymorphic loci in the human genome through simultaneous amplification and three-color detection of ten loci (Figure 1, Panel A; nine STR loci and Amelogenin for gender identification; including D21S11, THO1, TPOX, vWA, Amelogenin, CSF1PO, D16S539, D7S820, D13S317 and D5S818). These loci collectively provide a genetic profile with a random match probability of 1 in 2.92×10^9.

The StemElite™ ID System allows co-amplification and three-color detection of ten human loci (Figure 1, Panel B; one mouse locus, nine STR loci and Amelogenin for gender identification; including D21S11, TH01, TPOX, vWA, Amelogenin, CSF1PO, D16S539, D7S820, D13S317, D5S818). These loci collectively provide a genetic profile with a random match probability of 1 in 2.92×10^9 while the mouse primers simultaneously detect a 1% fraction of mouse contaminant in a human cell line.

The systems include a hot-start *Taq* DNA polymerase for convenient room temperature reaction assembly. Following amplification, samples are analyzed by capillary electrophoresis in a single injection in conjunction with

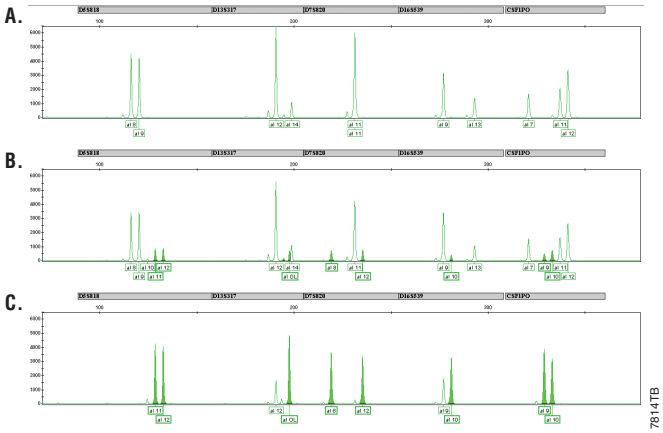

Figure 2. Determination of cell line contamination using the Cell ID™ System. Panel A. HEK293 cell line STR profile. **Panel B.** STR profile of HEK293 Cell Line with 29% HeLa cell line contamination. **Panel C.** HeLa cell line STR profile. DNA was extracted from 10^4 cells using the Maxwell® 16 Cell LEV DNA Purification Kit, then amplified with the Cell ID™ System. Amplified products were detected on a capillary electrophoresis instrument. For simplicity, only the JOE-labeled allele profiles are shown.

provided standards to assist in determining allele sizes for the different loci. In Figure 2, the Cell ID™ System was used to detect contamination of a cell line. The genetic profile is determined using allele-calling software.

Most research scientists have access to institute core facilities and service companies that have the instrumentation (capillary electrophoresis), software and experience to perform cell line profiling, even if they do not have the same capability within their own labs. General recommendations in the literature suggest that investigators authenticate an early passage (first week of culture) of their cells to establish the identity of the cell line. Cells should be authenticated again before freezing, once every two months that the culture is actively growing, and before publication. If a lab is using more than one cell line, all lines should be tested initially to rule out cross-contamination (8).

SUMMARY

Because of the importance of cell culture to biomedical research and technology, proper cell line authentication is in everyone's best interest. However, cross-contamination continues to be a problem. With the increasing number of new cell lines and the high rate of cell culture use in labs worldwide, significant gaps have been created in basic principles of quality control (i.e., cell line authentication). From research articles published with misidentified cell lines and resultant questionable results, to stem cell lines and other lines destined for clinical uses, cross-contamination affects science in all realms—from lab bench to clinic. Without significant change to the handling and treatment of cell cultures, it will become only a larger and more serious issue.

REFERENCES

1. Drexler, H.G., Dirks, W.G. and MacLeod, R.A.F. (1999) *Leukemia* **13**, 1601–7.
2. Drexler, H.G. *et al.* (2001) *Blood* **98**, 3495–6.
3. Cabrera, C.M. *et al.* (2006) *Cytotechnology* **51**, 45–50.
4. Chatterjee, R. (2007) *Science* **315**, 928–31.
5. MacLeod, R.A.F. *et al.* (1999) *Int. J. Cancer* **83**, 555–63.
6. Buehring, G.C., Eby, E.A. and Eby, M.J. (2004) *In Vitro Cell. Dev. Biol.* **40**, 211–5.
7. Liscovitch, M. and Ravid, D. (2007) *Cancer Lett.* **245**, 350–2.
8. FDA. General Requirements for Laboratory Controls. 21 CFR 211.160 and 21 CFR 610.18.
9. *ATCC Connection Newsletter* (2000) **21**, 1–2.
10. Masters, J.R. *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98**, 8012–7.

PROTOCOLS

- *Cell ID™ System Technical Manual* #TM074, Promega Corporation
  www.promega.com/tbs/tm074/tm074.html
- *StemElite™ ID System Technical Manual* #TM307, Promega Corporation
  www.promega.com/tbs/tm307/tm307.html

ORDERING INFORMATION

Product	Size	Cat.#
Cell ID™ System	50 reactions	G9500
StemElite™ ID System	50 reactions	G9530

For Research Use Only. Not for use in diagnostic procedures.

[a]U.S. Pat. Nos. 5,843,660 and 6,221,598, Australian Pat. No. 724531, Canadian Pat. No. 2,118,048 and other patents and patents pending.

[b]STR loci are the subject of U.S. Pat. No. RE 37,984, German Pat. No. DE 38 34 636 C2 and other patents issued to the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, e.V., Germany. The development and use of STR loci are covered by U.S. Pat. No. 5,364,759, Australian Pat. No. 670231 and other pending patents assigned to Baylor College of Medicine, Houston, Texas.

Patents for the foundational PCR process, European Pat. Nos. 201,184 and 200,362, expired on March 28, 2006. In the U.S., the patents covering the foundational PCR process expired on March 29, 2005.

[c]This product is sold under licensing arrangements with the USB Corporation for Forensic and Genetic Identity Applications Fields specifically excluding tissue typing related to transplantation or other medical procedures. Further licensing information may be obtained by contacting the USB Corporation, 26111 Miles Road, Cleveland, OH 44128.

[d]This product is sold under licensing arrangements with Stratagene for Forensic and Genetic Identity Applications Fields specifically excluding tissue typing related to transplantation or other medical procedures. Further licensing information may be obtained by contacting the Business Development Department, Stratagene California, 11011 North Torrey Pines Road, La Jolla, CA 92037.

Products may be covered by pending or issued patents or may have certain limitations. Please visit our Web site for more information.